



universität
wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„End-to-end Network Quality:
A Utility-based Approach“

verfasst von / submitted by

Dipl.-Ing. Dipl.-Ing. Patrick Zwickl, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Doktor der Wirtschaftswissenschaften (Dr.rer.oec.)

Wien, 2016 / Vienna 2016

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on the student
record sheet:

A 796 305 175

Dissertationsgebiet lt. Studienblatt /
field of study as it appears on the student record sheet:

Wirtschaftsinformatik

Betreut von / Supervisor:

Univ.-Prof. Dr. Peter Reichl, M.A. St

Erklärung zur Verfassung der Arbeit

DI DI Patrick Zwickl, BSc
Glockengasse 1/2/9, 1020 Vienna, Austria

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit, die anderen Werken oder dem Internet im Wortlaut, dem Sinn nach entnommen oder in Zusammenarbeit entstanden sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe. Diese Dissertation ist originär und wurde zuvor weder als Ganzes noch als Teil zur Erlangung eines universitären Abschlusses verwendet.

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text.

(Ort, Datum)

(Unterschrift)

Acknowledgements

Dedication.

This thesis is dedicated to my family, especially my wife, and friends.

Acknowledgments.

I would like to thank the members of my Ph.D. committee, Heikki Hämmäinen, Burkhard Stiller and especially my supervisor Peter Reichl. The pages are numerous, the process leading to those pages has been extensive, and on these grounds their efforts cannot be valued high enough.

My research stay at the Aalto University has clearly been an enriching and highly productive experience. I would like to thank the members of the “Netbiz” team (Department of Communications and Networking, School of Electrical Engineering) around Heikki Hämmäinen for their great hospitality.

Further, I would like to thank my closest collaborators from the University of Vienna (especially from the Cooperative Systems Group), FTW Telecommunications Research Center Vienna (especially Ivan Gojmerac¹ and Andreas Sackl), VTT Technical Research Centre of Finland (especially Martin Varela and Toni Mäki), the University of Zagreb (especially Lea Skorin-Kapov and Ognjen Dobrijevic), and the numerous research projects I have participated (especially the FP7 project Economics and Technologies for Inter-Carrier Services (ETICS), the Celtic-Plus Quality of Experience Estimators in Networks (QuEEN) project and several COST Actions). Language and research are challenging, that is why I am grateful to everyone who has shared some advice.

Finally, in good old personal tradition, I thank the reader for attempting to read the unbearable number of highly condensed pages. I hope this thesis will be of service to you.

¹ Our collaboration started at FTW, where we shared an office, and has been retained since then.

Preface

Similar to Net Neutrality (NN), the introduction of Quality of Experience (QoE)-aware charging on the Internet involves a conflict of interest where around the efficient utilisation of resources at least societal, cultural, business, economic, technological and judicial needs meet (also see [WAR11] for inspiration). A scientific work in such an area can profit from a clearly communicated but purely personal contentual stance: From the beginning, the Internet has taken an outmost important societal role. The core idea of the Internet, primarily driven by an academic need, was to connect previously separated Autonomous Systems to form a global network of equally accessible information. Without a surprise, predominant technologies and charging practices for Internet services have also followed a Best Effort (BE) principle over the Internet Protocol (IP) where non-discriminatory access to information is predominant nowadays. Today, the Internet forms a prosperous and balanced “ecosystem”, providing means of “transport” for any kind of utilisation, including a manifold of business models and non-commercial or even charitable exchange. In analogy to biology, where ecosystems refer to a community of organisms, the Internet provides a space where commonly accepted Internet technologies provide the means for interoperability, cooperation and even collaboration of stakeholders—a characteristic that should not be eliminated by any structural change.

Bringing in systematic changes to a seemingly egalitarian Internet system that may not only lead to a classification of users and usages but also centrally integrates profitability considerations, may appear retrogressive and potentially societally dangerous (contradicting acquired societal values). Such a system change is, however, induced when transitioning from the quantity- and price-oriented BE Internet to a system where quality is co-equally marketed. This argumentation, however, cuts corners in the discussion, as both consumers and operators can be profiteers of a change, sometimes even simultaneously. One example is the trending concept of QoE where human perceptions are studied in order to optimise networks. While such optimisations can be solely developed around the needs of either consumers or telecom operators, this thesis aims at collecting the data that provides indications on how to maintain a balance between stakeholder interests also

in quality-aware networks of the future.

The structure of the Internet ecosystem, as it evolves over time, also requires a careful design and continuous moderation: while in early markets profit margins seem to have favoured telecom operators, the rise of Internet giants like Google plus the diminishing profits of operators have turned the coin². Following an outmost libertarian position, the per se rationality of stakeholders and self regulation of markets³ may automatically correct imbalances whenever they occur, ie, protect telecom operators from sustained unprofitability. While such a view is debatable and probably dubious⁴, I do not believe in charitable features of profit-oriented companies (whether in the unrealistic case of a full-deregulation with self-controlled profits or the full regulation where providers work for the sake of the Internet's health). I also believe in the protection of societal rights around a functioning ecosystem of mutual respect to one another—with respect for heterogeneous roles on the Internet and the common environment—, with tools supporting educated decisions, and neutrality and deregulation wherever it allows for. From my point of view, the openness of the Internet refers to the proper moderation of regulatory strength, where both over- and underregulation are unwise alike. Overregulation antagonises the business rationale of telecom operators, while underregulation endangers the protection of consumers, small firms and the stability of the industry. Openness also refers to the often appreciated freedom of the Internet and its societal role:

“ Freedom of connection with any application to any party is the fundamental social basis of the internet. And now, is the basis of the society built on the internet. ”
Tim Berners-Lee

However, maintaining a functioning ecosystem, especially in its most open form, is especially challenging at the complexity scale of the Internet:

“ The Internet is the first thing that humanity has built that humanity doesn't understand, the largest experiment in anarchy that we have ever had. ”
Eric Schmidt, Google Inc.

David Ulevitch has further stressed the importance of heterogeneity (such as diversity of services) on the Internet, which also raises the question what “fairness” or “equality” could mean. From a legal point of view, in most Western countries a *ius respicit aequitatem* exists, where the colloquial summary of Austria's corresponding laws with “*Gleiches*

² Today's Network Service Providers (NSPs) are still widely profitable. As long as the Return on Investment (RoI) justifies the investment risks, market modifications may not be inherently needed.

³ Many of the available definitions for self-regulation are summarised in [Ogu95]. Our criticism mainly addresses the form of “Unconstrained Market Competition”, which creates an environment with limited or no regulator involvement and difficult comparability of products.

⁴ “Dubious” results from similar beliefs that have led to a financial collapse as a result of almost unrestricted speculations with risks of subprime mortgages in the US market around 2008; further reading [Fox11; Sti10]

gleich, Ungleiches ungleich” (Art. 7 of the Bundes-Verfassungsgesetzes, Art. 2 of the Staatsgrundgesetz; translating to: *equals equally, unequals unequally*) gets to the heart of the equality problem: equal handling may require inequality. In the context of telecommunications networks, different services or users may have different needs (and perceptions), which would require an unequal handling to obtain the required equality. The transition from perception-oriented network quality metrics, ie, from QoE to Quality of Service (QoS) (see the definitions in the main part of this thesis), follows this line of thought, but at the same time takes the concept beyond purely “technical” definitions to a more eclectic equivalent that is much harder to convey to humans, especially when it involves heterogeneous stakeholders in an ecosystem.

“ I think anything which promotes heterogeneity on the Internet promotes stability. Diversity in services, service providers, and separating the layers of the networking stack are all important. ”

David Ulevitch, OpenDNS

Heterogeneity among consumers but also content providers is especially of interest when extending the desire for openness also to liberal discrimination practices for the Internet transport and its pricing. In combination with QoE-awareness such a discriminatory Internet enables a series of objective functions such as the maximisation of user experiences, profits of content provider or network operators (eg, by exploiting weaknesses of consumers), or the *social welfare*. The objective function is likely not chosen independently of today’s market structures and competitive situation, but the new capabilities to discriminate may amplify existing traits in the one way or another. However, not all strategies to cope with heterogeneity can maintain the balance in the Internet ecosystem, which is not in the interest of this thesis. Reinterpreting the definitions of M. Friedman [Fri07], it is not the technology to be blamed for the usage of technology, it is not the economy to be blamed for the assessment of alternative outcomes, but the political dimension that has to define the environments for the Internet today.

By concentrating on a more pluralistic viewpoint integrating seemingly antagonistic positions and needs, this thesis is, hence, intended to provide questions and tools to overcome the black- and whiteness of fundamental Internet rights discussions while also commercialising QoE. On that account, it works through various perspectives of demand (users, customers) and supply (operators) interests and systematically questions discrimination practices from a pricing and structural point of view. One of the underlying questions, providing a consistent thread through the dissertation, is the essence of any kind of discrimination practice: *Does it pay off to discriminate contents, applications, users or anything else on the Internet?* Discrimination is an inherent concept when considering network quality differentiation, especially when focusing on the trending QoE research field and the burgeoning discrimination practices in the Internet business, eg, side payments as direct compensation payment among stakeholders of the Internet ecosystem. The integrated

view of supply and demand in a more holistic multi-method and utility-oriented analysis, thus, seems unavoidable, which is addressed in this thesis.

Abstract

In the prosperous days of both cellular and fixed line telecommunications, the introduction of flat rate pricing boosted profits due to the increased utilisation of available capacity. The growth of the customer base delivered the required revenue rise, which justified continuous infrastructure investments. Since then, the entire telecommunications industry has been undergoing a drastic change. Today, it faces an environment where profit margins are low, the customer basis stagnates and flat price curves narrow the scope for recovery strategies. The increase in bundle prices is difficult to communicate to consumers, especially when the level of competition remains high. Thus, especially if new usage paradigms like Machine-to-Machine (M2M) or Internet of Things (IoT) do not meet the high expectations of the industry, the entire industry needs to reinvent itself economically.

When following a modern inter-disciplinary understanding of telecommunications, as in [LKC12; Jai06], the Internet can be treated as a sort of experience good [Nel70]. Experience goods, contrary to objectively assessable goods, need to be perceived to be evaluated properly. Thus, the effective conveyance of Quality of Experience (QoE) products, as subjective representation of network quality, paves the way for the co-equal marketing of quality and price and associated business prospects. However, existing approaches for quality-oriented networks lack such clear communication strategies (cf. [Var+15]) apart from empirically-backed marketisation frameworks and technologies. The vital QoE research field has started to map technical network aspects (ie, Quality of Service (QoS)) to a user perspective (ie, QoE), but it has failed to tackle a successful introduction and establishment of QoE as a product on the market.

In response to these shortfalls, the present thesis presents a utility-centric solution for marketing end-to-end quality-assured network services that equally balances user, social, economic and technological viewpoints across various access & core networks. It will, therefore, clarify in what kind of quality-assured network services customers are interested, how much they are willing to pay and how this information can be utilised economically and technically. Thereby, quality guarantees provide certainty as an orthogonal quality aspect to network or service parameters. Special attention is devoted to understanding the proper parameterisation and configuration of successful market entrances of “premium”

network service offers.

By setting a focus on Video on Demand (VoD) and other media-rich services, a representative cross-section of the enormous problem space is addressed. This work traverses the problem field in a top-down manner from high-level economic over medium-level optimisation-oriented towards more fine-grained technological contribution layers. On each of these three contribution layers tailored methodologies are used: around known market fundamentals, a sustainable market configuration is required, which is addressed by studying various Value Network (VN) configurations on the upper layer. On the middle layer, empirical QoE and Willingness-to-Pay (WTP) studies are conducted in order to run empirically-backed analytical and game-theoretic optimisation models. Those models can maximise revenues, profits or social welfare aspects. On the lowest layer, a link to technological frameworks is created which accommodates the end-to-end perspective by studying the economically and technologically efficient assignment of network resources along the transmission path. For this purpose, Service Level Agreement (SLA) negotiations, path computation, access control and network scheduling techniques are used. In a final step, an evaluation of central input data is applied in order to reflect on the general applicability of results.

The results of this thesis indicate that third-degree price discrimination can yield substantial revenue growth potential, especially when an appropriate price is immediately set at the market entrance. User, NSP and/or social welfare gains can be generated when migrating to utility-aware access control, resource assignment and network path computation systems. Side payments, which break aspects of NN, largely appear economically unattractive and Transit NSPs have to face their low market powers. A generalisation of results to other cases in telecommunications seems likely, but deserves further attention. The use of developed empirical or Value Network (VN) toolsets alike can be used outside the context of communications services.

Kurzfassung

In den glorreichen Tagen der Telekommunikation führte die Einführung von Flatrates zu einer Profitsteigerung durch eine erhöhte Netzauslastung, welche umfassende Investitionen in die Netzwerkinfrastruktur ermöglichte. Doch der drastische Wandel der letzten Jahre ließ die Gewinnmargen substantiell schmelzen, was teilweise durch stagnierende Nutzerzahlen und starken Wettbewerb bedingt war. Da bei starkem Wettbewerb direkte Preiserhöhungen kaum ohne Marktanteilsverluste umsetzbar sind, limitierte die frühe Fixierung auf flache Preiskurven (ie, mittels Flatrates) auch den Veränderungsspielraum für Netzbetreiber. Falls neue Nutzungsszenarien, wie Machine-to-Machine (M2M) oder Internet of Things (IoT), keinen Umschwung bringen, muss sich die Industrie grundlegend neu erfinden.

Mit einem modernen, interdisziplinären Verständnis von Telekommunikation [LKC12; Jai06] kann das Internet als Erfahrungsgut [Nel70] verstanden werden. Derartige Güter können im Gegensatz zu objektiv quantifizierbaren Produkten nur subjektiv und basierend auf eigenen Erfahrungen beurteilt werden, was auch die Vermarktung von Quality of Experience (QoE) erheblich erschwert (cf. [Var+15]). Sowohl eine geeignete Produktkommunikation und Vermarktung als auch empirisch-parametrisierte Technologien legen jedoch den Grundstein um Netzwerkqualität auf den gleichen Rang wie den Paketpreis zu heben und damit eine nachhaltigere Vermarktung von Internetdiensten zu erwirken. Während das Forschungsfeld um QoE in den letzten Jahren die Verknüpfung von netzwerktechnischen (QoS) und nutzergetriebenen (QoE) Aspekten angestrebt hatte, konnten keine hinreichenden Lösungen für die kommerzielle Etablierung von QoE-Produkten am Markt präsentiert werden.

Als Reaktion auf die genannten Probleme verfolgt die vorliegende Dissertation einen nutzwertorientierten Lösungsansatz um Ende-zu-Ende-Netzwerkdienste mit Qualitätsgarantie zu kommerzialisieren, unter gleichwertiger Berücksichtigung nutzer-, sozial-, ökonomisch- und technisch-getriebener Gesichtspunkte. Sie erforscht insbesondere attraktive Marktsegmente für garantierte Netzwerkqualität, deren ökonomische Verwertbarkeit und Möglichkeiten für deren technische Realisierung. Während die verwendeten Garantien dabei Stabilität orthogonal zu anderen Qualitätsmerkmalen vermitteln, sorgt die Ende-zu-Ende-Eigenschaft für ein konsistentes Kundenerlebnis ungeachtet der globalverteilten jedoch kooperierenden Netzwerkanbieter. Be-

sondere Aufmerksamkeit wird auf die Parametrisierung und Konfiguration eines erfolgreichen Markteinstiegs gelegt.

Durch einen gezielten Fokus auf Video on Demand (VoD) und andere medienintensive Dienste kann diese Arbeit einen repräsentativen Querschnitt durch die umfassende Materie zeigen. In einem Top-down-Ansatz wird ein Lösungskonzept über drei Ebenen von grundlegend ökonomischen über optimierungsorientierten bis zu feingranularen technologischen Aspekten konstruiert. Auf jeder Ebene werden Methodologien des jeweiligen Fachgebietes verwendet: Auf der obersten Ebene werden fundamentale Marktkonfigurationen als Value Network (VN) verglichen. Die mittlere Ebene fokussiert auf empirische Nutzerstudien zu QoE und Willingness-to-Pay (WTP), um ökonomische Optimierungsmodellen realistisch zu parametrisieren. Diese Modelle ermöglichen die Maximierung von Umsätzen, Profiten oder der sozialen Wohlfahrt. Dazu werden empirische Laborstudien sowie analytische und spieltheoretische Modelle verwendet. Auf der untersten Ebene wird eine Anbindung an technische Lösungen geschaffen, die sowohl ökonomisch als auch technisch eine effiziente Ende-zu-Ende-Verteilung von Ressourcen ermöglichen. Dabei werden unter anderem Techniken zum Aushandeln von Service Level Agreements (SLAs), zur Netzwerkpfadberechnungen, für Zugangskontrollen und des Netzwerkschulings verwendet. In einem finalen Schritt werden Daten, die eine zentrale Rolle in dieser Dissertation einnehmen, in einer dezidierten Studie evaluiert.

Die Resultate dieser Arbeit deuten darauf hin, dass Preisdiskriminierung dritten Grades ein substantielles Umsatzwachstum ermöglichen kann, sofern ein geeigneter Preis bei Markteintritt festgelegt wird. Durch den Umstieg auf nutzenorientierte Zugangskontrolle, Ressourcenzuteilung und Netzwerkpfadberechnung können Vorteile für Nutzer, Anbieter oder die soziale Wohlfahrt geschaffen werden. Darüber hinaus erscheinen Kompensationszahlungen (ie, “side payments”), die mit NN im Konflikt stehen, aus ökonomischer Sicht weitestgehend unattraktiv. Insbesondere auffällig ist auch die niedrige Marktmacht von Transitnetzwerkanbietern. Die Übertragung dieser Ergebnisse auf andere Bereiche der Telekommunikation erscheint realistisch, bedarf jedoch einer weiteren detaillierten Untersuchung. Die empirischen Daten sowie das VN-Toolset können auch für Anwendungsgebiete abseits von Kommunikationsnetzen angewandt werden.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Problem	3
1.3	Methodology & Central Contributions	4
1.4	Solution Approach	6
1.5	Related Theses	8
1.6	Structure	10
2	Network Quality Markets	13
2.1	Background & Related Work	14
2.1.1	Definitions	14
2.1.1.1	From Graphs to Communication Networks	14
2.1.1.2	Quality of Service (QoS)	15
2.1.1.3	Quality of Experience (QoE)	18
2.1.1.4	Usability, Quality of Perception (QoP), Quality of Design (QoD)	21
2.1.1.5	Willingness-to-Pay (WTP)	22
2.1.2	Human Factors	24
2.1.3	Market Model	25
2.1.3.1	Objectives	25
2.1.3.2	Fundamentals	25
2.1.3.3	Consumers' Quality Surplus	26
2.1.4	From Value Chains to Value Networks	27
2.2	Ecosystem	29
2.2.1	Economic Dimensions	31
2.2.2	The QoE Value Network Flavours	33
2.2.2.1	Application-centric Pricing (cf. Fig. 2.3)	33
2.2.2.2	Access network-only (eg, local synchronisation)	34
2.3	Qualitative Analysis	35

2.4	Quantitative Analysis	38
2.4.1	Directions	38
2.4.2	Representation	40
2.4.2.1	Value Network Dependency Model (VNDM)	41
2.4.2.2	Value Network Graph (VNG)	43
2.4.2.3	Atomic Graph Operations	44
2.4.2.4	Eclectic Graph Operations	47
2.4.3	Value Network Quantification (VNQ) Requirements & Assumptions . .	48
2.4.4	VNQ Approach	50
2.4.5	Dependency Indicators	52
2.4.5.1	Quality	52
2.4.5.2	Scale	55
2.4.5.3	Risks	55
2.4.6	Revenue Sharing	59
2.4.6.1	Problem	59
2.4.6.2	Solution Overview	61
2.4.6.3	Calculation	62
2.4.7	Implementation & Evaluation	63
2.4.7.1	Tool Development	63
2.4.7.2	Sensitivity Analysis	64
2.5	The Video on Demand Value Network	65
2.5.1	Qualitative Assessment	66
2.5.2	Quantitative Assessment	73
2.6	Concluding Remarks	76
3	Demand & Spending	79
3.1	Internet Pricing & Customer Satisfaction (CS)	80
3.2	Does QoE Pricing Pay?	82
3.2.1	Model	83
3.2.2	Numeric Simulation	86
3.3	WTP & Demand	87
3.3.1	Related Work	87
3.3.2	2011 Trial	89
3.3.2.1	Demography	89
3.3.2.2	Test Design	89
3.3.2.3	Test Routine & Technical Setup	90
3.3.2.4	Results	91
3.3.3	2012 Trial	94
3.3.3.1	Demography	94
3.3.3.2	Test Design	94
3.3.3.3	Test Routine & Technical Setup	95

3.3.3.4	Results	97
3.4	Quality of Experience (QoE)	99
3.5	Motives	101
3.5.1	Interaction Behaviour	101
3.5.2	User Classification	107
3.6	Market Entrance Pricing	109
3.6.1	Monopolistic Market: A User Perspective	109
3.6.2	Market Challenges	112
3.6.3	From Monopoly to Competition	113
3.7	Evaluation & Concluding Remarks	114
4	Economics of Supply	119
4.1	Market Segmentation	120
4.1.1	Demand & WTP	120
4.1.2	Customer Segments	128
4.1.3	Generalisation	129
4.2	Single-class Service Provisioning	132
4.2.1	Base Model	132
4.2.2	Extended Model	135
4.3	Multi-class Service Provisioning	136
4.3.1	Background	137
4.3.2	Model	138
4.3.3	Results	140
4.4	Offloading Network Traffic	141
4.4.1	Model	142
4.4.2	Intra-provider Equilibrium	144
4.4.3	Inter-provider Equilibrium	145
4.4.4	Interpretation	146
4.5	Fairness & Regulation	146
4.5.1	Club Goods	147
4.5.2	Per-user Fairness & QoE	148
4.5.3	Net Neutrality (NN)	150
4.5.3.1	European Union (EU)	152
4.5.3.2	North America	154
4.5.4	Side Payments: Structural Market Power	157
4.5.4.1	Overview.	157
4.5.4.2	Quantification.	158
4.5.5	Side Payments: Game Theory	160
4.5.5.1	Effectivity.	161
4.6	Concluding Remarks	171

5	QoE & WTP Technologies	173
5.1	Background	174
5.1.1	Integrated Services Architecture (Intserv)	174
5.1.2	Differentiated Services Architecture (Diffserv)	174
5.1.3	Border Gateway Protocol (BGP)	175
5.1.4	IP exchange (IPX)	176
5.1.5	IP Multimedia Subsystem (IMS)	177
5.1.6	Connection-oriented Technologies	178
5.1.7	Energy- and Cost-aware Routing	179
5.1.8	QoE-aware Routing, Access Control and Resource Assignment	179
5.1.8.1	QoE-driven Path Optimization Model for Multimedia Services (Q-POINT)	180
5.1.8.2	ADmission control and resource Allocation for adaPtive mul-TImedia SErvices (ADAPTISE) and Long Term Evolution (LTE) Simulator	180
5.2	Inter-Domain Network Quality	181
5.3	Controlled Service Degradation: NSP and User Utilities	183
5.3.1	Approximation Model	185
5.3.1.1	Stage 1 (S1): Subjective QoE Data Without Price	186
5.3.1.2	Stage 2 (S2): User Utility	187
5.3.1.3	Stage 3 (S3): NSP Utility	188
5.3.1.4	Stage 4 (S4): Individualisation (Optional)	190
5.3.2	Case Study	191
5.3.2.1	Stage 1 (S1): Subjective QoE Data Without Price	191
5.3.2.2	Stage 2 (S2): User Utility	192
5.3.2.3	Stage 3 (S3): NSP Utility	192
5.3.2.4	Practical Implementation & Operational Use	195
5.3.3	Transmission Path	195
5.3.3.1	Parameterisation and Proof	198
5.3.3.2	Integration in QoE-driven Path Optimization Model for Multimedia Services (Q-POINT)	201
5.3.4	Access Control & Resource Assignment	202
5.4	Network Scheduler	202
5.4.1	Scheduling Techniques	203
5.4.1.1	Features	203
5.4.1.2	Capacity Dimensioning	205
5.4.2	Solution Approach	206
5.4.2.1	Scheduler Parameterisation	206
5.4.2.2	Generic Scheduler Design	208
5.4.2.3	Link to Policy Frameworks	210
5.5	Concluding Remarks	212

6	Evaluation	215
6.1	Demography	215
6.2	Technical Setup	216
6.3	Test Routine	218
6.4	Results	219
6.4.1	WTP	220
6.4.2	QoE	223
6.5	Recommendations	225
6.6	International Cross-Validation: Oulu Trial	227
7	Summary & Conclusions	229
7.1	Summary	229
7.2	Interpretation	230
7.3	Outlook & Future Work	233
A	Value Network Quantification (VNQ)	235
A.1	Inequality Measures	235
A.1.1	Concepts	236
A.1.1.1	Features & Calculation	236
A.1.1.2	Drawbacks	238
A.1.1.3	Voting Power	238
A.1.1.4	Interpretation	239
A.1.2	Lessons Learned	239
A.1.2.1	Decomposition	239
A.1.2.2	Overweighting	241
B	Net Neutrality (NN)	243
C	QoE and WTP	245
D	Utility	247
D.1	Numeric Results	247
D.2	Figures	249
	List of Conclusions	251
	Nomenclature	259
	Glossary	261
	Acronyms	263
	Own Publications	269

Standards	273
Bibliography	277

INTRODUCTION

1.1 Motivation

Communication is one of humanity's key intellectual and productivity advantages, undergoing a continuous technological change in recent years. By far the most noticeable shift has been the emergence of the Internet, which has fundamentally globalised human communication with a strong impact on daily information exchanges in almost all circumstances. Today's communications networks are heterogeneous, interconnected with each other¹ and generally highly complex, but are capable of providing low cost connectivity to almost all humans on earth. The high investment costs in communications infrastructures have been cushioned by promising prospects for a long time irrespective of unsustainable market approaches (which, at least partially, have also turned to practice, ie, \$127 billion in revenue by AT&T in 2012²). However, the entire telecommunications industry has been undergoing a drastic change during this time, with revenues and profits continuously falling³. Paired with soaring traffic demands due to multimedia services [Cis12] and associated continuous investments in network upgrades like LTE rollouts or fibre to the home, a highly pressured market has been rendered for NSPs—see [OFC11].

When scrutinising the potential causes explaining the loss of prospects, charging practices deserve a focal point. Today, flat rate pricing [Tuf04], ie, (almost) flat price curves, represent the predominant charging scheme for Internet services, which is a result of historic optimisations. In the early days of the Internet, the short term RoI for the enormous network infrastructure investments, ie, Capital Expenditure (CAPEX), called for quick economic utilisation of the available capacity. For this purpose, NSPs enticed customers to use Internet services intensively and to purchase bundles with higher access prices⁴ by

¹ Internet users expect to be able to reach anyone else on the Internet, irrespective of the involved NSPs, countries or devices.

² http://www.att.com/Investor/ATT_Annual/2012/downloads/ar2012_annual_report.pdf, last accessed: 20th May 2016

³ "Tellabs End of Profit": http://www.tellabs.com/markets/tlab_end-of-profit_study.pdf, last accessed: Oct 17, 2013

⁴ Steep price curves are constructed around quantities, keeping traffic demands low, but also shrinking the market.

keeping cost curves flat and, thus, being easily predictable for customers. This process is today colloquially referred to as “filling the pipe” strategy, which represents a low quality, low unit cost mass-market strategy. While this strategy limits action space for the per-user revenue growth potential, ie, Average Revenue Per User (ARPU) growth, to a manipulation of bundle prices, the per-user traffic volume has been soaring with the emergence of more services. In early markets, the growth of the **customer base** has compensated for the high investment costs which, however, does not apply to more mature markets any more, ie, the Internet markets seem to have reached their saturation. This is indicated, for example, by a cell phone subscription penetration of above 130% in European countries such as Austria or the United Kingdom (UK)⁵— especially if new usage paradigms such as M2M communications or IoT services do not catch on. All attempts to correct end user bundle prices have suffered from high competition in Western markets and have not been able to stop the downturn in falling end user prices and revenues. For this reason, more modern charging schemes leading to new, more fairly configured revenue opportunities, eg, quality differentiation, are required. A first step in this direction has been the trend to combine volume and bandwidth limitations for Internet access (see [PL13]), as forerunner of quality differentiation in networks.

Nowadays, it is broadly accepted that the Internet, especially when involving quality differentiation, extends far beyond the pure involvement of technology [LKC12; Jai06], which is mirrored in the emergence and vitality of QoE research and its absorbed business interest (see Section 2.1.1). Modern network solutions have to involve the users with their subjective perceptions, their interests in services, and their human susceptibility for various kinds of bias and irrationality alike (see Section 2.1.2). Such kinds of network, on the other hand, also involve markets and competition. Quality is subjective, it needs to be perceived and experienced, which leads to the difficulty of kick-starting an unknown network quality market and sustaining market success in the long run. Contrarily, the global context requires cooperation and orchestration among providers in order to provision the service expected by users. Finally, there is society defending acquired rights, conventions and benefits associated with the Internet. The “secret” to success in such a market has not yet been unlocked. This is reflected in the debatable market success of QoS offers which have not yet been able to trigger the turnaround for the entire industry. While insufficient market frameworks may have contributed the lack of customer-orientation has proved to be fatal for the market.

An (almost) flat pricing curve leads to an (almost) infinite traffic volume, which entices customers to use the Internet more intensively.

⁵ http://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2012/Mobile_cellular_2000-2011.xls, last accessed: 20th May 2016

1.2 Problem

New pricing schemes around network quality differentiation may provide the desired revenue prospects, but only if associated markets are properly parameterised and configured. The commercialisation of network quality, especially in its human-centric form of QoE, is afflicted with a substantial problem field, cf. Fig. 1.1, ranging across various economics via user-related to technical trouble spots. The root of all these problems relates to the by now unclear prospects, purpose, practicability, configuration and utilisation of multi-class service offers, providing reasons for previous market failures and the lack of determined actions. This has consequentially manifested the market position of low profit single class services such as BE Internet with flat price curves which, however, further foster the economic pressure on NSPs in the long run.

The more technically-oriented QoS-based quality differentiation, in particular, neither comes with a clear market framework nor does it establish sufficient links to users (or their perceptions and interests in provided quality levels), ie, it commonly concentrates on a single research discipline view that disallows an efficient and effective parameterisation and marketing for customers and NSPs. Apart from that, QoS technologies have not been tailored to provide a truly end-to-end service experience—potentially spanning over multiple independently operated Domains and centring on application and user needs at the end points—but have specifically been optimised for particular segments of the network, as in the access network for providing Voice over IP (VoIP) with LTE networks in the case of Voice over LTE (VoLTE) [IR.92].

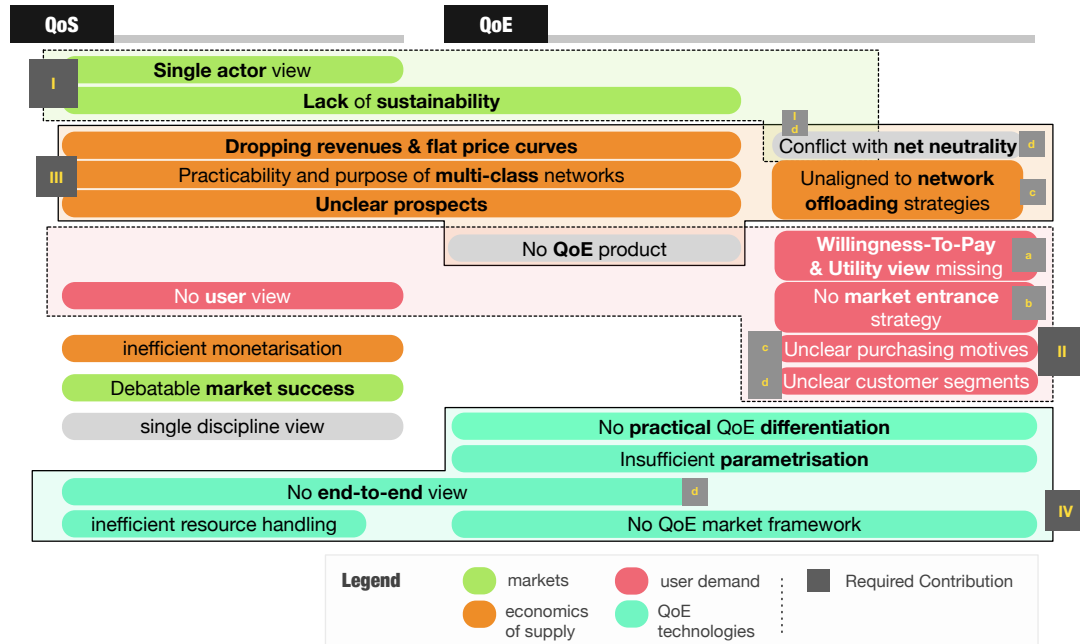


Figure 1.1: The QoE problem field.

In the sphere of QoE-based quality differentiation that adds a user perspective to QoS, similar deficits in the marketisation still exist. Today, only an indirect marketisation of QoE exists, where empirical results feed back a high granular Customer Satisfaction (CS) figure to the NSP. The efficient utilisation of resources like network capacity has only been targeted by high-level QoE management concepts such as in [AL08]. The otherwise vivid QoE community, both industrial and academic, has accumulated a backlog of investigating both user and NSP utility figures (and even how QoE is affected by pricing during purchasing situations) and the proper parameterisation of successful market entrances for quality-differentiated network services. The latter considerations are necessary in order to kick-start and sustain market success for novel network services. In addition, background motives for purchasing network qualities and services, and associated customer segments deserve more attention. On a technical level, to the best of our knowledge no practically realised QoE differentiation mechanism exists that takes empirically-backed customer demand and purchase behaviours into account. Thus, this opposes the underlying idea of QoE to deliver elevated or even guaranteed network quality to the end user based on their actual or estimated perceptions, preferences and WTP.

Orthogonally, other obstacles like the relationship to NN and other social values or the use of network offloading strategies for optimising the service QoE, social welfare or profits need to be considered when marketing QoE-differentiated services.

1.3 Methodology & Central Contributions

This thesis targets the co-equal marketing of quality and price (per quantity) and associated business prospects. The central contribution of this thesis is the transition from QoS or QoE to an associated utility perspective with promising commercialisation strategies. In response to the sketched heterogeneous multi-disciplinary problem field, the present thesis, therefore, utilises an equally multi-disciplinary set of techno-economic methodologies to obtain this goal.

One of the integral and most pressing problems relates to unclear foundations of quality-differentiated market design, where business interactions among actors in the QoE ecosystem have been studied insufficiently. Such sorts of problem are approached by clearer modelling of the transition from QoS to QoE with the help of VNs—see Fig. 1.2⁶. Relevant VN configurations are thereafter compared with the aid of both qualitative and quantitative VN assessment methods, where the latter requires the development of a first-ever quantification concept. The results will provide a basis for recommendations in order to create a solid foundation for the quality-differentiated network market that alleviates accumulated market pressures and fosters market sustainability in the long run.

The empirically-backed quantification of demand levels, WTP and utility figures requires the design and conduction of dedicated laboratory trials. The data obtained from

⁶ Bubble-shaped contribution indicators are given in Fig. 1.2 (the bigger they are, the more extensively topics are addressed) and associated with the numbered problems from Fig. 1.2.

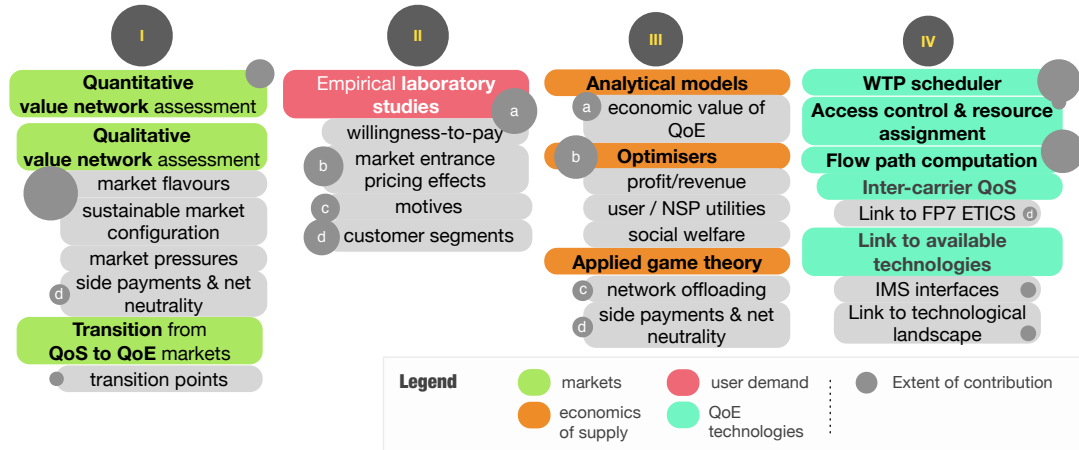


Figure 1.2: Methodologies and contributions for the QoE problem field (see Fig. 1.1).

these trials will characterise the product expectations of potential customers and can help to assess numerically the appeal for NSPs to enter such a market. The associated special case of market entrance is crucial for unfamiliar or novel services, especially when they are in the nature of an experience good [Nel70], such as network quality, and thus difficult to convey to customers. To date, no quality-differentiated network service market has been kick-started effectively. Thus, it is essential to investigate common pitfalls empirically in the configuration of immature markets. Apart from this, lucrative customer segments have to be targeted in order to keep technical complexities low while still creating an economically effective and technically efficient system.

The empirical data essentially further feeds our supply-side assessments which aim to provide recommendations on the optimal network resource provisioning without directing NSPs right into common configuration issues. In this part, analytic models, optimisers and applied game theory are used in order to work out revenue-, profit- or utility-optimal solutions for the single-class (referring to a single premium QoE class) and multi-class network service provisioning with quality or QoE guarantees. Especially due to its empirical parameterisation, this contribution clearly goes beyond the state-of-the-art in technical-centric, QoS-based alternatives. The NSPs are further provided with recommendations towards a revenue- or profit-optimal business configuration, but also provide the basis for user or social welfare optimisations. This is further supplemented by studying rationales for network offloading strategies and by assessing the economic value of side payments, ie, payments between Content Providers (CPs) and NSPs that potentially stand in conflict with NN viewpoints.

On a technical level, the practical realisation of market-oriented QoE is in focus by first constructing utility-aware network resource assignment mechanisms based on associated winner determination algorithms. The presented solution concepts will be of end-to-end character, support access and backbone networks, and aligned to leading technologies

of practical relevance, such as IP Multimedia Subsystem (IMS) used for VoIP in LTE. Additional effort is invested in the support of inter-Domain network quality agreements and associated technologies.

1.4 Solution Approach

Pairing the low cost and excellent scalability of today’s packet-switched networks with features supporting predictability and controllability—qualities rather known from the circuit-switched domain—, we will systematically introduce tools to moderate, market and provision QoE on the basis of today’s technological landscape and in parallel to standard BE Internet. This approach follows the position of Teitelbaum & Shalunov [TS03] by largely focusing on quality guarantees that effectively and efficiently manage the quality risks for customers, while limiting its scope to cases where QoE-sensitivity is sufficiently high.

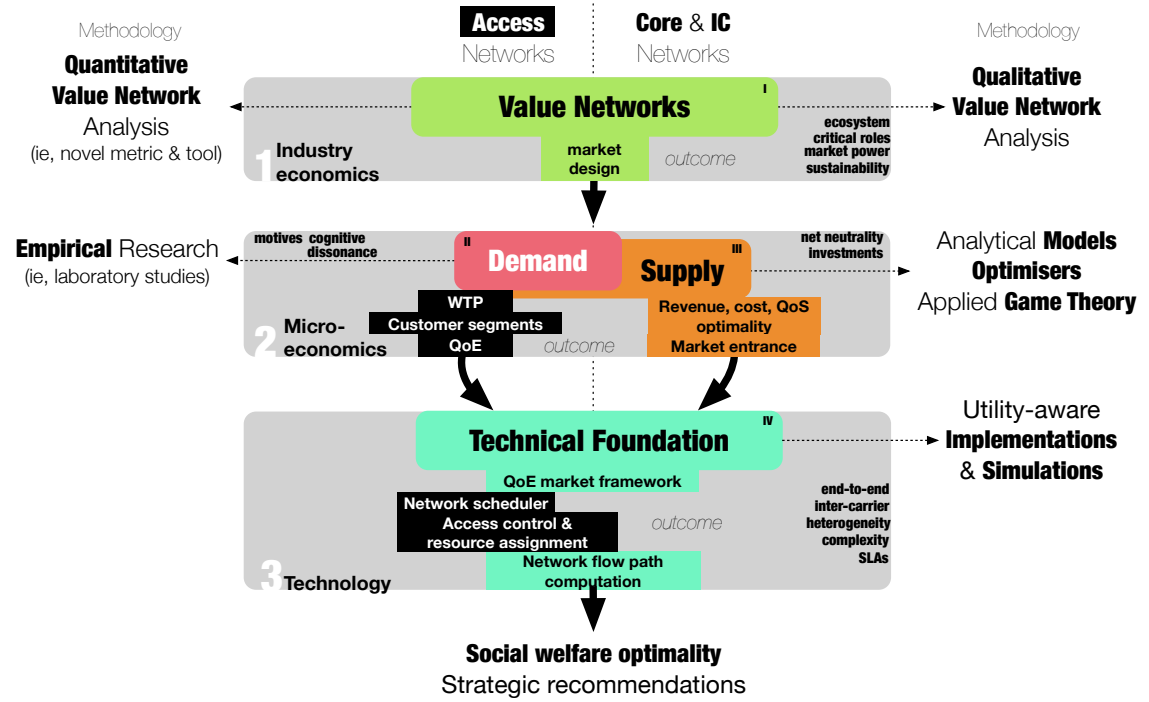


Figure 1.3: The big picture.

In its entirety, the multi-faceted QoE ecosystem (human, context, business, technology), in accordance with the definitions in [LKC12], requires comprehensive solution approaches that systematically integrate the domain knowledge of various disciplines. The present thesis will, thus, hierarchically deconstruct the problem scope in a top-down approach in order to ultimately come up with a profound and thoroughly parameterised

technical solution for quality-differentiated networks. This thesis will, therefore, collect a tailored selection of methodologies, as depicted and interrelated in Fig. 1.3, which are hereinafter introduced as common thread for the entire thesis.

Originating from the formulation of markets, qualitative and quantitative assessments of network quality markets and involved monetary flows are undertaken on **layer (1)**—ie, the layer of **industry economics**. Building on the existing qualitative means of analysis, a quantitative complement is designed and later on applied. Corresponding evaluations will set the scene for **layer (2)** linking **industry economics** to **microeconomics**.

The microeconomic layer (2) is bipartite: *On the demand-side*, service demand and associated spending behaviours are quantitatively modelled. The data for this model can be retrieved from a series of specifically designed laboratory-based trials, which among others assess concepts like WTP (short term spending), QoE (customer satisfaction supporting the loyalty), customer segments, motives and decision phenomena (eg, around cognitive dissonance effects), and finally transfer this understanding to utility figures. These trials augment the classical QoE perspective with novel economically- and partially also psychologically-driven assessment components. *On the supply-side*, providers need to perform profit optimisations in the short and long run that are based on the empirical evidence found in the demand-side assessment. This optimisation is orthogonally complemented by business strategies that increase the resource usage efficiency and lower the costs such as due to network traffic offloading. Strategies to cope with market limitations that may arise from NN boundaries—a conceptual antagonist to QoE-based differentiation—and societal aspects are derived from analytical models, applied game theory and structural VN analyses.

After bagging quantified demand and spending levels as well as revenue-optimal market configurations, a tailor-made technical realisation is created on **layer (3)**. Contrary, to purely technical works, this realisation is designed for users and practical applicability. The present work will, hence, illustrate solution approaches for the WTP-aware quality-differentiated scheduling, access control and resource assignment, and WTP and QoE-aware transmission flow path computation. Following an “end-to-end design principle” as coined by [SRC84] and extended by [TS03], the present thesis will further overcome the access-centric perspective, such as predominant in the QoE domain, by equally integrating core network and interconnection considerations. Only the demand considerations on layer (2) will remain entirely access network-oriented by definition.

As a common thread across all layers one **main research question** is formulated for this monograph:

Research question: Are there observable utility gains for the relevant stakeholders of the telecommunications ecosystem which drive them to transition from a single-class BE Internet to a market design with differentiated end-to-end network service qualities?

This research question can be transferred to **main contributions** w.r.t. the contribu-

tion layers given above. These main contributions include:

Main Contributions:

- Quantitative and qualitative VN analysis of the QoE ecosystem through representative case studies, where a novel assessment method is used for the quantitative part
- Investigation of market segments and purchasing motives for quality-enhanced network services and their relationship to the quality perception (ie, QoE)
- Profit and social welfare optimisation around NSP and user utilities derived from dedicated empirical WTP and QoE trials
- Game-theoretic and structural assessment of the economic feasibility of discriminatory actions such as side payments in communications networks
- Conceptual design of utility-aware technical network elements such as network schedulers, access and resource assignment, and flow path optimisations
- Integration of utility-aware technical network elements with IC and QoS technologies and marketisation frameworks

More detailed objectives are presented in the respective chapter introduction.

1.5 Related Theses

The present work has the closest interfaces to other theses in the area of QoE, most notably to those authored by former colleagues at the FTW Telecommunications Research Center Vienna (FTW): an ongoing work by Andreas Sackl on QoE expectations and a finished work on QoE impairments induced by delay by Sebastian Egger [Egg14a]. Although this thesis shares two trials with the first work, which have been conducted jointly in 2011 and 2012, the analysis and data usage are largely different. Contrary to both works, the present thesis much stronger focuses on the interworking of business, technological and user aspects related to quality differentiation in telecommunications (ie, empirically-backed techno-economics for communications services).

A strong collaboration link has also been established to colleagues of the CELTIC QuEEN⁷ (esp. Martin Varela and Toni Mäki from VTT Technical Research Centre of Finland (VTT)) and COST Qualinet⁸ (esp. Lea Skorin-Kapov, Ognjen Dobrijević and

⁷ <https://www.celticplus.eu/celtic-plus-project-queen/>, last accessed: 20th May 2016

⁸ <http://www.qualinet.eu/>, last accessed: 20th May 2016

Krunoslav Ivešić from University of Zagreb, Faculty of Electrical Engineering and Computing (UNIZG-FER)) projects. In this context the ongoing thesis of Toni Mäki (supervised by Kalevi Kilkki from the Aalto University) is of high interest that aims at planning and optimising networks and the associated businesses of the NSPs around empirical purchasing data for network qualities. Apart from the jointly conducted evaluation trial in 2015, which is reported in Chapter 6, his thesis is probably best positioned as continuation of the material in Chapter 5.

Another interface exists in the business context to works such as Antonio Ghezzi’s thesis (Politecnico di Milano) [Ghe11] where the quantitative view of VNs, which is in focus of this thesis, is supported by a conceptual link between VNs and business models. Contrary to the present monograph, this related work concentrates on business modelling design and extensive surveys of available qualitative literature. While also considering VNs, Thomas Casey (Aalto University) [Cas13] in contrast to Antonio Ghezzi studied value dynamics for the provisioning of wireless access as temporal evolution of an ecosystem (“value system”). While the chosen application area and detailed methodology differs from the approach of this dissertation, a conceptual relationship with our quantitative VN assessment can be acknowledged (a detailed separation is provided in Section 2.3). Among other techno-economic methods, the ongoing thesis of Nan Zhang (Aalto University) also uses VN analysis in order to identify economically attractive means of content delivery on the Internet. Her work specifically focuses on market perspectives of in-network caching solutions and the associated field of Information-Centric Networking (ICN).

Interesting complementary works also exist in the sphere of theoretic (mainly collaborative) game theory. Gideon Blocq (Technion; Israel Institute of Technology) [BO12], a project colleague within the *Seventh Framework Programmes for Research and Technological Development (FP7) ETICS* project⁹, has introduced a new concept denoted as the *price of selfishness*, as manifesto for intensified cooperation. This view has largely influenced the outcome of the FP7 ETICS project and as such provides a relevant boundary information for this thesis. Isabel Amigo [Ami13] (Telecom Bretagne) has further proposed QoS-aware concepts for allocating bandwidth and share resources in multi-Domain alliances in her dissertation, which is also placed in the context of ETICS. Contrary to the present thesis, both works do not centrally focus on empirical parameterisation in order to optimise communications networks.

The ongoing thesis of Christos Tsiaras (University of Zurich) focuses on countermeasures to the de-facto existing monopoly in the termination business of cellular network termination (ie, a single NSP that controls the traffic termination to its customers). For this purpose, he applies an auction-based approach that not only aims at improving the QoE of end consumers but also centrally integrates QoE into the auction. The area of traffic termination modalities is, however, out of scope of the present thesis.

On a more technical level, the present thesis will, among other concepts, make use of adaptive streaming technologies. A deep view on adaptive streaming technologies includ-

⁹ <https://www.ict-etics.eu/>, last accessed: 20th May 2016

ing various kind of large-scale real world measurements is presented in the work of Florian Metzger [Met15] (University of Vienna) for the case of cellular networks.

1.6 Structure

The remainder of this thesis is structured as follows (cf. Fig. 1.4): a QoE charging ecosystem is formed in Chapter 2 around an extensive literature review which, among other factors, shapes a taxonomy followed throughout this thesis. The QoE charging ecosystem is detailed by a case study, making use of qualitative and quantitative assessment techniques. Thereafter, the QoE ecosystem is parametrised on the demand side by empirically characterising customer spending behaviours in network quality markets in Chapter 3. Specific aspects such as purchase motives, market entrance pricing strategies and the relationship between monetary expenditure and QoE are considered. The supply-side counterpart is elaborated in Chapter 4, where social constraints, cost-cutting strategies and empirical demand results are integrated in various kinds of optimiser and model. While the previous sections have focused on a techno-economic viewpoint, Chapter 5 entirely shifts the attention towards technology-centric investigations relating to the operationalisation of techno-economic recommendations. This includes strategies to apply controlled service degradation while integrating QoE and expenditure data in solutions for each segment of the network. This section focuses on a truly end-to-end service experience, which encompasses inter-Domain heterogeneity and contractual agreement issues. Due to the fundamental reliance on empirical data, primarily originating from the assessment in Chapter 3, Chapter 6 will focus on the retesting of essential system parameters for cross-validation purposes. This work closes with concluding remarks in Chapter 7, which aggregates and harmonises the specific conclusions presented in previous sections. Orthogonal material is presented in the Appendix such as more extensive backgrounds on used methods or additional figures and data tables.

	Ch. 1: <i>"Introduction"</i>	
	Ch. 2: <i>"Network Quality Markets"</i>	Value Networks
Demand	Ch. 3: <i>"Demand & Spending"</i>	Supply
	Ch. 4: <i>"Economics of Supply"</i>	
Technical	Ch. 5: <i>"QoE & WTP Technologies"</i>	Foundation
	Ch. 6: <i>"Evaluation"</i>	
	Ch. 7: <i>"Summary & Conclusions"</i>	

Figure 1.4: The chapter structure of this thesis.

NETWORK QUALITY MARKETS

In his book entitled “Wealth of Networks” [Ben07], Benkler describes the emergence of a new and radically different kind of economy, the **networked information economy**, as a disruptive successor of common information economies such as financial services. The spreading of the network service ubiquity has started with the era of smartphones and cellular data connectivity and has since then been entangled with our daily life routines—also see IoT trends [Gub+13]. The available communications technologies have paved the way for a change from industrial production-oriented to increasingly connected information economies. The new industry’s distributed and highly parallelised character renders production-oriented market approaches, dominated by the sequentiality of processes, ancient and increasingly inadequate to meet the demands of modern society (also see [Ben07]). Apart from that, Benkler’s viewpoint further supports the economic value of modern network services and as such underlines the importance of understanding network service markets as targeted in this section.

The present thesis applies tools both reflecting the described economic transition and being tailored to the assessment of marketing novel quality-differentiated network services. At the core of our investigations stands a proper understanding of new economic roles in an increasingly dynamic and complex environment. While value chains have focused on the old paradigm of hierarchically organised activity sequences, this viewpoint seems less and less related to this new kind of economic world. More modern approaches such as business models suggest a simplistic single-entity viewpoint, which neglects the new increasingly networked nature of the Internet economy. As a consequence, other analysis methodologies are required which are being thematised, formed and utilised hereinafter.

Such a matching technique has been found in VN analysis, which allows for the critical assessment of the whole QoS Internet market eventually. Building on some essential background information as a starting point (eg, required definitions), this chapter constructs a simplistic market model capturing the essentials of QoS-based differentiation. A subsequent step reviews and extends the QoE pricing ecosystem and its relationship to underlying technical realisations. The rest of this section analyses most relevant VN con-

figurations in order to provide recommendations on the sustainability of the entire industry. In addition to specifically tailored qualitative techniques, this part relies on a quantitative assessment tool for generating more precise answers as a fully original contribution.

This chapter, hence, primarily targets the following objective:

Objective: Systematic representation and assessment of QoE market designs for communications services

This objective can be subdivided into the following four contribution blocks:

- Review of essential definitions and network quality viewpoints
- Detailing of the QoE ecosystem around fundamental market considerations and driving market forces
- Creation of means for identifying promising market configurations
- Illustration of sustainable market designs in the QoE and VoD context

2.1 Background & Related Work

The first part of the present background analysis is dedicated to the definition of most relevant terms being used throughout this thesis. Further terms and definitions can be looked up in the glossary section. The second part reviews human factors that have the potential to disrupt marketisation strategies whenever they are neglected. The third part illustrates the necessity for discussing “premium” quality markets, ie, quality levels going beyond classical BE offers. For this purpose, a simplistic market model is constructed, which apart from that revisits a set of economic terms relevant for this thesis alike. Finally, the methodological shift towards VN and the corresponding analysis techniques are reviewed.

2.1.1 Definitions

2.1.1.1 From Graphs to Communication Networks

This part will review syntactical elements in the assessment of VNs with graphs leading the conceptual way.

A graph consists [BM82] of leaves (“vertices” or “nodes”) and edges. Graphs may be directed or undirected, weighted or unweighted, labelled or unlabelled, etc. Thus, a graph inherently requires a characterisation. While for a generic graph theory introduction this work kindly refers to [BM82], context-specific definitions are given in Section 2.4.2.2.

The term *network* is regarded a synonym for graphs. Especially for empirically formed graphs such as social networks or communication networks the notion of network seems to

be preferred over graphs. Our analysis will be rooted fundamentally in graph theory, but will make use of specific applications and analysis mechanisms known from other fields. It, thus, explicitly also involves the notion of (social) networks, eg, [Fre06; Pin11].

Communications networks, including the Internet, are specific forms of networks that describe the local or global connection of individual devices or entire ASes. This view has further been narrowed down by Jordan & Schaffer in 2014 [JS14].

“ The term Communications Services means all services; (a) over a network that uses a public right-of-way; and (b) that reside at or below the network layer or are required to manage the network. ”
Jordan & Schaffer, 2014 [JS14]

The network layer in this case refers to the well known International Organization for Standardization (ISO)/Open Systems Interconnection (OSI) model (cf. ISO/IEC 7498-1 [IEC7498]), which formulates seven layers: 1) Physical (bitwise point-to-point connection), 2) Data link (reliable bitwise point-to-point connection), 3) **Network** (packet delivery in networks), 4) Transport (reliable packet delivery), 5) Session (data sessions), 6) Presentation (data representation), 7) Application.

Alternatively, RFC 3439 [RFC 3439] presents a different and more cleaned up approach, which is often denoted by the “Internet layer model”. For the present thesis, both models serve our purposes.

2.1.1.2 Quality of Service (QoS)

The quality of communications networks can be measured and moderated. QoS is a generic concept for characterising network quality. The QoS notion has yielded a variety of interpretations ranging from more technology-oriented to more user-centric views, concentrating rather on communications service-centric concepts or taking a broader scope. Some noteworthy definitions are given in chronological order as follows:

User-centric views:

“ ...the collective effect of service performance which determines the degree of satisfaction of a user of the service ... ”
International Telecommunication Union (ITU)-T E.800, 1994 [E.800-94]

“ The totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service. ”
ITU-T E.800, 2008 [E.800-08]

Technology-oriented views:

“ ...a set of qualities related to the collective behavior of one or more objects ... ”
ISO/OSI, 1998 [IEC 13236]

“ The most extensive measure is Quality-of-Service (QoS), comprising all aspects of a connection as voice quality, delay, loss, reliability etc. We consider a subset of these, Grade-of-Service (GoS) or network performance, which only includes aspects related to the capacity of the network. ”
ITU-D SG 2/16 & ITC Draft, 2001 [D SG 2/16]

“ The Quality of Service (QoS) provided by a transport service is defined in terms of the way a traffic stream is affected when it is transported through the network. This is typically in terms of the probability of cell loss, delay, and cell delay variation. ”
Courcoubetis & Weber, 2003 [CW03]

“ QoS refers to the capability of a network to provide better service to selected network traffic over specific underlying technologies. In packet-switched networks, QoS is described by the following parameters: **Timeliness**, which indicates QoS time dependence and involves the sub-parameters **Delay** and **Jitter**, **Loss Probability**, **Reliability**, **Throughput** and **Delivery Time**. ”
Markaki et al., 2007 [MCN07]

The chronological order extrapolates a trend over time towards more technology-centric and more specific definitions of QoS. While this may clarify technical details on interpreting QoS in practice, the user focus has apparently faded away. For more details on a historic evolvement of the QoS term, its flavours, and its march towards QoE we kindly refer to [Rei10].

Teitelbaum & Shalunov [TS03] capture some other comprehensible but early viewpoints around the transition from classical circuit-switched to cheaper packet-switched network paradigms. They introduce the end-to-end principle and its relationship to QoS:

“ ...what both customers and providers need are not true guarantees, but rather tools to manage risk.
At the level of packet forwarding, risk management tools might include adequate capacity provisioning or, perhaps, statistically provisioned QoS techniques aimed at providing a level of assurance without the full cost of a *guarantee*. ...The low cost of the Internet stems not just from the simplicity of the best-effort service model, but from the end-to-end design principle [SRC84] as well.

Referring to the early definitions in [SRC84] and the corresponding notion in [Kil12], the end-to-end network design principle refers to the application-aware selection of network solutions. For quality-differentiated networks, the end-to-end principle may refer to an intelligent QoS and/or QoE adaptation, which is parameterised at the end points. The end points provide sufficient information about the needs of the user, the application and type of usage, and potentially expected revenues or costs for optimally adapting the network transmission. The cost perspective for meeting customer expectations has already become apparent with the contrasting of quality guarantees [Kes97] and risk management techniques for network quality in [TS03].

“ The function in question can completely and correctly be implemented only with the knowledge and help of the application standing at the end points of the communication system. Therefore, providing that questioned function as a feature of the communication system itself is not possible. ”

Saltzer, Reed and Clark, 1984[SRC84]

“ **end-to-end principle:** the principle that, whenever possible, communications protocol operations should be defined to occur at the end-points of a system. ”

Kilikki, 2012[Kil12]

For modern network quality-differentiated networks, the scalability, flexibility and cost advantages of modern packet-switched networks have to be related to the predictability and controllability of circuit-switched networks. Mechanisms for packet-switched networks have to follow an end-to-end design principle [SRC84; Kil12], which integrates the knowledge about the used application. Known “qualities” from telephony, in terms of network characteristics, have been linked to the “Internet world” where concepts like Asynchronous Transfer Mode (ATM) [ATM93] have emerged, which centrally build the Internet around QoS considerations. Today, ATM has no practical relevance anymore and is thus left aside in this thesis.

Teitelbaum & Shalunov [TS03], on the other hand, direct researcher towards the expectations of QoS, eg, whether absolute guarantees or some weaker form of elevated quality is targeted. They direct their line of thought towards a deliberate *risk management* tool-set, which is capable of providing technical or economic responses to failure scenarios. Thus, the authors construct an orthogonal view to quality concerns due to *congestion* by centrally discussing failure in the QoS mechanism design. Amongst others, they propose to decouple the technical and economics QoS understanding by considering “Warranted QoS” concepts, which combine “statistical over-booking” and other cost optimisations

with insurance mechanisms—potentially provided by third-party insurers and traded on dedicated marketplaces.

For the design principles for “elevated QoS”, Teitelbaum & Shalunov [TS03] refer to a quotation of Keshav [Kes97] that essentially describes the technical roots of the QoS discussion. Teitelbaum & Shalunov themselves, however, draw attention to a more risk of failure-oriented point of view.

“ The Holy Grail of computer networking is to design a network that has the flexibility and low cost of the Internet, yet offers the end-to-end quality-of-service guarantees of the telephone network. ”
Keshav, 1997 [Kes97]

While both quality degradations from congestion and failures (risk of failures) will be considered in this thesis, a focus is laid on congestion. Other causes for low quality, such as artificial scarcity, are further acknowledged.

2.1.1.3 Quality of Experience (QoE)

Countervailing this trend of rather technology-centric quality definitions, the notion of QoE has emerged. QoE essentially focuses on the needs and perceptions of users. Over the years, QoE has found its place in academia and in industry (as one metric for measuring or estimating CS), although, as concluded in 2010 [ETSIW], it “is not a universally well understood concept”. While since then the terminology has clearly evolved, a global agreement on one definition does not yet exist. Hence, some noteworthy definitions of QoE are given as follows:

Standardisation:

“ The overall acceptability of an application or service, as perceived subjectively by the end-user.
NOTE 1 – Quality of Experience includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.).
NOTE 2 – Overall acceptability may be influenced by user expectations and context.
ITU-T Recommendation P.10/G.100 “Vocabulary for performance and quality of service”, Amendment 1 [P.10] (also see [D.197] ”

“ A measure of user performance based on both objective and subjective psychological measures of using an ICT service or product.

NOTE 1: It takes into account technical parameters (eg, QoS) and usage context variables (eg., communication task) and measures both the process and outcomes of communication (eg., user effectiveness, efficiency, satisfaction and enjoyment).

NOTE2: The appropriate psychological measures will be dependent on the communication context. Objective psychological measures do not rely on the opinion of the user (eg, task completion time measured in seconds, task accuracy measured in number of errors). Subjective psychological measures are based on the opinion of the user (eg, perceived quality of medium, satisfaction with a service).

ETSI [TR 102]

Research projects & Academia: A versatile set of QoE definitions also exists in literature. A notable definition was coined by the Qualinet project:

“ Quality of Experience (QoE) is the degree of delight or annoyance of the user of an application or service. It results from the fulfilment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the user’s personality and current state.

Qualinet (working version) [LMP12]

A distinction is often made between *subjective* and *objective* QoE (cf. [Lag+12]; also see [Sch+13]):

- **Subjective QoE** is based entirely on empirical ratings of end users, in the field or controlled laboratory testings, ie, the subjective perceptions of the users. Following the recommendation of [P.800], subjective QoE is mostly measured in Mean Opinion Score (MOS) on an Absolute Category Rating (ACR) scale, which ranges in the interval [1, 5]: 1: “bad”; 2: “poor”, 3: “fair”, 4: “good” and 5: “excellent”. This thesis will concentrate on the ACR scale for measuring and acquiring subjective QoE data, and complement it with well-known binary acceptability (or acceptance) ratings. According to [TS 103], subjective ratings of resources (network resources, devices, etc.) are influenced by the type of application (eg, video streaming for a particular content, bitrate, codec etc.), the interface (eg, display, sound), the context (temperature, lighting, noise, etc.) and the user (preferences, personality, etc.). Subjective QoE tests are regarded to be “expensive and time-consuming” ([TS 103]).
- **Objective QoE**, as the complement to the subjective perspective, predicts human behaviour based on a model or set of equations taking technical and potentially also

empirically-backed human parameters into account [Lag+12]. According to Laghari *et al.* [Lag+12], only solutions focusing on particular quality aspects have appeared, eg, Structural SIMilarity (SSIM) (for images) or Peak Signal-to-Noise Ratio (PSNR), which insufficiently integrate “influential factors” like the context. The parameterisation of objective QoE models with empirically-backed subjective QoE data is often referred to as hybrid QoE assessment approach. Without empirical parameterisation, the assessment converges back to a QoS-based quality perspective.

Orthogonally to this distinction, there is also room for interpretation regarding both the mapping between QoS and QoE and the design of QoE models, which aim at uncovering important factors and relationships that shape the quality perception of humans:

First, the relationship between QoS and QoE is often argued to be non-linear—mostly an exponential [FHT10; JL10] or logarithmic relationship [Rei+10] is expected. The latter theory mainly builds on historic observations from psychophysics regarding the roots of human perception. Especially, the Fechner scale¹ (see [Hei04]) provides evidence for a logarithmic perception curve of various stimuli, eg, temperature, weight, sound volume, and likely also (network) quality. Despite continuous empirical testing efforts, no undisputed ground truth exists for mapping QoS to QoE in all details. This thesis will rely consistently on empirically-backed QoE results for specific applications with a slight preference towards logarithmical mappings. We will only revert to QoE approximations from QoS parameters, eg, IQX hypothesis [FHT10] or ITU-T’s E-model [G.107]², whenever insufficient data exist.

Second, the conceptual relationship between QoS and QoE was expressed in a multitude of QoE models over time: As surveyed in [LC12], only a single model addresses all relevant QoE factors including technical, contextual and business aspects as well as human roles and demographic influences (see the definition of WTP for follow-up information on business aspects). All models, however, do not provide direct (and empirically-backed) strategies to utilise a given network infrastructure and customer base optimally in order to optimise a given objective function.

A QoE monitoring approach was proposed in [TS 103], which implements a specifically constructed QoE model, ie, the ARCU model, with a specifically designed agent. The authors intended to provide means for estimating the QoE based on a multi-layered monitoring approach involving all aspects of QoE — see [TS 103] and previous discussions.

¹ The Fechner scale is frequently also referred to as Fechner law or erroneously as Weber-Fechner law [Mar07]. The validity of the Fechner scale was questioned in [Ste61] for input extrema with the proposition and subsequent parameterisation of the alternative Stevens’ law.

² The E-model standard captures various kinds of impairment on the audio quality (eg, room noise, one-way and round-trip delay, or equipment impairment), which is reflected in the E-model rating as results often compared to other QoE measures (eg, in [SRV04]).

2.1.1.4 Usability, Quality of Perception (QoP), Quality of Design (QoD)

Usability has already been defined by ISO 9241 standard in 1998 [ISO 9241], and as such looks at effectiveness, efficiency and satisfaction matters for software:

“ The effectiveness, efficiency and satisfaction with which specified users achieve specified goals in particular environments.
effectiveness: the accuracy and completeness with which specified users can achieve specified goals in particular environments
efficiency: the resources expended in relation to the accuracy and completeness of goals achieved
satisfaction: the comfort and acceptability of the work system to its users and other people affected by its use ”
ISO 9241 [ISO 9241]

Today, the term is often associated with the field of Human Computer Interaction (HCI), which conceptualises and tests means of interaction with digital systems and often involves means of visualisation or content representation. This is reflected in web usability expert Jacob Nielsen’s extended definition³:

“ Usability is a **quality attribute** that assesses how easy user interfaces are to use. The word “usability” also refers to methods for improving ease-of-use during the design process.
Usability is defined by **5 quality components**:

- **Learnability:** How easy is it for users to accomplish basic tasks the first time they encounter the design?
- **Efficiency:** Once users have learned the design, how quickly can they perform tasks?
- **Memorability:** When users return to the design after a period of not using it, how easily can they reestablish proficiency?
- **Errors:** How many errors do users make, how severe are these errors, and how easily can they recover from the errors?
- **Satisfaction:** How pleasant is it to use the design?

There are many other important quality attributes. A key one is **utility**, which refers to the design’s functionality: Does it do what users need? ... ”
Jacob Nielsen, 2012

³ Jacob Nielsen, “Usability 101: Introduction to Usability”:<http://www.nngroup.com/articles/usability-101-introduction-to-usability/>, last accessed: 20th May 2016

In [TR 102], the concepts of *usability*, *Communication Service Performance* (ie, QoS and QoE) and *User Experience* (eg, appeal) are distinguished and aligned to each other. As in our understanding, QoE, *User Experience*, and CS only seem to represent different flavours of the same matter in the context of network quality, this thesis concentrates on differentiation of technical quality level, ie, QoS, perceived/experienced quality, ie, QoE, and application specific appeal, eg, usability.

Other terms such as Quality of Perception (QoP) [TR 102] or Quality of Design (QoD) [Rei07] have almost vanished from the academic jargon and have been absorbed by the QoE concept, which has been playing an overwhelming role in the telecommunications industry in the last years.

2.1.1.5 Willingness-to-Pay (WTP)

The concept of WTP for network services (and in particular network qualities) has recently regained substantial attention. While not necessarily bound to telecommunications WTP provides a valuable ingredient for further enriching the vital QoE research field, extending its boundaries from user- and technological considerations to an additional inclusion of an economic dimension, ie, formulating the roots for a QoE ecosystem. Proper assessment of WTP has the potential to substantially alleviate the mismatch between QoE and the deduced (approximated) purchasing behaviours (understanding QoE as a market product). The improved understanding of purchasing behaviours will allow for a series of optimisations in the telecommunications market. The roots and definitions of WTP will be reviewed below.

The mere assumption that more satisfied customers would yield an increased WTP is only backed by “anecdotal evidence” [HKH05] (also see Finkelmann [Fin93] and Reichheld & Sasser [RS90]). Nonetheless, unexpectedly limited research is available that links CS (see details in Section 3.1) to monetary means (outcomes). Such a knowledge gap creates substantial marketisation problems, as revenue figures, obtained from the customers’ spending behaviour and often approximated from CS variations, provide central information for any market optimisation. Nonetheless existing research provides two main viewpoints on the link between CS and outcomes:

1. **Endurance:** The well-studied Willingness-to-Accept (WTA) concept describes the required minimum monetary compensations for persuading a customer to “endure” or “accept” an undesirable condition. Thus, WTA forms an “endurance metric” as reversal of WTP. This may relate to the minimum discount that is necessary to prevent a customer from switching to a competitor or opting out. In Experiment 1 [FP502] of the FP5 project Market Managed Multi-Service Internet (M3I), this position has been empirically confirmed by obtaining lower acceptability rates for higher price bands than for lower alternatives.
2. **Maximisation (spending):** WTP, however, refers to the maximum amount customers are willing to pay without losing their interest in a given product (quality).

Separating the WTP from a general disposition to spend money for a given product (or quality of a product), ie, Readiness-to-Pay (RTP) [RNR08], some groundbreaking definitions exist in (economic) literature:

- Cameron and James [CJ87], Krishna [Kri91], Ries [RNR08]: Maximum amount of money customers are willing to spend.
- Monroe [Mon90]: Reservation price customers assign to a product/service/experience.

While several works in literature such as by Haneman [Han91], Cummings *et al.* [CBD86] and Fisher *et al.* [FMS88] have revealed substantial disparities, Shogren *et al.* [Sho+94] even observes divergences between WTP and WTA, contradicting prior intuition (see, for example, [Wil76]). Both Haneman and Shogren conclude with the existence of persistent disparities between WTP and WTA whenever no obvious substitute for a product exists⁴. On the contrary, the effects vanish under the existence of clear substitutes. The WTP and WTA concepts are subsequently treated to be potentially disparate, as no clear substitution strategies exist in most cases⁵. The distinction of WTP and WTA is continued in Section 2.1.2.

In the context of network qualities, the endurance-based perspective seems to originate from the BE business model where low prices (as discounts from the price for perfect conditions) are implicit compensations for anticipated fluctuations of the service quality. The present thesis, however, targets the premium segment where quality above the average is guaranteed⁶, which highly correlates to the maximisation viewpoint. While the given WTP definitions are in no way contradictory, this thesis will follow the formulation given in [CJ87] for clarity reasons.

Van Moorsel [Van01] further claims that improved network quality, ie, QoE, would increase the WTP of customers, which is in analogy with the generic CS and WTP relationship (see [HKH05]) not self-evident and requires empirical backing. Nevertheless, it is reasonable to assume that improved network quality will not lower the WTP of rational customers.

Following the recommendations of Anderson [And96] and Gotlieb *et al.* [GGB94], Hobmurg *et al.* [HKH05] study “the link between CS and WTP” backed by empirical evidence. This link will serve as reference point for WTP analysis for network qualities in this thesis. Such related studies and the above given considerations also provide a firm stand on utilising empirical studies as vehicles for confirming fundamental relationships

⁴ This may be seen equal to no or limited “choice” for users that would be willing to switch to a competing service or substitute technology/product.

⁵ In some cases WTA and WTP may converge to equality, especially in rather unusual polypoly cases. Current trends in Europe, however, rather sketch a densification or consolidation of the telecommunications market, especially in the mobile segment. Treating WTP and WTA as unequal substantially reduces testing risks.

⁶ Providers typically do not aim to price their products at the brink of annoyance, but to design products for lucrative customer segments.

between CS and obtainable monetary outcomes, ie, WTP. This research perspective is absolutely in line with the intentions of the present work.

2.1.2 Human Factors

Perceived risks of losses or chances of gains play an important role in the valuation (ie, utility, WTP) of alternative options. By mainly referring to the findings of Bernoulli's *St. Petersburg paradox*⁷ [Ber96], Friedman [FS48] points out that the belief of risk involvement through plain expected utility maximisation would be "ancient" and "unrealistic". Following the considerations of von Neumann and Morgenstern [VM44], Friedman has aimed at coming up with numeric utility considerations incorporating this pseudo-irrational behaviour of users. With his work he could highlight the tendency of users to underestimate moderate utility gains at moderate risks, while overstating high gains at high risks, eg, playing lottery despite reducing their expected utility. More than 30 years later, the *prospect theory* of Kahneman and Tversky [KT79] has confirmed Friedman's claims empirically. They could also show that risk aversion is favoured if users face losses at moderate risk levels while risk seeking seems to be dominant in the case of gains with identical probability. In communication networks, the risk of unsatisfactory QoS is often spatially and temporarily limited, which renders guarantees attractive for humans (if fairly priced), ie, certainty as an orthogonal quality attribute. On the other hand, elevated BE Internet is a less effective instrument as only a limited number of services and customers may profit from an improved QoS (ie, a low probability to profit from upgrades).

The utility gain interpretation by users is affected by the status quo as *reference point* [KT79]. A shift of reference points may occur whenever expected gains or losses are already taken as given. In the past, for example, one may have observed a QoS improvement (eg, in terms of access speed) at a given rate each year, which cannot be maintained anymore in the future; then the reference point may be on the expected QoS to attain rather than on the status quo. According to [KT79], further influences may also be the unwillingness to accept losses, ie, one may consider gambling in order to compensate unaccepted losses. A productive change of reference point may be the formulation of "decision problems in terms of final assets, as advocated in decision analysis" [KT79], which seems to eliminate risk seeking in most cases. For the conduction of empirical trials, reference points may form a testing bias that may limit the practical applicability of results (see Chapter 3 for response strategies to various kinds of testing bias).

According to Knetsch [KS84], required compensation payments tend to exceed WTP figures, ie, WTP and WTA are disparate. He further questions classical indifference curves in which temporal losses could be compensated by equivalent future gains without affecting the mean CS. Among users, this may support concepts such as the *Scitovsky paradox* [Sci41] where social welfare may be increased by compensation payments between losing

⁷ The St. Petersburg paradox refers to an experimental lottery with low perceived value for humans, but with infinite expected value from a mathematical point of view. Paradox describes that test subjects pseudo-irrationally turn down a winning lottery ticket.

and gaining parties. In order to avoid such paradoxes, per-user “fairness” moderation is integral in quality-differentiated communications networks.

2.1.3 Market Model

Minding the objectives of both users and NSPs a fundamental market model is formed in order to clarify the scope and potential of QoE-based user discrimination. A central point of investigation will be the consumers’ quality surplus—the quality provisioned above the demanded level—as source of economic inefficiency. This surplus may be used as service-specific or customer-specific discriminator. The following paragraphs will characterise this concept.

2.1.3.1 Objectives

According to Sen *et al.* [Sen+13], the **users** have the objective to selfishly maximise their **utility**⁸ levels being subject to charged price and the provided quality. **NSPs** by contrast aim at maximising their **profit** subject to the charged price and the corresponding costs. In the case of quality-differentiated network services, the quality, ie, QoS, needs to be considered as determining cost dimension. On a microeconomic market level, both perspectives work together to form possible equilibria. In contrast, classical BE Internet provisioning represents a club good [ST97], which renders a maximally intensive network resource usage dominant for users.

Beyond this very classical perspective, societal objectives may also include qualitative aspects like the fairness among users or the avoidance of various kind of defective strategies, which are discussed in more detail in Section 4.5.1.

2.1.3.2 Fundamentals

The network infrastructure, and its usage Z , is limited by an overall capacity C ⁹:

$$Z = \sum_{i \in Q} \left(\zeta(q_i) \cdot d_i \right) \leq C \quad (2.1)$$

where q_i is a continuous quality function¹⁰, and ζ is the amount of network resources associated to the provisioned QoS and the demand d_i for the set of all considered quality classes Q in the form $\text{symQoSClassPL} = \{\text{symQoSClass}_1, \dots\}$. C may be raised by infrastructure investments or may degrade through the omission of required maintenance efforts in the long run. Infrastructure investments require long-term business prospects (ie, profits).

⁸ The definition of utility in the context of quality-differentiated services deserves specific attention, which is targeted among others in Section 5.3.

⁹ Depending on the used technologies, especially in wireless networks, this capacity (eg, measured in throughput) may not be static over time or location.

¹⁰ Depending on the context, the quality level may refer to overprovisioning practices as multiplicative factor for the requested resources, eg, bandwidth in Mbit/s, or may represent content quality settings, eg, video bitrates, with different network resource demands.

For networks, an effort function c of the form $c(Z)$ exists that is subject to the total quantity of provisioned network resources Z . The effort transitively relates to the offered QoS and the respective demand levels d . Based on a single-class Internet assumption (eg, BE), the level of q is fixed for all provisioned services d , while specific q_i may exist for each quality class Q_i in a multi-class case.

An invalid utilisation of the network exceeds the overall capacity C . Each efficient, valid configuration is Pareto-optimal. An assignment is efficient whenever $Z = C$ holds, and thus no single user's position can be improved without deterioration for other users. Whenever $Z < C$ the assignment is regarded to be inefficient to some degree. The fewer users are provisioned for a static C , the higher the average quality q . Whenever C is fully utilised the quality can only be moderated by shifting resources between quality classes in Q .

The quality q^* is the quality at the equilibrium point. The supply (quantity, quality of infrastructure) matches the demand curve at an attractive position¹¹. The quantity ζ^* is the equilibrium quantity. The demand at d^* (following a demand function of users) is subject to q^* and a given price point p^* , $d^*(q^*, p^*)$ (p^* can be neglected for flat price curves). In turn, due to C limitations, also the q is expressed relative to d : $q^*(d^*)$. In the longer term C may be extended by additional investment costs I .

2.1.3.3 Consumers' Quality Surplus

The consumers' quality surplus α ,

$$\alpha = \left(\zeta^*(q) \cdot d^* \right) - \beta \quad , \quad (2.2)$$

can be quantified as resource demand difference between the statically provisioned quality level q and the actual requests β , where β results from the resource demand for the provisioned quality q^u of each user u and the overall user demand equilibrium d^* .

A utility gain for individual users exists when their desired minimum quality level q^u (ie, the acceptance quality where customers do not refrain from using the service with the equilibrium price level p^*) is exceeded by the equilibrium quality q^* without any extra charge. The difference between the utility $\mathcal{U}(q^u)$ and $\mathcal{U}(q^*)$ characterises the individual utility gain due to the existence of a quality surplus.

When assuming that a considered population is not entirely homogenous, as in practice, a resource usage inefficiency typically occurs, as in practice q^u will not be identical for every customer u . Thus, q^* may not satisfy the acceptance quality level of all customers (which will be excluded from the service), while it may exceed the demands of a set of other users. Price discrimination especially when combined with quality differentiation, as will be centrally investigated in this thesis, provides means for a more individual service provisioning that aims at both lowering the rate of excluded customers while also providing low-cost offerings to other customers.

¹¹ The used optimisation strategy, typically profit or revenue optimisation, will determine the equilibrium point.

Technically this refers to a multi-class service design that due to higher resource efficiency skims the consumer's quality surplus when the system is parametrised properly. Such services are expected to be economically more attractive for NSPs¹² and may increase the social welfare of the system. Such a multi-class design clearly goes beyond the BE scope and, thus, sets the scene for the elaboration of a multi-class QoE marketisation concept as pursued by this thesis.

Conclusion #2.1. *It is economically preferable to transit from a single-class to a multi-class Internet service provisioning when providing deliberately selected quality classes.*

On the other hand, a multi-class quality differentiation may create complexity issues and societal conflicts (see Section 4.5.3). A notable other example for increasing the resource efficiency in telecommunications is the transition from charging for QoS to QoE (see Section 3.2), which will be combined with a multi-class service provisioning in this thesis (esp. see Section 4.3).

2.1.4 From Value Chains to Value Networks

♣¹³ The concept of value chains originates from the age of industrial production while today's economy is increasingly complex, dynamic, distributed and especially information-oriented. Though, the traditionally prevailing value chain concept [Por85], envisaging the development of value through a series of chained activities, is not sufficient for the strategic assessment of such complex business ecosystems anymore.

According to Benkler's view on modern information economies [Ben07], the Internet constitutes a driver of a major economic shift. In this context, the network service industry, including the surrounding "New Economy", is an especially interesting case that is characterised by non-linear value creation processes, ie, a parallelisation of business interactions, with limited hierarchical control, for which VNs appear to be the only feasible analysis technique nowadays. This view supports the position of Peppard in [PR06] who sees the VN perspective "more suited to New Economy organisations, particularly for those where both the product and supply and demand chain is digitized". In the context of mobile networks, they further describe the old value chain concept as the "walled garden" approach of the early phase. Strict hierarchical organisation has made it difficult for content providers to participate in the market, which substantially impairs the innovation diffusion process. In modern Internet economies such "walled garden" scenarios, however, do not meet the demands of customers anymore and require reconsideration and appropriate other tools.

Triggered by this economic change, a wide range of literature on VN principles (external view of enterprises—the business relationships of firms) is available—eg, [HS89], [NR98], [GNZ00], or [GLS09]—whose understanding has to be set in relationship to business models

¹² The ease of conveying products to customers when retaining a single-class BE Internet service provisioning may provide scalability effects that are not considered in this perspective.

¹³ *The materials used in this section have been reported in [ZRG11] and [ZR13] before.*

(internal perspective of an enterprise)—eg, [Tim98], or [Tee10]. The linking role between autonomous business model islands is best described by Peppard’s unembellished definition of VNs (see below). Casey & Töyli [CT12] further claim that market success of any business model “largely depends on how it interacts with models of other players in the industry”.

“ A service delivered over a network requires the use of two or more network components. Think of value networks as a set of relatively autonomous units that can be managed independently, but operate together in a framework of common principles and service level agreements (SLAs). Firms in the network are independent; otherwise they would fall into a case of ‘vertical quasi-integration’ [Jar88]. However, the relationships enjoyed by the firms in the network are essential to their competitive positions. The structure of the network plays an important role in firm performance and in industry evolution [MKP98]. ”
Peppard, 2008 [PR06]

The importance of understanding roles within a VN is further strengthened by Jacobides’ [JKA06] concept of an industry architecture¹⁴, which conceptually explains the value generation distribution among roles within a VN. The authors, based on prior works [Tee86], especially stress the importance of complementarity (especially in modular systems) and mobility (“number of assets that can potentially enter a combination”, ie, a kind of fungibility measure for a player) when evaluating the distribution of bargaining powers in an industry architecture. Based on their business strategy and technical capabilities, players may be able to take on fortunate roles that control bottlenecks in the industry architecture, ie, those players ‘become the “bottleneck” of their industry’ or may become ‘gatekeepers’ [Bal09]. Bottlenecks may create asymmetric dependencies in the architecture that may substantially shape bargaining powers in the industry. An interesting case for studying bottleneck roles is given in [Bal09] where the increasing “platformisation” in the cellular telecommunications market leads to various kinds of stable industry architectures, ie, various kinds of VNs configurations.

From a reversed perspective, the broadly applied intra-firm concept of business models has emerged about a decade ago [Tee10], and has been interlinked with VN concepts recently [Bal07]. Both Ghezzi [Ghe12] and Ballon *et al.* [Bal07] have proposed concepts for introducing VN parameters for the business modelling design in the telecommunications sector while Ghezzi has even specifically considered the inter-Domain case. Nevertheless, the VN-related parameters still require qualitative assessment, eg, based on expert experiences.

¹⁴ While in the original work, the industry architecture has been mainly aligned to the value chain concept, the argumentation line systematically follows the broader industry understanding of VNs.

2.2 Ecosystem

✦¹⁵ The conceptualisation [Kil08] (also see Section 2.1.1), modelling [Rei+10; FHT10; LKC12; Wu+09] and empirical testing of QoE [De +10; KRD12; Rei+10; Egg14b] in relationship to objective technical parameters and combinations thereof has received notable attention. Despite the undeniable interest in the commercialisation of the increasingly systematised QoE understanding, the interest in specifically designed charging concepts has been much lower in telecommunications. The present section will target the most fundamental step in characterising charging for QoE by revisiting and extending recently appeared concepts on the QoE ecosystem [Rei+13a].

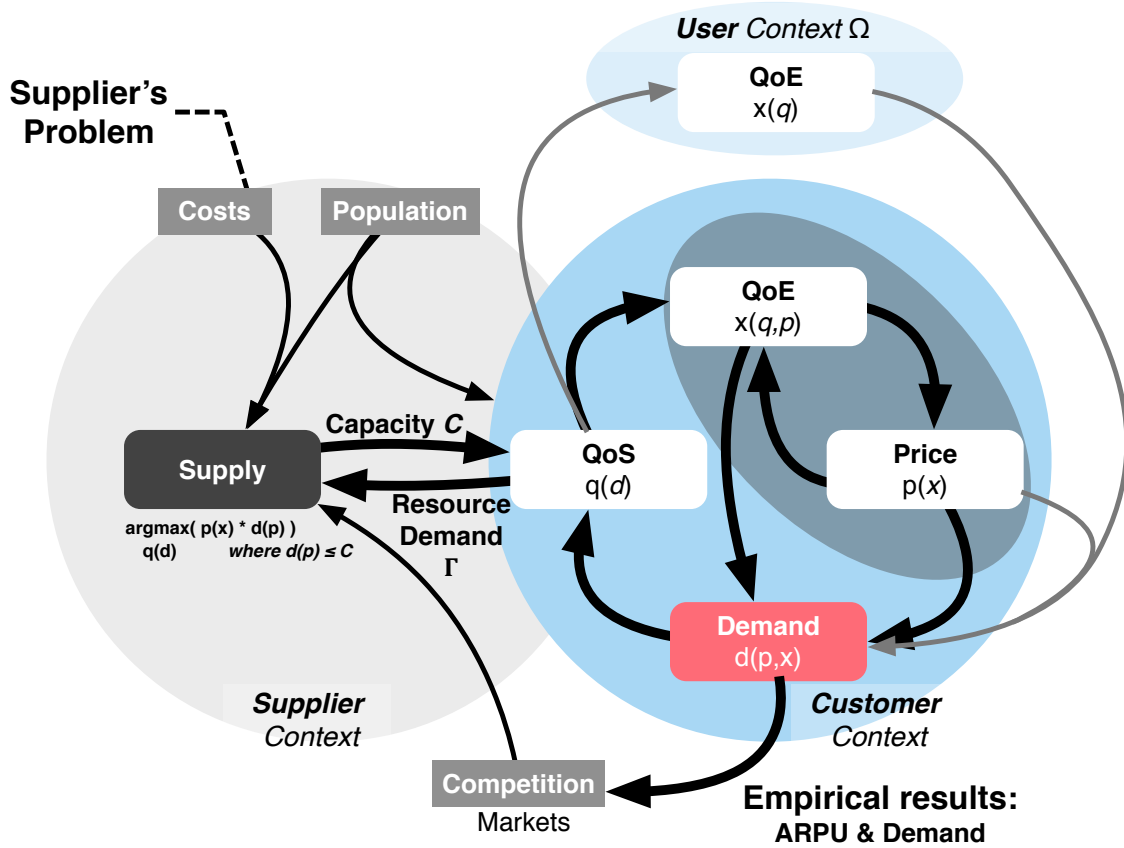


Figure 2.1: The ecosystem of charging for QoE (Source: [ZRS14], extending view in [Rei+13a])

Taking the infrastructure as given, such as implicitly assumed in [Rei+12; Rei+13a], the **demand-centric module** of the QoE ecosystem can be formulated in accordance with the definitions in [Rei+13a] (circle on the right side in Fig. 2.1): the QoS $q(d)$ is subject to the fraction of the demand d and the (constant) overall capacity C . The transition

¹⁵ The materials used in this section have been reported in similar but shorter form in [ZRS14]. Other original sources are indicated below.

from QoS to QoE x refers to the mapping of technical quality parameters to a perceptual and very individual representation of quality. The function x is not only characterised by the test population, the application and user context parameters Ω , but primarily by the two variables q and p where p is the charged price for the current level of x . Apparently, this leads to a non-trivial fixed point problem¹⁶ where the settling of x and p is loop-wise dependent on each other. Whenever both x and p have eventually been set, the corresponding demand $d(p, x)$ can be measured, which in turn characterises the current QoS. This demand loop is iteratively repeated whenever individual factors are altered. In a large enough network such as the global Internet, we may assume that the population, demand levels, available capacity, used services, but also the perception of quality always slightly changes—thus, this loop is quickly iterated in order to efficiently use the available network. While QoE assessments are typically conducted in the user context (ie, usage, but no purchasing context), the commercialisation, as depicted in Fig. 2.1, requires a more intensive focus on the customer context (ie, purchasing context).

For QoS charging, Reichl *et al.* [Rei+12] assume monotonically falling $d(p)$ (convex) and $q(d)$ (concave) functions, whereas $p(x)$ (linear relationship to x) is assumed to be monotonically growing. They conclude after rescaling all parameter to the interval $[0, 1]$ that only two trivial fixed points exist: 1) Free of charge Internet service, ie, $(p^*, d^*, q^*) = (0, 1, 0)$, or 2) Premium high QoS Internet services, ie, $(p^*, d^*, q^*) = (1, 0, 1)$. Unfortunately, as both trivial fixed points yield no revenue ($\mathcal{R} := p \cdot d$) and thus no profit for providers, they are commercially not interesting. Thus, less trivial boundary equilibrium points are required. While the existence of such interior equilibria has been proven in Reichl *et al.* [Rei+13a], the authors have also disproved the possibility of an interior equilibrium in the case where customers are only quality-sensitive. For detailed specifications of their models, the proofs, and details on the outcome, this thesis kindly refers to [Rei+13a] but also to [Rei+12].

Contrary to prior works quantifying or estimating functions affecting the QoE (including the price; see the dark blue circle in Fig. 2.1), the present work focuses on the charging process itself, ie, obtained revenues, demand levels, etc. This is mainly inspected on the axis of empirical WTP measurements for a given QoS level in a monopolistic market. Such measurements both provide quantitative figures of the existent demand for a QoS level and the corresponding ARPU values.

The demand module, as sketched above, has been complemented in [ZRS14] by formulating a corresponding **supply-centric module** (ie, “Supplier’s Problem” as detailed in Section 4 and sketched in Fig. 2.1 on the left-hand side), which is linked inherently to the rest of the QoE ecosystem through the aggregate resource demand Z and the available capacity C . This module has to take care of both strategic as well as operational decisions regarding the revenue-optimal provisioning of communications services for an extrinsically given population. Strategic decisions refer to a considerate capacity management in the longer run, which weighs up the RoI for the CAPEX required to increase the capacity C . On a shorter timescale, this module is responsible for provisioning the right quality

¹⁶ A formalisation of this fixed point is extensively discussed in [Rei+12; Rei+13a].

levels to a lucrative segment of the population—thus (implicitly) taking care of admission control, quality and pricing decisions. On both timescales a more holistic view including competition has to arise in this module, which progress from the typical monopolistic considerations such as associated to classical QoE or WTP trials. While techno-economic issues are being handled in Section 4, Section 5 is dedicated to the corresponding technical details.

Another specific challenge refers to the often neglected perspective of time. Explicitly, alternations of quality perceptions but also corresponding purchasing behaviours may be imposed by an adjustment of price or quality levels. Market entrance pricing recommendations, which are worked out in Section 3.6, will specifically address this perspective. Due to the absence of corresponding data, changes in WTP in the course of time (on the scale of years) and other time-related effects are left for further work.

2.2.1 Economic Dimensions

♣¹⁷ Especially when following an inter-domain perspective where several providers are cooperating in order to be able to provision global Internet services, the role of QoS and QoE needs an economic formulation. Following the economically crucial **revenue** and **cost** dimensions, which interact within business models especially wrt service pricing, we can systematically separate QoS from QoE.

The revenues are inherently linked to the purchasing decisions of customers. Customers subjectively evaluate offered prices and qualities, which will have a substantial impact on CS and thus QoE figures. The perceptive evaluation of customers eventually leads to a purchasing decision¹⁸, which is captured by WTP measurements. In turn, the price is the dominating moderator for the revenues.

In the backend, costs take a more substantial role, as they provide a natural limitation of what can be provisioned, ie, creating a direct relationship to technical constraints such as the overall capacity C . Network dimensioning, and associated investments, are CAPEX- (eg, trenching cables) and Operational Expenditure (OPEX)-intensive (eg, reserving capacity from third parties, energy costs, etc.) which requires justification and minimisation. The cost of network quality directly refers to the QoS being provisioned or even guaranteed. By contrast, QoE first needs to be translated to QoS in order to create a technical and more objective representation.

Subsequently, QoS and QoE will be attached as dominating factors to **access**, and **core** and **inter-carrier** networks (cf. Fig. 2.2):

Access. The direct customer access, including the collection of service fees, renders CS and, thus, QoE to be the key factor in access networks. After collecting the access fees, the provisioning of the network service remains cost- and thus QoS-centric for

¹⁷ This section is structured in alignment to our contribution in deliverable D5.100 [CEL14] of the QuEEN project: <https://www.celticplus.eu/celtic-plus-project-queen/>, last accessed: 20th May 2016.

¹⁸ Users might decide not to spend any money at all.

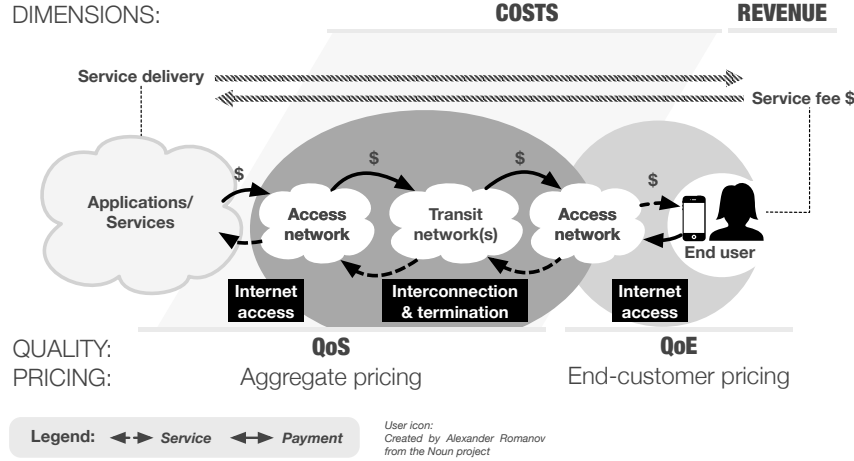


Figure 2.2: Classical economic QoE charging dimensions.

NSPs. The transition from QoE to QoS is handled pragmatically in access networks (ie, the transition point). Depending on the VN configuration the transition point may be migrated, eg, to a CPs.

Core. Due to the absence of a direct customer access, core networks are cost-centric. Individual requests and QoE demands are attached to aggregate QoS classes.

Inter-carrier. On a European or global scale, inter-carrier agreements such as (paid) peering or transit services, which are purchased from other providers, are essential. In analogy with the core networks, the cost for reserving aggregate resources (with QoS guarantees) is central. Whenever regulatory authorities consider to intervene, QoE may become a useful basis for defining threshold values for characterising a good service quality for a representative traffic mix. Any controlling measurements will likely remain on QoS level, however.

On the supply-side, a cost-revenue equilibrium needs to be identified by both targeting high CS for a wide range of customers (ie, maximisation strategy), and minimising costs (ie, minimisation strategy). Whenever an entity is capable (and in charge of) transitioning between QoE and QoS, which requires direct or indirect customer access, a strategic advantage is obtained. In particular, such a position comes along with information advantages (ie, gains from information asymmetry), which may allow the absorption of the gain when transitioning from a QoS- to a QoE-centric network provisioning—see Section 3.2.

This thesis concentrates on the revenue perspective, which has largely been neglected in related work so far; QoE management perspectives [AL08] have, for example, focused on the optimal resource utilisation and costs only.

2.2.2 The QoE Value Network Flavours

♣¹⁹ The transition between QoE and QoS requests is handled at dedicated **transition points**, where user-centric information, such as QoE and/or WTP, is aligned to underlying network mechanisms, ie, QoS. The location of the transition points depends on the business case and thus VN configuration. Classically, the transition takes place at the premises of the Access NSP facing the end user. At this point, bundles of individual user requests (or per-session QoE demands and characteristics) are converted to aggregate QoS demands, required for the core and IC network perspective. Further VN configurations are discussed after this.

2.2.2.1 Application-centric Pricing (cf. Fig. 2.3)

Supported by the spread of service-oriented and subscription-based business models, eg, Software as a Service (SaaS) such as Office 365 replacing classical software purchases, and the stunning market success of “app stores” in the mid-00s, the responsibilities of the CP have been substantially extended. Today, the access to and consummation of content, payment services and many customer relationships functionalities have been taken over from NSPs—increasing the stakeholder’s market power at the expense of NSPs. In this light, this trend may be continued towards moderating QoE (advancements) not only by NSPs but also by CPs—cf. Fig. 2.3.

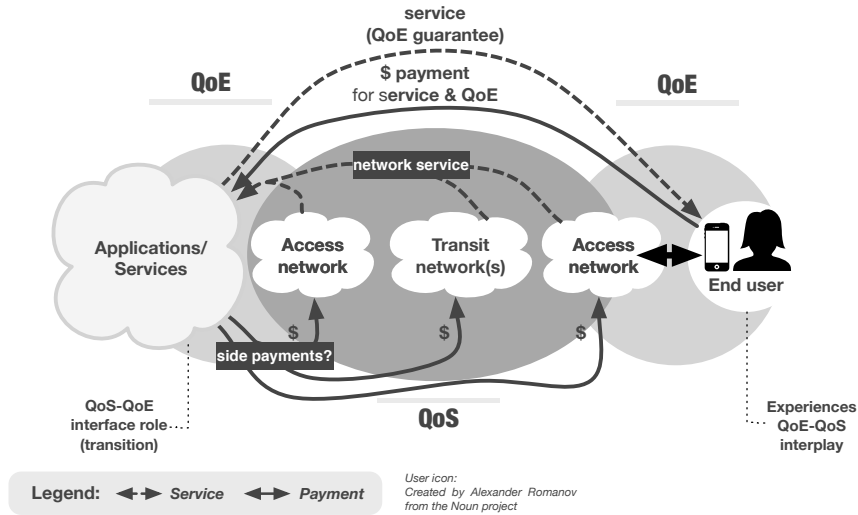


Figure 2.3: App-centric pricing [CEL14].

When customers can directly purchase the entire service (including the network transmission) from a CP, eg, a video streaming platform, no customer relationship with any

¹⁹ The hereinafter presented contents are based on our contribution to [CEL14] (project consortium-wide publication). The materials have been reworked for this thesis.

NSP needs to be established. The overall service experience (eg, network quality, video quality and purchasing process) is attributed to the CP in charge—ie, the assessment of individual service components such as network QoS are replaced by an aggregate experience. Such a structural change also positions CPs as moderators of the network (and service) QoE for their customers. The NSP’s role is reduced to a supplier of network services, ie, a “bitpipe” provider.

In such a scenario, the classical Internet access business is affected alike. Due to the absence of an end customer relationship, NSPs may not necessarily have any Internet access agreements with customers anymore²⁰. Their business may be entirely subject to side payments received from CPs—issues and their relationship to NN are discussed in Section 4.5. Such payments are compensations for the invested network resources (including QoS) and are likely insensitive to QoE demands of customers. The pricing model for network services will, hence, be cost-centric. QoE advancements, as part of the overall service experience, are translated to QoS (ie, bandwidth, latency, etc.) purchase requests by the respective CP.

Apart from the gap between QoS- and QoE-centric network segments, another one exists between QoE and monetary expenditures of customers (see Section 3.3 and 4). At both gaps, transition points are positioned that essentially determine the commercialisation strategy. These points may be moderated by large enough service providers or may be assisted by NSPs due to their QoE competencies and existing high-volume inter-Domain business relationships. NSPs may in the later case act as *brokers*²¹ that provide an end-to-end QoE provisioning and monitoring even with service guarantees. By selling the QoS-to-QoE and/or QoS-to-WTP transition services, such a broker holds a strategic position in the VN. In other words, the broker profits from information asymmetry advantages and economies of scale.

2.2.2.2 Access network-only (eg, local synchronisation)

For example due to clever caching, some services, mostly of end-to-end nature involving multiple network Domains, may be localised partially. Thus, the service operates to a large extent in access networks or even in Personal Area Networks (PANs). The transition responsibilities may consequently shift towards the Access NSP, the end user Domain (PAN management), or device manufacturers supporting the PAN. Despite the variations, the charging for network quality will again be QoE-centric unless an M2M use case is targeted where the user involvement may be very indirect. The present thesis will, however, concentrate on cases where multiple NSPs are involved, ie, end-to-end inter-Domain scenarios.

²⁰ This VN configuration illustrates an extremum. In practice, the importance of Internet access contracts may be reduced, but their existence may remain relevant for QoS- and QoE-insensitive services or small CPs. Some CPs may not have the capability or scale to negotiate agreements with all required NSPs separately.

²¹ The role as a broker may also be taken on by other stakeholders.

2.3 Qualitative Analysis

♣²² The relatively limited amount of related work on VN analysis [GNZ00; All00] concentrates on qualitative means. Despite the common usage of VNs since the 1980s, structured assessment techniques have received limited attention (also see [BC08]). High-level VN building blocks have been studied in [BC08] based on a review of existing VN modelling approaches such as *e3value* [GA01] (“lack of a clear strategic focus” [BC08]; more driven by business modelling considerations), *c3value* [Wei+07] (rather myopic strategic analysis relative to direct competitors) and VN Analysis (VNA) [All00] (as discussed in more detail below). The resulting model differentiates in “capabilities” and “assets” of each actor and their required “supplies”. The business interactions are based on the exchange of various kinds of resources (“offers”), eg, “brand”, “product”, “information” or “service”.

The recent proposition of VNA²³ [All00] has been leading the way for systematising the qualitative understanding of particular VNs. VNA provides means for detecting elements of interest within a VN, eg, misspecifications, inefficient value flows, role interdependencies, response-delivery relationships, risks, analysis on tangible and intangible deliveries and heart beat roles of VNs by studying a VN’s business interactions, irrespective whether they exchange tangible or intangible assets (cf. [All08]).

In contrast, the main contribution of the Network Value Analysis (NVA) approach [PR06] is the definition of VN analysis phases. The analysis process itself is only rudimentarily systematised. Nevertheless, the authors suggest the assessment of per role “value dimensions” and interrelations between roles, as well as the consideration of the expected dynamics of a VN. These suggestions support the view of this thesis that a better understanding of dependencies between VN roles is required with more advanced analysis methods.

In 2012, Casey & Töyli [CT12] had come up with a method for modelling VN dynamics for the case of two-sided platforms [RT06]. For this purpose, they defined a structural feedback model being operationalised by a simple analytic model, which captures VN dynamics over time. This technique expressively illustrates the dynamics of roles within a VN, eg, the cardinalities (eg, how many users are participating?), and as such can provide valuable illustrations for complex market constellations. On the other hand, the quantitative assessments are not based on syntactical considerations of the VN but on a feedback model that has been defined context-specifically.

Despite the high value of qualitative assessment approaches such VNA, a full syntactical quantification may increase both the accuracy and the comparability of VN assessments and may, thus, positively support VN reconfigurations. However, quantified VN representations, eg, reflecting bargaining powers in a VN, have not yet emerged. Thus, essential research questions in the IC network service business—like the configuration of

²² The hereinafter presented materials are advancements of our initial contribution to [CEL14] (project consortium-wide publication).

²³ Verna Allee: <http://www.valuenetworksandcollaboration.com/>, last accessed: 20th May 2016

the Transit NSP’s role—cannot be resolved by quantified answers, but have to rely on expert assessments only.

Nonetheless, the systematic qualitative analysis may provide essential first indications on VNs and may also help to narrow down the analysis scope for more elaborate quantitative assessment techniques. This section, hence, constructs a coherent qualitative analysis template in agreement with the available literature subsequently. Apart from the approach in [All00], which is tailored directly to the needs of qualitative assessment techniques, the advancements from the related area of Business Model (BM) design are also included. In particular, the work of [Bal07] gives an extensive overview of BM design literature also bridging the gap to VNs. While a similar approach specifically designed for IC networks has been provided by [Ghe12], we will subsequently focus on Ballon *et al.* [Bal07]’s well-established approach.

The below presented analysis template has previously been reported in [CEL14] and has only slightly been reshaped in this thesis. Special attention has been paid to the completeness of the assessment by aggregating similar categories from [Bal07] with those from [All00]. Structurally the template primarily follows [Bal07] as almost identically reported in [CEL14]:

- **Value Network Parameters**

- **Combination of Assets / Heartbeat roles** [All00] [Bal07]: Concentration of relevant assets on individual roles—relating to heartbeat roles in (‘What role(s) is most essential in the life of this network?’, ‘What would happen if the person in a “heartbeat” role is replaced with someone else?’ [CEL14]). From [All00] we can also specifically address the following *related* “network indicator”:
 - * **Role Contribution and Structure:** Assessment of distribution of ties (eg, the existence of central roles) and the associated value creation structure.
- **Vertical Integration:** Degree of internalisation of activities or actions.
- **Customer Ownership:** Role serving as a contact point for customers with essential influence on their actions (eg, product choice), which is supposed to “own” the customer.

- **Functional Architecture Parameters**

- **Modularity** of products / services
- **Distribution of Intelligence / Exchange of Value** [All00] describes the distribution of critical processes within the VN, eg, dominant value exchanges or key interactions (pulse points) may exist [All00]. According to [Bal07], more distributed network topologies, for example, encompass a better distribution of intelligence.

- **Interoperability:** The degree of interworking of systems, processes, roles with others, eg, the fungibility of the exchanged products. In combination with modularity, the degree of meshing and collaboration is defined in the network.
- **Financial Model Parameters:** Mainly **revenue and cost-sharing modalities** and underlying revenue/cost models (eg, fairness or sustainability) are of importance on a network level.
- **Value Proposition Parameters**
 - **Positioning:** The positioning refers to the question of which kinds of products (eg, substitute or complement to an existing product or new niche product) are delivered to which kinds of target customers (also wrt customer segment, regional aspects, competition level etc.).
 - **Customer Involvement:** Customers may play a more active or passive role in shaping the product. Two-sided markets [RT06] may yield scalability gains with actively participating users or customers may develop their own product on top of the initial product. For example, Application Programming Interfaces (APIs) may provide data access to social media platforms, which may enable new business ideas to be realised. High customer involvement may thus foster their commitment, but may also support innovation diffusion processes.
 - **Intended Value & Perceived Value [All00] (Network Indicator):** “Intrinsic as well as extrinsic attributes (network effects)” that express the degree of “operational excellence”, “product leadership”, “competency” and “customer intimacy” [Bal07] (wrt custom-made solutions, lock-in effects, etc.).
- **Network indicators (complementary subset only) [All00]**
 - **Sequence** is a description of main value flow paths.
 - **Reciprocity** describes the structural balance of business interactions, ie, each business interactions may be counterbalanced by another business interaction (a product is sold for money to a customer; two distinct counterbalanced business interactions). The reciprocity is also influenced by the fungibility of exchanged resources or services, eg, some goods may be harder to be traded with another role.
 - **Stability:** The stability of a network is positively affected by the degree of meshing in the network. In other words, the more intensively roles collaborate (eg, trade or knowledge transfer), the closer is their relationship and the entire VN in the long run.

2.4 Quantitative Analysis

♣²⁴ In contrast to the qualitative technique given in Section 2.3, the subsequently introduced methodology, denoted as “Value Network Quantification (VNQ)”, aims at providing novel quantitative means for assessing the dependency of a single entity on a specific VN (as currently modelled) and its role within a VN. Thus, whenever an entity is quantitatively very dependent on the VN, it may experience high market pressures when, for example, negotiating prices with other players. This concept can also be referred to as **entity dependency model** to underline its semantical heading. When we, for example, suppose that a machine builder only sells a single type of product to another entity, then it may reduce its dependency through a diversification strategy. Diversification may in such a case be achieved by targeting a second customer entity with a different type of product. Similar kinds of dependencies (mirroring market pressures) are of interest for the quantitative assessment of this section.

The construction of the VNQ method consists of six phases:

1. Essential **directions** for the formation of the VNQ method, such as the identification of key analysis dimensions and cases, are addressed in Section 2.4.1.
2. **Representation** forms for VNs that enable their quantitative assessment are introduced in Section 2.4.2.
3. **Requirements & assumptions** for a practicable solution are worked up in Section 2.4.3.
4. The **VNQ approach** is then designed in Section 2.4.4 and followed up by calculation details in Section 2.4.5.
5. Additional usage scenarios of the VNQ method are illustrated with the example of a **revenue sharing** technique for information economies in Section 2.4.6.
6. The **technical realisation** of the VNQ tool and a sensitivity analysis in Section 2.4.7 provide a technical closure.

In Section 2.5 (for the results see Section 2.5.2), this technique will be contrasted to qualitative techniques in a representative case study on VoD services.

2.4.1 Directions

This section clarifies the essential directions for designing a quantitative analysis technique (eventually VNQ), which is able to reveal each entity’s (ie, actor role) dependency on the

²⁴ *This section extends the quantitative analysis materials for VNs as reported in [ZR13] and [ZR15]—successor works of [ZRG11; ZR12]. In an early work [ZRG11], we have been able to illustrate the expression power of novel quantitative assessments. Such techniques can provide clear figures on the dependency of individual actors on the current industry setup, ie, the ecosystem or VN.*

surrounding business ecosystem, ie, a VN, in order to illustrate potentials for improvement. The intended outcome is an easily interpretable dependency metric providing a statement on the dependency of one entity on the VN relative to others—such values range in $[0, 1]$. This outcome supports the understanding of dominant (lowest dependencies; closer to 0) and dominated actors (highest; closer to 1). Actors with high dependency may lose interest in this business in the long run, which may cause disruptions in the ecosystem in this form. For this reason, this approach targets the comparison of various flavours of a business ecosystem in order to allow the creation of a sustainable configuration with mutual interest of all required actors.

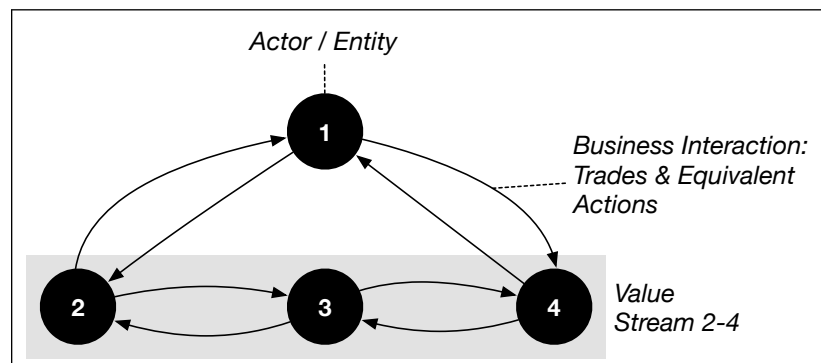


Figure 2.4: The Structure of Trading.

The business ecosystem refers to a VN, which captures business interactions between actors (cf. Fig. 2.4). The quantification of VNs requires a more systematic understanding and definition of VNs around which a tailored metric is defined. A VN captures the market structure of an industry or market. It is capable of representing multiple value streams. For example, the business interactions between actors 2, 3 and 4 (see Fig. 2.4) could form a value stream with different time horizon than the interactions between actors 1 and 2, and 1 and 4. An individual value stream, if detached from others, is similar to a value chain representation. Value chains conceptually concentrate on activity chains, while value streams represent a looser concept that is more suitable for information economies. Actors are subsequently treated in a more technical sense as entities. Their integration in the ecosystem is represented by business interactions.

The quantitative assessment will be conducted on the basis of a candidate VN along the dimensions of

1. **Quality**—ie, distribution of value among business interactions—,
2. **Scale**—ie, market volume or the size of the entity as scaling factor)—and
3. **Risk**—ie, risk associated to an entity—,

which are separately assessed for the case of

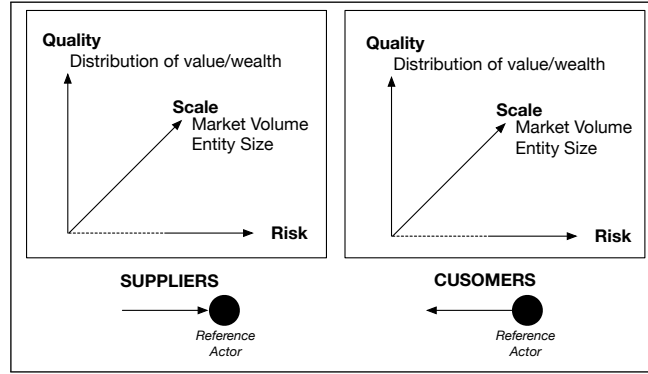


Figure 2.5: VNQ Dimensions and cases.

1. **Customers** and
2. **Suppliers**

of each entity. While, of course, an outgoing business interaction points to a customer, the customer entity faces a business interaction from a supplier. This reciprocity is deceptive, as each entity may be more or less exposed to a single business interaction depending on its business diversification strategy. As a result the assessment along the *Quality* and *Scale* axes may be affected. Hence, a separate assessment of supply- and customer-side cases is necessary.

The core of the dependency metric computation is the *Quality* dimension, which requires the access to the most detailed market data. This perspective fundamentally relies on a utility assessment and the understanding of the value distribution among business interactions or competing alternatives, in analogy to wealth distribution. For this purpose, the Gini coefficient [Bre+84] is used, which is a well-known metric for inequality measurement in the field of information theory and economics, and has a series of highly beneficial properties over, for example, entropy-based alternatives—see Section A.1 in the Appendix. The used Gini coefficient ranges in $[0, 1]$ where 0 is total equality and 1 refers to maximal inequality.

The *Scale* dimension, as will be shown within this section, is directly linked to the *Quality* assessment. The separated *Risk* dimension, however, requires a specific integration strategy in the overall metric.

2.4.2 Representation

VNs have been widely assessed in the past, but typically with qualitative means of analysis. When designing a complementary quantitative technique, the representation of VNs requires an augmentation and a stricter framework, which is established hereinafter.

2.4.2.1 Value Network Dependency Model (VNDM)

♣²⁵ A central consideration in the field of industry analysis are *Porter's five forces on firms* [Por79]—industry rivalry, (bargaining) power of suppliers and customers, substitutes, and threats of new market entrants—which, however, originate from a single-enterprise perspective that stands in contrast to VNs' external view. Though, the concepts such as suppliers or customers can also be linked to directional relationships being used within VNs. Each relationship from a source to a target entity forms a supplier-customers relationship. The substitutes, eg, switching from Transit NSPs to Content Distribution Network (CDN) providers, may be modelled by the effort for reshaping the VN. On the contrary, market entries may be quantified as costs (investments) for managing the entrance to an existing market. While a visual representation does not exist and the intended view is limited to a per-entity perspective, the used forces, therefore, clarify some essential data needs for a powerful VN representation.

In literature, eg, [BC08], Value Models are known for modelling business relationships with defined interfaces for exchanging goods with economic partners, ie, a mechanism for visually modelling Value Networks [BC08] such as *e3Value* [GA01]. An extension of *e3Value* named *c3Value* [Wei+07] introduces the goal modelling technique from [Yu95] and Porter's strategic competition analysis [Por80] to *e3Value*. In contrast to *e3Value*, it applies a resource-centric representation targeting the dimensions of customers, capabilities, and competition through strategic resources modelling [BC08]. These dimensions are still modelled qualitatively in *c3Value* in order to maintain the fundamentals of *e3Value* (cf. [Wei+07]), which still leaves bargaining powers without quantified evidence.

Especially Value Models and Porter's five forces will serve as inspiration for creating the Value Network Dependency Model (VNDM) that is specifically designed for the representation of VNs. The new quantitative focus requires the inclusion of quantitative business and market data: costs, resource values and cardinalities. Cardinalities—as indicator for internal rivalry—define the number of instances being addressed by one relationship, eg, an *CP* can make business with a multitude of End Customers (*ECs*), which requires a certain degree of competition among *ECs*—also see the formalisation introduced in Definition 5. Similar to [BC08], a differentiation in resource types (values) transferred via relationships is required in order to capture their economic influence on the VN. The classification of resource types will concentrate on the fungibility of exchanged resources, where notable types are money, uncustomised exchangeable goods (ownership is linked to exchange), customised exchangeable goods, promises/requests/treaties (indirect ownership exchange), and intangible goods. The VNDM satisfies all these requirements and provides a visual representation that covers most aspects.

A exemplary VNDM is illustrated for Internet-based content distributions, eg, for VoD services, in Fig. 2.6. In this example, the *EC* pays for a service requested from a *CP*. The service provisioning through the network is established by the Edge NSPs—providing

²⁵ This section is aligned to the contents and presentation given in [ZRG11].

access to the Internet—of the *CP* and the *EC*, as well as by the *Transit NSPs* connecting *Access NSPs*. Both the *CP* and the *EC* have already paid (initial value ϕ) for their Internet access (solid line) exposing certain service costs to the *Access NSPs*. The *Access NSPs* (at the network edges) in turn have to compensate the IC fees charged by the *Transit NSP*. Complementary, all actors have to face certain role-dependent costs for their participation in the VN such as investments in own Customer Premise Equipment (CPE) by the *ECs*, network investments by *NSPs*, or content production costs by the *CPs*. As a result of the non-linear value streams of the access and the service provisioning, traditional value chains cannot capture this example.

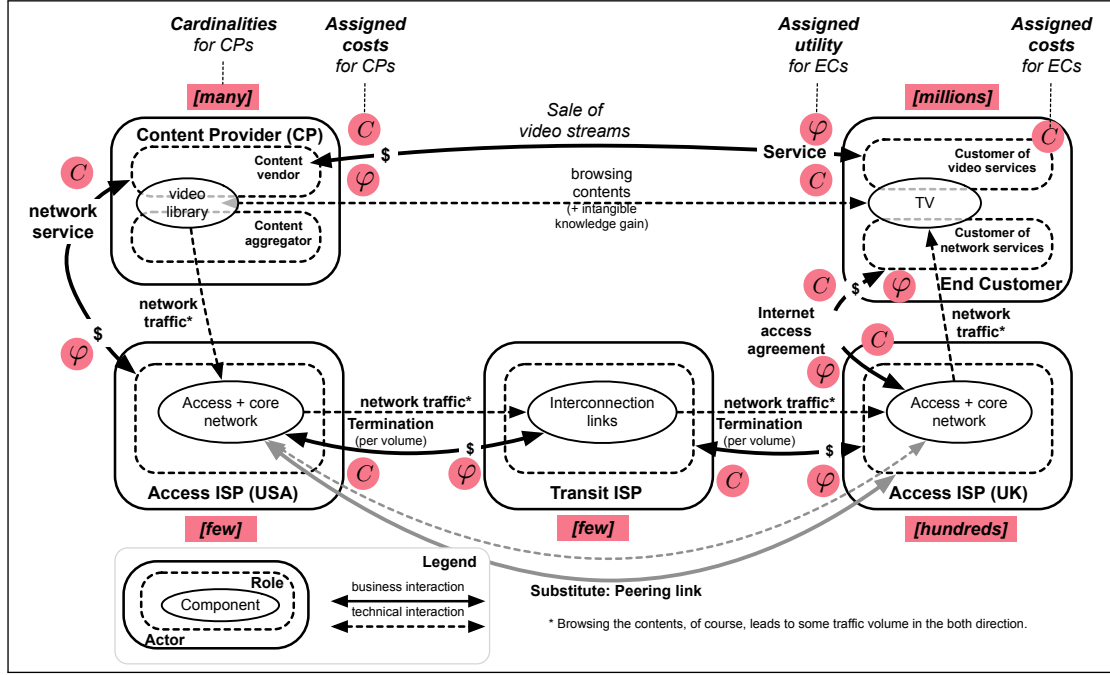


Figure 2.6: An exemplary Internet market as VNDM

The VNDM in Fig. 2.6 uses an intermediary abstraction level (level of detail)—a differentiation which is very common in Model Driven Engineering (MDE), eg, [JK06]. This allows the differentiation of *Access NSPs* (at the network edge) providing their services to *CPs* from those serving the *ECs*. Moreover, instead of the notion used in [ZRG11], the visual representation is aligned to the approach in [CT12] that differentiates technical and business interactions. For the VNQ approach, the business interactions are focal.

The data augmentation is highlighted in pink. Specific values are omitted for illustrative purposes. Cardinalities refer to the relevant set of alternatives for the other side, eg, a customer can pick from many *CPs*, while *CPs* can market their service to millions of *ECs*. The difference in cardinalities for the *Access NSP* market in the United States of America (USA) and in the UK arises from the available choice for the respective *ECs*,

which differs from region to region. Due to the sketched chained network transmission and revenue sharing model, the sending *NSP* temporarily bears the transmission costs. A compensation is ultimately achieved by the Internet access and transmission agreements of *ECs* and *CPs* that create a functioning value stream.

Despite the omission of specific values in Fig. 2.6, the VNDM's readability is limited, which may cause problems for larger assessment cases. For this reason, we recommend the usage of separate data tables referencing to business interactions shown in a VNDM. Technically this information needs to be encoded in a meaningful way, which is addressed in the next section.

2.4.2.2 Value Network Graph (VNG)

♣²⁶ Irrelevant of the visual representation, technically a VN is still a graph. This section will transfer the VNDM idea to graphs, which requires some degree of formalism. For a higher degree of details, cardinalities are encoded by alternatives (instances) of individual business interactions. This allows a richer perspective on VNs, as, for example, competing offers with different prices can also be encoded. This represents a clear progress over the representation in [ZRG11].

For a clean handling of the data, any unassigned costs, which were considered in the VNDMs approach, are eliminated. Unassigned costs have to be assigned to business interactions. If no better knowledge exists, they can be distributed proportionately among business interactions according to their market volumes. In addition, the representation of bilateral trades is represented through two unidirectional business interactions.

The remainder of this section will also translate the economic terminology, eg, business interactions, used by VNDMs to a formal language required by graphs, eg, edges. The formalisms are primarily captured in hereinafter definitions:

Definition 1 (Directed Labeled Graph). *The ordered pair $G = \langle E, R \rangle$ of the form $R \subseteq E \times E$ —where E is a finite number of vertices and R a finite number of edges—is a directed graph. A graph of the form $G = \langle E, R, \mathcal{L} \rangle$ where $\mathcal{L}$ is set of labels (a set of tuples formed from the set of cost labels $\mathcal{C}$, the set of value labels $\mathcal{W}$, and the set of type labels $\mathcal{T}$) and $R \subseteq E \times E \times \mathcal{L}$ is a **tuple** of the form $r_i = \langle e_1, e_2, L \rangle$ (with $e_1, e_2 \in E$ and $L \subseteq \mathcal{L}$; each r_i in R is labelled) is called directed labeled graph.*

Based on the definitions of Directed Labeled Graphs and their consistency—cf. Definition 2—a Value Network Graph (VNG)—cf. Definition 3—can be formulated.

Definition 2 (Consistency). *Let $G = \langle E, R, \mathcal{L} \rangle$ be a directed labeled graph, then G is called non-empty, iff $|E| \geq 1$, $|R| \geq 1$, and $|\mathcal{L}| \geq 1$. We call a non-empty graph G consistent, iff every node e_i is connected with the rest of the graph, using such that G cannot be subdivided into subgraphs without removing a single edge r_i from G .*

²⁶ This section reworks the definitions used in [Zwi11] to introduce a series of specific labels required for quantifying VNs in a subsequent step.

Definition 3 (Value Network Graph (VNG)). *Let $G = \langle E, R, \mathcal{L} \rangle$ be a directed labeled consistent graph, then G is a graph for a VN, if each node $e \in E$ represents an entity of the VN, each edge $r \in R$ corresponds to a business relationship between two entities of this VN, and label L represents a set of tuples of the form $l = \langle c_j \in C, w_j \in \mathcal{W}, t_j \in \mathcal{T} \rangle$ corresponding to the properties of each instance j of r . Such a graph G is denoted as Value Network Graph (VNG) G_{VN} for the labelled VN.*

Each edge $r \in R$ is further characterised by a series of properties being detailed by the specified labels in Definition 4.

Definition 4 (VNG Labels). *Let $G = \langle E, R, \mathcal{L} \rangle$ be a VNG, then for each $L \in \mathcal{L}$ for each vertex instance and each linked edge $r \in R$ a tuple of the form $\langle c, w, t \rangle$ has to exist where each $t = t^r$ for each vertex instance—thus, the resource type is only determined by the edge r rather than its edge instances. In particular, the set $C \subseteq \mathcal{L}$ is called set of cost labels, iff each $c \in C$ describes the cost of a business relationship, ie, an edge $r \in R$ of G . The set $W \subseteq \mathcal{L}$ is called set of value labels, iff each $w \in \mathcal{W} \subseteq \mathcal{L}$ describes the value of a business relationship of G . Conversely, iff each $t \in \mathcal{T} \subseteq \mathcal{L}$ describes the type of a business of G , then $\mathcal{T}$ is called the set of type labels.*

A vertex (node) instance in a VNG corresponds (cf. Definition 5) to an instance of an entity (ie, role)—eg, Samsung as an instance of the mobile phone manufacturer entity. We further assume that the labels of each instance i of e_i (representing market players of similar structure, eg, two bakeries) may not be identical with other labels. Hence, each edge holds a set of labels (eg, values, costs) where a representative value may be used for visual reference.

Definition 5 (Vertex Instance). *Let $e \in E$ be a vertex of the VNG $G = \langle E, R, \mathcal{L} \rangle$, then ψ is the set of vertex instances of e . $|\psi|$, as number of instances for e , is called the cardinality of e . In our case, at least one vertex instance $i \in \psi$ exists for every e in E .*

Relationship instances represent vertex instances that can maintain a given business relationship. Typically, the set of vertex instances is mirrored to the set of edge instances (ie, relationship or business interaction instances). This provides a higher flexibility for modelling competing offers and substitutes.

2.4.2.3 Atomic Graph Operations

Graph operations aligned to the business context, can provide valuable information on noteworthy forces characterising a VNG and its future evolvement. Especially atomic graph operations, as syntactic manipulations of a prior version of the graph, systematically reflect each minimal difference induced by an industry change.

By essentially building on the given definitions, all possible atomic manipulations of an existing VNG G , ie, alterations of a single edge r or node in G , are investigated around any arbitrary vertex e (representing an entity). On that basis, more sophisticated special cases are considered.

2.4.2.3.1 Edges.

Addition: A new *incoming edge* $r_{in} \in R$ for node e is considered as new structurally-different supplier for any process or demand of e . This may be caused by *product developments*, eg, new products or product evolutions requiring different relationships, or new sources of *substitution* for existing edges (one edge to e may directly or transitively be substituted). Hence, this covers new business opportunities for e , which may help to diversify the value streams (ie, positive influences on *bargaining powers*). A new *outgoing edge* r_{out} for node e implies selling the service of e to new customers—ie, a customer relationship is formed to another node. Such relationships do not cover industry rivalry where one entity instance (node instance) gains market shares at the expense of another instance. Beyond that, such an augmentation operation may also be triggered when adding *substitutes* for any existing relationship, eg, caused by product development or adaption to demand variations. Associated edge deletions or alterations, which may mitigate the positive influence on bargaining powers, are covered in the following procedures.

Deletion: The opposite effects occur when relationships are deleted (removed from G). Business relationships (or alternatives) could dissolve, hence making e more dependent on other relationships²⁷, ie, increasing bargaining powers of partners. In association, the overall market volume may be reduced (ie, downsizing the entity role), which would reduce the *entity size* and transitively affect the negotiation power.

Alteration: Often *volume*, *price* or *cost* levels may change amongst others entailed from product cycles, industry rivalry (product supply), regulation, product demand, or other market factors. An alteration of the resource type, eg, the exchange of customised goods in place of money, may due to its structural influence on G be most fundamental²⁸. Besides any semantical shifts, the *fungibility* of exchanged resources is affected, which may strongly influence a firm's positioning in G . Thus, a marker representing structural dependencies referring to the resource dependency is necessary in order to illustrate the compatibility of a relationship to other vertices in the case of alterations. Another change can be seen in the *redirection* of edges, which may result from Business Model evolutions (eg, also due to ongoing vertical integration). A redirection may mean diversification of business partners (improving *bargaining powers*) or may, on the other hand, represent a *substitution* for another mechanism. Beyond that, properties of relationships and their instantiations may change. While from an inter-entity perspective such shifts are captured by bargaining powers, ie, a relationship may become more important for an entity, intra-entity shifts may

²⁷ This also correlates to the concepts from 'VN Analysis' listing such effects as risk from structural dependencies – <http://valuenetworks.com/public/item/236476>, last accessed: April 18, 2012

²⁸ The alteration, in this case, is the replacement of an existing relationship by a new but different relationship. If one relationship can easily be redirected to any other node (eg, money), but the new relationship cannot anymore, a substantial structural dependency is created that needs to be considered.

rather indicate changes in entities' market sizes (*entity size* as aggregate of *volumes* assigned to relationships) and *competition* situation.

Granularity: Changes of the granularity (as changes in abstraction levels or perspectival change) may be caused, for example, by variations in the degree of vertical integration. The very same VN may be formulated on a very fine- (G^f) or coarse-grained level (G^c ; also due to vertical integration of roles), leading to important implications on the analysis of VNs: whenever several edges are unifiable to a single edge, it may be considered as single business relationship. Nevertheless, the aggregation of all edges between two nodes, may not preserve semantical considerations (eg, covering individual business streams with varying fungibility). Thus, quantitative analysis approaches should not enforce the aggregation beyond the elimination of redundancies in G . Due to the static nature of graphs and, thus, VNQ, product developments or dynamics vertical integration are not covered over time. This may be quantified by comparing G with a redrawn alternative G' , eg, due to scenarios to be tested against each other, or by introducing new vertically integrated edges/entities in parallel (also see Section 2.4.2.3.2). Market dynamics models, like presented in [CT12], could also be used to assess market evolutions of this sort.

2.4.2.3.2 Nodes. Changes in edges or their configuration may have a strong impact on the nature of G , however, altering the configuration of a node may entail even more severe consequences. We assume that every node depicted in G has to be related to the common outcome(s) of the VN represented by G , eg, the video consumption of an end customer, establishing an indirect relationship to any other node in G . In all other cases, a node is considered to be an external node beyond the scope of G (and the corresponding VN). From this definition, we can entail that at least one edge is affected whenever a node is added or deleted.

Addition: Whenever a node e is added it also requires the addition of one or more edges connecting it to other nodes and to G in order to avoid trivial cases. Accordingly, effects of adding edges influence the positioning of e in G and all other nodes being involved.

Deletion: If the VN can exist without the supply or demand of e , a *product evolution/development* has occurred or a *substitution* has been possible, which justify the deletion of e from G . The effects for all other nodes losing their edges to e also have to be taken into account.

Alteration: Common graph operations may alter the number of vertex instances—ie, a change of the *cardinalities*. In particular, when the number of instances of a node e rises (or conversely decreases), a *market entrance* has occurred (or a market exit conversely), ie, affecting *competition* aspects and the utility of business interactions.

Whenever all instances of e go bankrupt or decide to quit, the role of e at least temporarily vanishes from G . Such symptoms obviously represent a special case, which may be caused by passive product evolution (potentially targeting the whole VN) or by (misplaced) public regulation, eg, special taxes for communications services that may render a business unprofitable or new legal restrictions cannot be met.

Perspectival Change: Similar to edges, also nodes may be represented on different levels of granularity or degrees of vertical integration, which obviously affects the involved edges.

2.4.2.4 Eclectic Graph Operations

The prior assessment has concentrated on atomic graph operations, which affect a single vertex e (eg, creating e or adding an edge r to e) or edge r . This section analyses changes on a broader *eclectic* scale. Graph operations such as *substitutions* completely bypassing initial nodes have been ignored so far. For example, when the postal service²⁹ is replaced by an electronic mail system, not only the final delivery of the mail may be rendered obsolete, but international delivery agreements or billing processes (cf. Figure 2.7) are affected alike. Those graph operations eventually affect the value streams (or value creation processes) within a VN. In contrast to atomic graph operations, any such eclectic alteration is out of the control of intermediary nodes. Hence, source and sink nodes of prominent value streams carry more responsibilities for the entire VN and have more capabilities of moderating their dependency on the VN. When considering Internet and network services, more complex value streams are a reality and are, thus, potentially affected by eclectic changes alike.

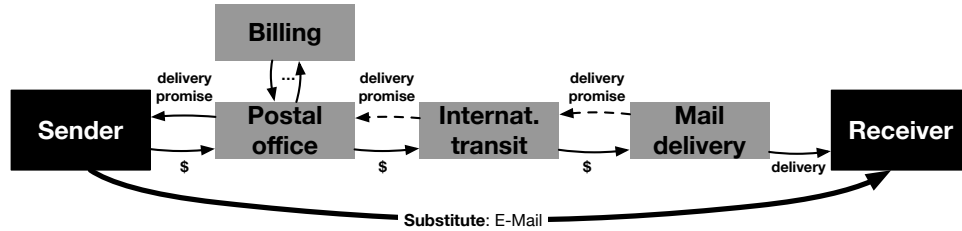


Figure 2.7: Postal service being substituted by an e-mail service.

Of relevance is the proximity or distance to *critical nodes*, which are end customers (or final nodes providing resources)—ie, *customer ownership*—, or critical assets and raw materials. Typically, there exist customer nodes within a graph G , as a representation of a particular VN. In manufacturing industries, critical assets would be placed at the source of a value stream, but in telecommunication the most critical assets may be the Edge NSP's Internet access infrastructure, which moderates the access to end customers.

²⁹ The mail delivery system works on the premise that a payment and the corresponding promise of delivering the e-mail triggers a series of actions.

The weaker the relationship of a node to critical nodes, the higher the difficulty to influence the entire VN. This lack of control is paired with an increased risk of being substituted, which accumulates along the path to the end customers: by the assumption of similar substitution probabilities for each edge, the likeliness of being replaced strongly increases with the distance. On the other hand, it could be claimed that a focal position of a firm could represent a similar herd of interest, ie, only one node being fully interconnected with all others (direct access to resources, customers, high bargaining power ratios, etc.). These focal positions are often referred to as “intermediary roles”, which take a bottleneck role that moderates demand and supply from a central position. For example, Internet-based travel agencies such as Expedia³⁰ provide an intermediary service, which does neither “consume” nor “produce” the service or product in a classical sense. The assessment of bottlenecks in the VN is covered by the factor of *bargaining powers* (as known from Porter’s five forces), known from atomic graph operations. The access to critical nodes is captured by the new *customer ownership* factor.

Such kinds of risks form additional factors that are out of the scope of a single edge or node but may affect the entire VN. They are subsequently aggregated by an additional *risk* factor.

2.4.3 VNQ Requirements & Assumptions

♣³¹ This section first briefly recapitulates on the identified forces around the graph-based assessment. As identified forces are considered to be mandatory for a VN dependency model, they form explicit **requirements** for the VN dependency metric, as systematised in Section 2.4.4.

Entailed from potential modifications of VNs, seven forces have been obtained that potentially cause the illustrated graph modifications: *bargaining power of suppliers*, *bargaining powers of customers*, *substitution*, *industry rivalry*, *market entrance*, *resource type*, and other *risks*. Obviously, the majority of these factors (the first five factors) relate to Porter’s five forces on firms [Por79]. As these factors seem to have a high influence on the VNG, we argue that they also have to be considered as forces on entities (nodes) in the inter-firm case of VNs.

Concerning the remaining two factors, on the one hand the *resource type dependency* (fungibility of resources) may result from swapping an edge by a different type of edge. Such a force has already been considered in [ZRG11; ZR12] entailed from the concepts of [BC08], and should be considered in the approach presented later on. On the other hand, additional *risks* including the examples of customer ownership, broader-scope substitutions, and product evolutions of other entities may represent more complex special cases with sufficient influence on the VNG G . Unmentioned risks like financial risks or unstable business partners may pose further disutility, which is to be considered in assessing the

³⁰ <http://www.expedia.com>, last accessed: 20th May 2016

³¹ The list of assumptions below are based on our prior publication in [ZR12].

strategic positioning of a firm in a VN as well. Beyond risks, the *size of entities* (also relating to market volume or market shares) requires a handling in graph-oriented quantification approaches, such as VNQ. The latter force represents a scaling factor to other forces. In the context of economies of scale (such as for the market of Internet services), the size of companies may have considerable influence on entities' dependency on a VN.

For the construction of our entity dependency model we further formulate a series of **assumptions and interpretations** followed throughout the remainder of this chapter:

Monopoly. Our model relies on the monopolistic decisioning case fully eliminating resource scarcity and competition for resources. Hence, all alternatives are assumed to be available when bearing a certain investment effort. Competitive games for resources may have to be considered in future work.

Stationarity. The core of the quantified VN options are assumed to remain stationary throughout the analysis, ie, prices, costs, players (entities and their instances), relationships, substitutes, investment costs and other properties remain constant. Hence, all substitution options can be explicitly modelled a priori.

Dynamicity. The dynamicity of real-world VNs is captured through modelling the entrances and exits of players, as well as through the analysis of substitutes.

Utility. Added values are used in the sense of utility gains—based on the Von Neumann-Morgenstern utility theorem [VM44]—already incorporating values, costs, risks, chances, and further impairments.

Time. Utilities are standardised for a predefined *time span* with a clearly defined *starting time*. This adaptation is required to integrate investment costs that may be required to switch to a substitutive alternative sufficiently. This concept strongly relates to depreciation mechanisms used in accounting.

Complete Knowledge. The information required in order to allow a sufficient dependency reasoning process (complete knowledge) is taken as given.

Individual Rationality. We assume an individual rationality of demanding customers (ie, only positive utilities for buying a product may trigger a purchase) and involved entities, ie, customers and entities always choose the individually best VN option.

Entity. The identification of the best VN configuration is subject to the perspective of an entity or a group of entities minimising its dependency towards the VN.

2.4.4 VNQ Approach

Based on the assessment dimensions (Quality, Scale, Risk) and cases (Suppliers, Customers) defined in Section 2.4.1, the VNG formulation from Section 2.4.2.2, and the requirements and assumptions listed in Section 2.4.3, the quantification process can be systematically derived, as depicted in Fig. 2.8.

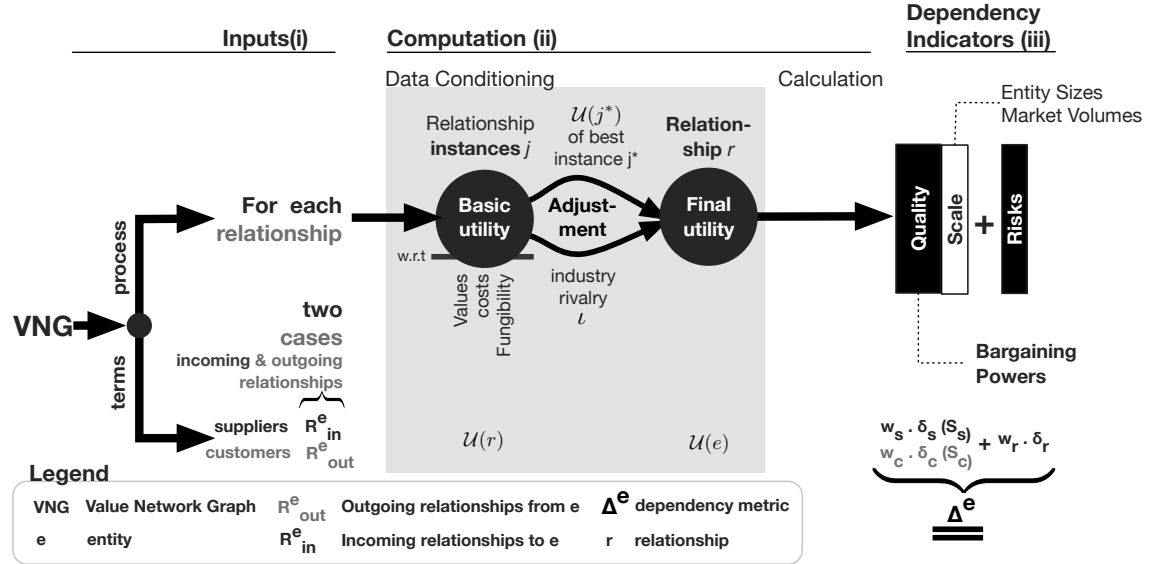


Figure 2.8: Big picture of VN dependency quantification.

Three main phases are identified: (i) inputs for the quantification in the form of a VNG and visually represented as VNDM; (ii) the computation phase of the actual dependencies on the basis of the input information, ie, a graph; and (iii) the dependency indicators as outcomes of the computation.

Phase (i). The first phase will distinguish two cases, as previously introduced: *suppliers* (set of incoming relationships R_{in}^e) and *customers* (set of outgoing relationships R_{out}^e), which separately trigger phase (ii).

Phase (ii). Following the graph-theoretic assessment in [ZR13] and the analysis dimension presented in Section 2.4.1, the three dependency indicators δ_s , δ_c and δ_r plus the orthogonal scaling factor S are used and assigned to the quantification dimensions of the second phase as follows:

Quality & Scale:

Quality: The bargaining power of suppliers δ_s and of customers δ_c form two dependency indicators on the quality dimension. These indicators primarily assess the utility distribution of business interactions (relationships) of an entity.

Scale: The entity size or market volume acts as orthogonal scaling factor S , which is integrated in δ_s and δ_c resp. The integration of the otherwise separate dependency indicator into $\delta_{s/c}$ is enabled through the technique discussed in Section 2.4.5.2 and Section A.1.

Risks:

The third independent dependency indicator δ_r is formed around a risk assessment, also covering eclectic effects within a VNG.

Other forces listed in Section 2.4.3 will be used in the quantification of the three dependency indicators. The indicators $\delta_{s/c}$ are designed from a top-down perspective around supply-side or customer-side relationships of an entity e as follows:

1. **Calculation of dependency indicators** w.r.t. utility distributions of relationships
2. **Calculation of utilities** w.r.t. the following subprocesses:
 - **Basic utility** calculation w.r.t. initial resource values, investment and operations costs, and the fungibility of the resource type
 - **Adjustment of utilities** w.r.t. industry rivalry

Phase (iii). On this basis a final dependency metric Δ^e for each entity e relative to all other entities of the VNG is formed, which aggregates the supply- and customer-side results of phase (ii) to a final dependency metric for each entity e relative to all other entities of G :

$$\Delta^e = w_s \cdot \delta_s(S_i) + w_c \cdot \delta_c(S_i) + w_r \cdot \delta_r \quad , \quad (2.3)$$

where $\delta_{s/c}(S_{s/c})$ refers to the inclusion of the orthogonal scaling factor in the computation of the dependency indicators $\delta_{s/c}$, and $w_{s/c/r}$ represent weighting factors for the integration of all independent dependency indicators $\delta_{s/c/r}$ in Δ . For all weighting factors $w_{\{s,c,r\}}$ no empirical parameterisation exists from literature to the best of our knowledge. They can, however, always be normed such that

$$\left(\sum_{k=s,c,r} w_k \right) = 1 \quad (2.4)$$

holds.

While Δ^e represents the ultimate result of the VNQ process (where a Δ^e close to 1 refers to a high dependency of entity e on this VN and close to 0 refers to the opposite), individual indicators may still be strategically considered in order to reveal sources of dependencies for an entity e . The calculation of dependency indicators δ composing Δ^e are discussed in Section 2.4.5.

2.4.5 Dependency Indicators

Subsequently the detailed calculation of the dependency indicators is introduced along the defined VNQ dimensions.

2.4.5.1 Quality

The quantitative metric for the quality dimension centrally refers to the bargaining powers of each entity e in the VNG. Following the discussed top-down presentation of dependency indicators, hereinafter first the indicators themselves are introduced:

1. **Dependency from Bargaining Power of Suppliers:** The bargaining power of suppliers (demanded resources by entity e) is calculated by analysing the value distribution among all requested resources (incoming relationships R_{in}^e ; a subset of all relationships of R of the VNG), of an entity e , eg, an NSP. Value in this context refers to the utility $\mathcal{U}$.

The utility of a single incoming relationship $r \in R_{in}^e$ to entity e is denoted by $\mathcal{U}(r)$. The bargaining powers can then be measured by analysing the value distributions of relationships, eg, on the basis of information entropy [Sha48], as indicator for disorders in systems. In this thesis it is, however, proposed to use the Gini coefficient calculation, also used as inequality measure in information theory (as metric for information impurity [Bre+84]), and further augments it to accommodate the extension given in [Yit83]³²: For this purpose, we define $p_r^{e,in}$ as fraction between the utility $\mathcal{U}(r)$ of one incoming relationship $r \in R^{e,in}$ and the sum of values from all relationships of entity e ,

$$p_r^{e,in} := \frac{\mathcal{U}(r)}{\sum_{k \in R^{e,in}} \mathcal{U}(k)} \quad . \quad (2.5)$$

³² The presented solution is a newly designed Gini coefficient calculation that works in analogy to the design presented by [Yit83]. Due to the different original Gini coefficient calculation form, the calculation is substantially different in this thesis.

Note that the larger $p_r^{e,in}$, the more dependent is the entity on a single other entity in the VN. The Gini coefficient itself is then the sum of $p_r^{e,in}$ values for all relationships $r \in R_{in}^e$ to the power of v , where $v = 2$ is defined for the case of δ_s ³³,

$$Gini(R^{e,in})^{v=2} := \sum_{r \in R^{e,in}} (p_r^{e,in})^2 \quad . \quad (2.6)$$

Finally, an appropriate dependency indicator δ_s^e is calculated in relation to all other entities $k \in E$, where E is the set of all entities of the VNG and $e \in E$ is satisfied, with the maximum inequality $Gini(R^{e,in})$ in the VNG,

$$\delta_s^e := \frac{Gini(R^{e,in})^{v=2}}{\max_{k \in E} \{Gini(R^{k,in})^{v=2}\}} \quad (2.7)$$

Further we establish the following notions for the rest of the thesis: R are without further information all relationships of a VNG, R^e are all relationships of entity e , and R_{out}^e and R_{in}^e are all outgoing and incoming relationships of e resp. For these sets the following holds:

$$R^{e,\{in,out\}} \subseteq R^e \subseteq R \quad . \quad (2.8)$$

2. **Dependency from Bargaining Power of Customers:** The bargaining power of customers (to whom resources are sold) can be represented by the analysis of value distributions of provided resources to other entities. The customer-side bargaining dependency δ_c^e thus is calculated exactly like δ_s^e , where $R^{e,in}$ is replaced by $R^{e,out}$ (the outgoing relationships of entity e).

Continuing in the hierarchical top-down introduction of VNQ, the calculation of bargaining powers is continued on the level of competing offers. Competing offers refer to the assessment of each instance j of relationship r in order to determine the **utility of a relationship** r , ie, $\mathcal{U}(r)$.

The utility $\mathcal{U}(j)$ of a relationship instance j is subject to its initial valuation³⁴ φ_j , operation costs c_j , investment costs i_j , and the fungibility³⁵ $\xi(j)$ of its resource type³⁶. Investment costs i_j are used in the sense of depreciation costs for the period under consideration reflecting the required investment costs, eg, for switching to a substitute or investing in new production techniques. $\xi(j)$ is calculated by the fraction of entities to

³³ The utility calculation will later on also require a different Gini calculation when inspecting relationship instances.

³⁴ The initial value typically corresponds to the revenue resulting from a business interaction that does not yet consider costs or the relative attractiveness of an offer in comparison to competing offers.

³⁵ This may also be referred to as modularity [Bal07]

³⁶ Business interactions exchange some kind of resource in a loose sense. Even the promise of a service delivery or intangible assets such as knowledge gain are a sort of resource.

which the resource of relationship instance j (exchanged via a relationship of an entity e) could be supplied or sold $k(j)$ out of the number of all entities $|E|$ of the VNG—cf. (2.9) where $0 \leq k(j) \leq |E|$ holds:

$$\xi(j) := 1 - \frac{k(j)}{|E|} . \quad (2.9)$$

The utility $\mathcal{U}(j)$ for the j^{th} best alternative for relationship r of entity e is then formulated by

$$\mathcal{U}(j) := y(j) \cdot \left[\varphi_j - c_j - i_j \right] \cdot \left[1 - \delta_f^e(j) \right] , \quad (2.10)$$

where $y(j)$ is an exponential spreading factor that ranges from $\frac{1}{n_r}$ to 1 (n_r is the number of instances for the relationship r for the case of outgoing customer-side relationships $r \in R_{out}^e$, and vice versa for $r \in R_{in}^e$) of the form

$$y(j) := \frac{1}{n_r} \cdot e^{\frac{\ln(n_r)}{n_r-1} * (n_r-j)} , \quad (2.11)$$

with

$$j \in \{1, 2, \dots, n_r\} .$$

The factor $y(j)$ (as introduced in [ZR12]) ensures that more likely picked alternatives—ie, the higher the utility, the higher the probability of being picked—are considerably valued more preferable. Hence, an interval $[(1/N), 1]$ is used where 1 is assigned to the best alternative $y = 1$, while $1/N$ is assigned to the least attractive option n_r for relationship r . The spreading factor is an essential prerequisite for analysing the value distribution of alternatives using a Gini coefficient. Despite the Gini coefficient's capability to measure value distributions or disorder, it does not cover the sequential preferability as targeted by $y(j)$. Following the *Gini* extension of [Yit83], the spreading factor $y(j)$ may also be replaced by a weighted *Gini* ^{v} coefficient, which overweights inequalities in the high utility range whenever $v > 2$ ($v = 2$ equals the classical Gini coefficient). According to our assessment in Section A.1, we believe that v realistically ranges in $[1, 3]$, and recommend $v = 3$ for the replacement of $y(j)$ and the default Gini with $v = 2$ for all other usages. The corresponding Gini coefficient with $v = 3$ is formed as follows:

$$Gini(e)^{v=3} := \sum_{j \in \{1, \dots, n_r\}} [p(j)]^3 , \quad (2.12)$$

where j remains a relationship instance³⁷ of r , which leads to the industry rivalry force ι ,

$$\iota_c(r) := 1 - \frac{Gini(r)^{v=3}}{\max_{k \in R_{out}} \{Gini(k)^{v=3}\}} , \quad (2.13)$$

³⁷ The ordering of relationship instances is irrelevant for the Gini coefficient calculation.

where R_{out} refers to all outgoing relationships known in a given VNG with $r \in R_{out}$. The case of incoming relationships R_{in} works in analogy to the described case and yields $\iota_s(r)$.

By the definition of $\mathcal{U}_r(j)$ (the utility of j^{th} best alternative³⁸) and the industry rivalry forces $\iota_{\{in/out\}}$, the utility $\mathcal{U}(r)$ for r can now be calculated by

$$\mathcal{U}(r_{in/out}) := \mathcal{U}_r(1) \cdot (1 - \iota_{\{in/out\}}) \quad . \quad (2.14)$$

The $\mathcal{U}(r_{in/out})$ result for each business interaction r (relationship) concludes the utility assessment and provides the required input for the calculation of the bargaining powers.

2.4.5.2 Scale

The dependencies resulting from entity sizes (market volumes) may be captured best by the market volumes of each entity e , eg, calculated through the share of resources exchanged with customers (S_c^e) and suppliers (S_s^e) from the market volume of the entire VN. If represented on the level of market volumes³⁹, the market share $S_{s/c}^e$ for entity e can be formed recursively in the interval $[0, 1]$,

$$S_{\{s/c\}}^e := \frac{\sum_{r \in R_{\{in/out\}}^e} \mathcal{U}(r)}{\sum_{k \in E} S_{\{s/c\}}^k} \quad , \quad (2.15)$$

and directly integrated in the computation of $\delta_{\{s/c\}}^e$ (cf. (2.7)), in analogy with Weighted Theil indices [WL03], ie, the result of (2.7) is multiplied with $S_{\{s/c\}}^e$,

$$\delta_{s_c}^e(S_{\{s/c\}}^e) := \delta_{s_c}^e \cdot (1 - S_{\{s/c\}}^e) \quad . \quad (2.16)$$

Through the multiplier $(1 - S_{\{s/c\}}^e)$ the influence of the bargaining power is scaled according to the market volumes, where lower market volumes increase the pressure and higher volumes do the opposite. This unweighted integration of $S_{\{s/c\}}^e$ in $\delta_{\{c/s\}}$ is realised in analogy with the decomposable Gini coefficient form by Lerman and Yitzhaki [LY85] (cf. Section A.1). Theoretically, a separate dependency indicator could be created, which would, however, create a redundant step (in analogy to (2.7)) in the calculation.

2.4.5.3 Risks

Whenever an entity (or one of its instances) is low on profitability, there may be the risk of a market exit in considerable time, eg, due to individual defaults or market failures. Such exits, hence, create risks for business partners. Comparable risks may originate from

³⁸ Highest *negated* (unnegated) worth, ie, $-\mathcal{U}(j^*)$ for any instance j^* of $r \in R_{in}^e$ ($r \in R_{out}^e$, resp.)

³⁹ Whenever the problem scope is simplified to a single service provisioning case, the entity size needs to be calculated from the market volume shares retrieved from external figures. In this case, the factors δ_s and δ_c are of heuristic nature where it is assumed entities proportionately distribute their business activities according to the profitability of their business interactions. In other words, the higher the profit for one interaction, the proportionately more business is conducted in this area.

many intra-firm sources, hardly exogenously controllable, as well as from the access to critical resources. Entities may then propagate risks via *role dependencies*⁴⁰ to others, eg, slowdowns of value flows. Transitively, substitutions of one entity's relationships may even render relationships with higher distances obsolete (eg, from neighbours of neighbours; 2 hops). Hence, customer ownership measured by the distance (ie, hops) in the graph is critical. Graph modifications may however especially have difficulties in capturing (eclectic) factors spanning over multiple nodes while existing risk modelling concepts may provide suitable complementary means. In turn, the utility calculation as described in (2.10) may be revised correspondingly. Nevertheless, future VNG variations may be even capable of directly capturing more of such risks by extending the set of required input information.

As a starting point, we will further briefly revisit context-independent risk analysis methodologies being utilised for systematically capturing risks for the networking IC use case. Required adaptations are established on the basis of meaningful BM Design parameter crossings with VN concepts such as presented in [Bal07; Ghe12]. By turning this to a relative dependency indicator on par with $\delta_{\{s,c\}}$, a computational integration is achieved in the overall quantification concept. This adaptation is complemented by a brief thematisation of risk perspectives on the utility calculation.

2.4.5.3.1 Risk Models. There exists a broad range of risk analysis concepts in the literature [NW05; Whi95] that range methodologically from quantitative mechanisms like Monte Carlo simulations [Whi95] to qualitative alternatives. Although approaches targeting e-commerce applications are known [NW05], to the best of our knowledge no analysis technique specifically dedicated to network services is available. Thus, more generic risk evaluation frameworks need to be adapted for our purposes:

While Ericsson Risk Management Evaluation Tool (ERMET) [NJ04] provides a qualitative risk evaluation tool for firms along five categories (financial, business control, hazards in the surrounding, hazards at the site and business interruption handling risks), they may be rather suitable for the production line-centric enterprises. A straightforward alternative is provided by the Composite Risk Index (CRI)⁴¹, which assesses risks in terms of *risk impact* v and its *probability of occurrence* pr . To ease the assessment of the impact, typically a rating on a defined scale, eg, between 0 and 1, is applied. Due to its simplicity and flexibility, the CRI is of widespread use today, providing a solid basis for risk evaluation. However, the specification of risk assessment categories, eg, the risk of fire in the production hall, have to be defined on a per-use basis. We will subsequently construct such an adaption for the networking industry on the basis of feasible VN crossings with BM design parameters.

⁴⁰ <http://valuenetworks.com/public/item/236476>

⁴¹ "Equivalence of Risk: A Mathematical Approach". <http://www.system-safety.org/conferences/2011/papers/Equivalence%20of%20Risk%20-%20A%20Mathematical%20Approach.pdf>, last accessed: 20th May 2016

2.4.5.3.2 Crossing of BM Design & VNs. By explicitly crossing BM design parameters with VN considerations, this part aims at revealing untargeted risk categories with network impact⁴². For this purpose, the analysis template from Section 2.3, mainly building on the works of [Bal07; All00] and reported in [CEL14], will be used to derive relevant new risk categories. In particular, Ballon *et al.* differentiate [Bal07] *control parameters* (VN Parameters & Functional Architecture Parameters) and *Value Parameters* (Financial Model Parameters & Value Proposition Parameters) which are partially already mirrored in the used quantification approach. Each parameter category from Section 2.3 that is marked in the list below with *Risk Category* will need expert estimations for optimal coverage, as they are insufficiently integrated into the current quantification process. Categories marked with *Partly Covered* will need an additional estimation for full coverage while all other categories are fully represented in the quantification process.

- **Value Network Parameters**

- **Combination of Assets:** While heartbeat entities are widely covered by the entity size aspect and bargaining powers of suppliers and customers (as indicators for interaction ties with others), the access to critical assets as well as more eclectic access to critical resources is not entirely captured by $G \rightarrow \textbf{Risk Category}$.
- **Vertical Integration:** This kind of integration relates to a change of granularity that are captured by comparing multiple VNG representations ($G, G', \dots$) or coexisting entities in $G \rightarrow \textbf{Partly Covered}$.
- **Customer Ownership:** Customer ownership is not directly reflected by our VNQ approach, as clients and the access to them are not addressed by specific concepts at that point $\rightarrow \textbf{Risk Category}$.

- **Functional Architecture Parameters**

- **Modularity:** The modularity is not target by VNQ, besides clear interlinkages to the covered interoperability of resources $\rightarrow \textbf{Risk Category}$.
- **Distribution of Intelligence / Exchange of Value:** This factor is insufficiently covered by VNQs $\rightarrow \textbf{Risk Category}$.
- **Interoperability:** Interoperability is assumed in telecommunications due to high standardisation efforts, but may be eligible for special business case considerations. This category is captured by the resource fungibility, which expresses the degree of exchangeability of resources during a trade $\rightarrow \textbf{Covered}$.

- **Financial Model Parameters**

⁴² Comparable to [All00], we assume that the risks in the value creation process, ie, yielding a utility outcome, are directly accommodated for by individual entities.

- **Cost (Sharing) Model:** The underlying cost and cost-sharing models are reflected by costs and relationship properties. Specific cost-sharing modalities may not be captured in a VNG → *Partly Covered*.
- **Revenue (Sharing) Model:** This case is identical to the cost (sharing) model → *Partly Covered*.
- **Value Proposition Parameters**
 - **Positioning:** The VNG has no particular understanding of the product, target customers and underlying novelties, competencies etc. → *Risk Category*
 - **Customer involvement:** This factor is not covered by VNQ → *Risk Category*
 - **Intended Value & Perceived Value:** This factor is mirrored in the difference between revenue and cost in VNGs. The graph does, however, give no reasoning why this value is accumulated → *Covered*.
- **Network indicators (complementary subset only) [All00]**
 - **Sequence (see Combination of Assets):** Despite the encoding of each business interaction, the main value flow paths are not highlighted in the VNG. The overlap of paths at a single entity may represent the position of a critical asset in the VN → *Risk Category*.
 - **Reciprocity:** The reciprocity of business interactions is not covered, as bargaining powers are separately assessed for the customer and supplier case and may, thus, involve different entities → *Risk Category*.
 - **Stability:** The stability is covered on a per-entity view. Each entity's bargaining power of suppliers and customers is affected by the diversification of business interactions. With a low degree of diversification (or concentration of value on one business interaction) these values will increase to the disbenefit of the studied entity → *Covered*.

The mentioned risks of product development, ie, BM adaptation, is reflected in the value proposition category by the *Positioning* (Product/Service Delivered) item. *Customer ownership* is introduced explicitly as a VN risk category. Hence, the discussed open issues can be sufficiently targeted by the stated set of risk categories, denoted by F , when they are sufficiently quantified (or estimated).

2.4.5.3.3 Computation. In order to facilitate the exchange of risk models according to context-specific requirements, additional risks are integrated as relative factor in the interval $[0, 1]$ where 1 is the highest and 0 the lowest risk. For this purpose, the sum of risk estimations $\sum pr_f \cdot v_f$ (where the impact v_f is, for example, estimated in the interval $[1, 5]$ and the probability pr in $[0, 1]$) for a single entity over all risk categories $f \in F$ is scaled in relation to the highest value in the VN,

$$\delta_r := \frac{\sum_{f \in F} pr_f^e \cdot v_f^e}{\max_{k \in E} (\sum_{f \in F} pr_f^k \cdot v_f^k)} \quad , \quad (2.17)$$

and can afterwards be integrated as additional dependency indicator in the overall calculation, as demonstrated in Fig. 2.8.

2.4.6 Revenue Sharing

When considering an ecosystem where multiple entities collaborate or cooperate, as common practice in the Internet ecosystem, revenue sharing mechanisms play a significant role. Revenue sharing refers to the task of “fairly” distributing collectively generated revenues among participating entities (players). Today’s revenue sharing mechanisms are typically rooted in cooperative game theory where the by far most notable concept is the Shapley⁴³ value [Rot88]. The concept of Shapley values is grounded on the measurement of the outcome reduction, ie, decreasing profits or increasing costs, when a player leaves an alliance. For example, when a series of factories work together to reach a common goal and one factory opts to start producing something different, how much of the revenue or profit is lost due to the lowered production performance. These concepts have been applied to the cases of coalition formation involving task divisions [ZR94], supply chain profits [NS08], power transmission planning [CKY98], multi-domain NSP federations [ABV12], and many other cases. In telecommunications, other revenue sharing concepts inspired by game theory exist alike, such as by He & Walrand in [HW05].

The subsequent analysis concentrates on a dedicated revenue sharing mechanisms on top of the Shapley value. This mechanism will take the VN dependencies of entities (players) into account, which are inferred from our proposed VNQ technique. An acceptable revenue sharing mechanisms requires such knowledge to assess the longer-term strategic positioning of firms fairly. In other words, suppose all players in the market are deciding to cooperate, then based on their strategic importance for a VN their “fair” payout from the entire revenue should be calculated.

We will introduce our novel revenue sharing mechanisms by first working out the shortfalls of today’s solutions in Section 2.4.6.1. On this basis, we will utilise both the state-of-the-art in literature and our advancements in quantitative VN analysis in order to augment the highly recognised Shapley value in Section 2.4.6.2.

2.4.6.1 Problem

From a value creation perspective, each entity (and its instances) of a given VN is required for producing the targeted output. This fact may draw the illusion of equality within the VN, which, however, can be disproved by the specification of specific dependency values

⁴³ Alvin E. Roth and Lloyd S. Shapley were awarded the “The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2012” for “the theory of stable allocations and the practice of market designs”.

for each entity. As given before, these dependency values characterise the nature of how entities are integrated in the VN, ie, their strategic positioning, mainly based on their business interactions but also on the basis of their capability to generate profit. In co-operative models, eg, alliances, bargaining mechanisms often rely on such inequalities of entities—such as through the prominent Shapley value [Sha53] as well-studied revenue sharing mechanism—for facilitating cooperation for achieving joint goals among heterogeneous (and competing) entities. As a result, these inequalities need to be made explicit, which is by definition not the case for unprocessed VNs.

Thus, we can acknowledge the following insufficiencies of Shapley value-based revenue sharing when transferring it to the context of VNs.

- All entities in a VN are considered to be equal while this may not reflect the real world. The removal of a single player (to assess its value properly) would require a systematic capturing of alternatives and their value for the VN in order to understand the utility loss for the remaining VN.
- While the Shapley value complies to superadditivity on coalition level, it implicitly also relies on superadditivity of relationships, ie, the distribution of value among business interactions of a single entity does not affect the worth of the entity. However, bargaining powers of business partners may substantially influence the market position of individual firms and, hence, may affect the “fair” share of revenues. Such forces are reflected in the common business practice of diversification in order to moderate the risk exposure within a VN. This perspective is unfortunately not covered by the plain Shapley value.
- A differentiation in the entity, eg, the Access NSP entity in the USA with for example AT&T on the instance level, is not foreseen. While the replacement of a single business alternative may be feasible with a more fine-grained instance-based perspective, on the VN level each entity is required for achieving the targeted output.
- Shapley values provide a short-term snapshot analysis, which is not capable of reflecting long-term market convergence tendencies. From our point of view, the integration of longer term issues such as risks (eg, of financial failure) deserves some integration in revenue sharing considerations.
- The number of alternatives may increase/decrease, ie, market entrance or market exit, which is not covered so far.
- Business interactions may be substitutable on medium- to long-term, which is not covered so far.
- The strategic relevance of exchanged resources or goods are not covered, eg, low fungibility may lead to disadvantages in price negotiations on medium- to long-run. This may not apply to business interactions of comparable quality, but may be of

interest for low latency services (LTE), for example, in comparison to higher latency alternatives (Universal Mobile Telecommunications System (UMTS)).

- Shapley values do not provide means for parameterising the quality of a relationship. Hence, it is implicitly assumed that such an assessment has been conducted before.

Solutions of the specific technique provided in Amigo *et al.* [ABV12] have not yielded results within the core [SS69]. The core is a concept in cooperative game theory describing the set of winning imputations that are not dominated by any other imputation. Imputations are allocations that are individually rational and efficient. In the case of Amigo *et al.* [ABV12], imputations refer to feasible bandwidth allocations based on conditions specified by the corresponding NSP.

2.4.6.2 Solution Overview

In response to the stated shortfalls of revenue sharing approaches, the present section introduces an alternative that refines the well-known Shapley value. As VNQ is capable of modularly characterising the economic nature of cooperative relationships, it can mirror the expected dynamicity of a VN with respect to medium- to longer-term strategic positioning of firms (ie, market and bargaining powers). The linkage of VNQ’s and Shapley value’s capabilities especially qualify for a revised revenue sharing mechanism for the following set of reasons:

- VNQ comes from graph-theoretic roots [ZR13], comparable to network topologies and representation of business collaboration in general.
- VNQ can spot the medium- to long-term convergence perspective.
- VNQ understands the concept of an “entity” (a role in the VN) and its competing instances (“firms”) with a single conceptual framework.
- Shapley value provides a well-recognised mechanism for fairly sharing revenues, eg, within an alliance.

While the intermingling of both concepts seems promising, the following obstacles need to be resolved in the quantification process:

- A VN relies on all entities; thus substitution of entities needs to be covered in the assessment of the utilities—as proposed in [ZRG11];
- Monetary payouts need to be assigned to entity instances.

Aligned to the specified obstacles we propose a dedicated **two or three stage quantification algorithm**, which leverages on the utility-based VNQ approach and the Shapley value concept:

1. Calculation of the VN dependency for each entity e and derivation of the associated power ρ^e as opposite of their dependency on the VN,

$$\rho^e := (1 - \Delta^e) \quad ; \quad (2.18)$$

2. Calculation of the payout to each entity according their ρ^e , ie, importance for the VN;
3. *Distribution Among Instances (Optionally)*: The distribution among instances needs to be computed, in addition, if and only if several instances are required to satisfy the role of one particular entity or a particular business interaction. Contrary to the prior stage, a classical Shapley value can be used (based on assessing the one-by-one removal instances). This is possible as only properties of a single business interaction are temporarily altered while the rest of the VN remains as it is.

The details of stage 2 is introduced in Section 2.4.6.3.

2.4.6.3 Calculation

In the context of VNs for modern Internet economies, see the definition by [Ben07], equally capable entities that can easily overtake the role of any other entity do not exist. On the other hand, all entities of the VN are required for retaining the industry in its current form. In other words, dropping a single entity may not be possible in many cases without disrupting the entire VN. However in this case, an entity's worth among *prima facie* equals can be captured by calculating the power of entities as the opposite of their dependency on the VN (cf. Equation (2.18)), which leads to an augmented calculation of the Shapley value's payouts φ .

By computing a weighted and adapted Shapley value (cf. Equation 2.22), the payouts⁴⁴ $\varphi(VN)^e$ for each entity e in the VN are derived. In contrast to classical coalitions, all actor entities in the VN are supposed to be required cooperators in the coalition T (the VN is the coalition in this case), which is reflected in the simplified term A (cf. Equation 2.19).

$$\begin{aligned} A &= |T|! \cdot \frac{(|E| - |T| - 1)!}{|E|!} = \\ &= (|E| - 1)! \cdot \frac{(|E| - (|E| - 1) - 1)!}{|E|!} \\ &= (|E| - 1)! \cdot \frac{1}{|E|!} = \frac{(|E| - 1)!}{|E|!} = 1/|E| \\ &\quad \text{when } |T| = (|E| - 1) \quad . \end{aligned} \quad (2.19)$$

⁴⁴ φ was previously used to indicate the initial value of a relationship. In analogy, it represents the payouts in this context without considering costs, and at the same time follows the common notion for the Shapley value payouts.

In lieu of profits varying with the entity-side composition of the VN, ρ^e can be used for weighting the importance of each entity in the VN in relationship to the overall profits (cf. B in Equation 2.20).

$$B = \rho^e \cdot \sum_{k \in E} \pi^k, \text{ (Weighting profit gain; normalisation)} \quad (2.20)$$

where π^k represents the profit for entity $k \in E$.

As the terms A and B imply a normalised output for all $|E|$ entities, a denormalisation C of payouts (scaling the profits to the original units) is required in the last time, which establishes a link between Δ^e and the VN's overall weights for all entities E (cf. Equation 2.21)

$$C = 1 / \sum_{j \in E} \rho^j \text{ (Denormalising payouts)} \quad (2.21)$$

The resulting payout for each entity e from the overall $\mathcal{R}$ in the VN,

$$\varphi(VN)^e := A \cdot B \cdot C \cdot \mathcal{R}, \quad (2.22)$$

may be useful for a discrepancy analysis with the achieved profits in practice. Whenever the achieved profits surpass the calculated payouts for $e \in E$, this may indicate indicator an increasing medium- to longer-term risks of jeopardised profits for e (lower margins or substitution). In the opposite case, the e 's market power may be leveraged to increase its margin in the future.

Each entity may also assess VN configurations by comparing expected payouts in various variants, which may lead to stability issues requiring a game-theoretic assessment in the future.

2.4.7 Implementation & Evaluation

The introduced approach has been implemented and evaluated. The evaluation process is twofold: in this section, we will concentrate on a sensitivity analysis while a separate section will practically apply this approach to use cases in telecommunications.

2.4.7.1 Tool Development

A prototypical implementation has been created with the programming language Python, which is available at <https://github.com/pzwickl/PZ-Thesis/tree/master/VNQ>. The implementation is provided with an MIT license for the academic and non-academic usage, extension and commercialisation.

Each VNG is represented as a graph and accessed via dedicated wrapper methods, which aim at assuring syntactic consistency.

The computation itself consists of several phases:

1. **Input:** The central input information is a complete and directed *networkx* graph, which allows multiple edges between any pair of nodes (ie, *MultiDiGraph* in *networkx*). This graph can be specified using a Python script or deserialised from persisted graphs (class *Persistence*). Deserialisers for JavaScript Object Notation (JSON) and Bytecode are delivered with the prototype. Due to the modular design of our implementation, other serialisers can be added with low effort. Due to the integration of the *networkx* library, import and export of graphs in the Graph Modelling Language (GML) format are available too. Those graphs may have to be augmented for satisfying the information needs of our technique.
2. **Calculation:** In the calculation phase the graph has to be traversed twice (see file *vnq.py*): once for assessing and aggregating relationship instance information to reasonable utility representations and another time for calculating and aligning the VN dependency indicators on top of this utility information.
3. **Output:** Textual output is provided, which encompasses the VN dependency metric Δ^e for any entity e (node in the VNG), fundamental descriptive statistics and interpretations, and a series of consistency checks (ie, a lazy reporting mechanisms, which provide a warning in the case of incomplete, suspiciously configured, are untypical input graphs). Visual representations are further created as a snapshot, which can be stored as a picture. The code is further created in a way to allow the export of data in the future, which also allows for utilising external graph visualisation libraries.

Unit tests have been included in the implementation.

2.4.7.2 Sensitivity Analysis

In 1994, Hamby [Ham94] has collected a series of sensitivity analysis methodologies varying by their “sensitivity ranking”. From the subset of quantitative methods, the efficient “one-at-time” and related “sensitivity index” methods have been broadly applied. While the first method only quantifies local sensitivities “relative to the point estimates chosen” rather than capturing parameter distributions [Ham94], the second more precisely measures the output difference between the minimum and maximum of each independent variable.

For this purpose, the sensitivity index ς will be analytically derived for all dependency indicators in order to characterise their impact on the overall outcome. The Python code served as basis for creating analytical bounds. The indicators δ_s , δ_c and δ_r , as well as the entity size factor, share some sensitivity characteristics:

The weighting factors w_s , w_c and w_r are exogenous forces. The respective entity size S provides an additional orthogonal scaling factor, which is directly derived from the market volumes encoded in the VNG and assesses separately. The dependency indicators range in the interval $[0, 1]$, so that the minimum is 0 and the maximum is 1. Due to factual minimum of 0 (see Table 2.1) the sensitivity index always equals the maximum value. The

Table 2.1: Sensitivity index for $\delta_{s/c/r}$ and $S_{in/out}^e$.

Dependency Indicator	Bounds & Index	
	Min	Max = <i>Index</i> (ς)
δ_s	$w_s \cdot 0 \cdot (1 - S_{in}^e) = 0$	$w_s \cdot (1 - S_{in}^e)$
δ_c	$w_c \cdot 0 \cdot (1 - S_{out}^e) = 0$	$w_c \cdot (1 - S_{out}^e)$
δ_r	$w_r \cdot 0 = 0$	w_r

maximum values for the assessed indicators are most directly influenced by the respective market volume factors S . Within the maximum and minimum ranges, the functioning of the Gini coefficient characterises the input sensitivity of the respective indicator's output value. The input is affected by $\iota_{s/c}$ indicators. The only exception is the indicator δ_r , which is subject to the chosen risk assessment framework.

The sensitivity $\varsigma(\iota_{s/c})$ of ι_s and ι_c are primarily characterised by the Gini coefficient—see (2.6) for the case $v = 3$ rather than $v = 2$ as for the other dependency indicators:

$$\varsigma_l(\iota_{s/c}) := Gini^{v=3}([\mathcal{U}_1, \dots]) \quad , \quad (2.23)$$

where $[\mathcal{U}_1, \dots]$ represents the list of utilities for all tested relationship instances, on which basis the value distribution is assessed using the Gini function. By design, the considered value range of $Gini^3$ is $[0, 1]$, which leads to a relevant index value of 1 (ie, $1 - 0 = 1$)—corresponding to the usage in Table 2.1. These two dependencies are further affected by the fungibility of business interactions, which alters the utility values.

The factor ξ is characterised by Fig. 2.9. Whenever no relationship to other entities exists for entity e , the maximum dependence of 1 results. The minimum of 0 is returned for the case of existing relationships to all other entities. For any intermediary position between these extremes a linear input-output relationship exists.

Thus, we have the following input-output relationships for the discussed dependency indicators:

2.5 The Video on Demand Value Network

This section provides a use case analysis of our VNQ technique for a rudimentary VoD scenario as reported in [CEL14] (qualitative assessment) and [ZR15] (quantitative assessment). More complex flavours may explicitly include entities like CDNs, content distributors or content owners, and other resource types such as intangible assets. This section focuses on an end-to-end network transmission involving QoE guarantees and using com-

Table 2.2: Input-output sensitivity of dependency indicators.

Factor	Relationship	
	Endogenous	Inherited
ξ	Linear in $[0, 1]$	
$l_{s/c}$	Non-linear, based on the Gini coefficient in $[0, 1]$	Linear influence from ξ
$\delta_{s/c}$	Two-stage non-linear, based on the Gini coefficient in $[0, 1]$	Inherited non-linear influence from $l_{s/c}$
δ_r	Depends on the chosen risk framework, normed in $[0, 1]$	

mon IC techniques such as transit⁴⁵, peering⁴⁶ or CDN⁴⁷—also see the definitions and business considerations in [Nor10] and [FP711].

2.5.1 Qualitative Assessment

•⁴⁸ Subsequently, several VN flavours are constructed for a VoD scenario, which are comparable to the scenarios empirically tested in Section 3. The following cases are distinguished: 1) NSP-centric, 2) VoD-platform-centric, 3) fully distributed and 4) combinations thereof.

In case 1), the NSP has created a VoD platform in order to extend its classical Internet access business to complementary content-related services. This case is supported by the current trend to bundle services, ie, triple- (often fixed line broadband, telephony, television) or quadruple-play offers (additional cellular services), in the European telecommunications market.

As depicted in Fig. 2.9, the central entity is the Access NSP that “owns” the customer relationship (“End user” entity). The content business has been specified in the form of the “Broadcaster” (live content; if offered) and “Film Studio” entity that may act as “Content Distributor” or “Content Producer”. The distributor is the holder and moderator of the distribution channel for the content created by producers. Before the merger of *Disney*⁴⁹ and *Pixar*⁵⁰, Disney was the distributor of *Pixar*’s content. By taking a commission on

⁴⁵ Transit refers to the paid and hierarchical IC. Typically NSPs are distinguished in Tier-1, Tier-2 and Tier-3 providers where the degree of IC decreases from Tier-1 to Tier-3. Tier-1 providers have a direct IC to any other Tier-1 and resell transit services to NSPs on lower tiers.

⁴⁶ Bilateral and bidirectional IC or multilateral mutual IC via dedicated peering points. Only the rare case of paid peering involves fees.

⁴⁷ CDNs replace IC needs by nearby storage facilities.

⁴⁸ *This section summarises our results initially published in [CEL14]. The description concentrates on the use case analysis on VoD (correlating to the efforts in other chapters).*

⁴⁹ <http://disney.com/>, last accessed: 20th May 2016

⁵⁰ <http://pixar.com/>, last accessed: 20th May 2016

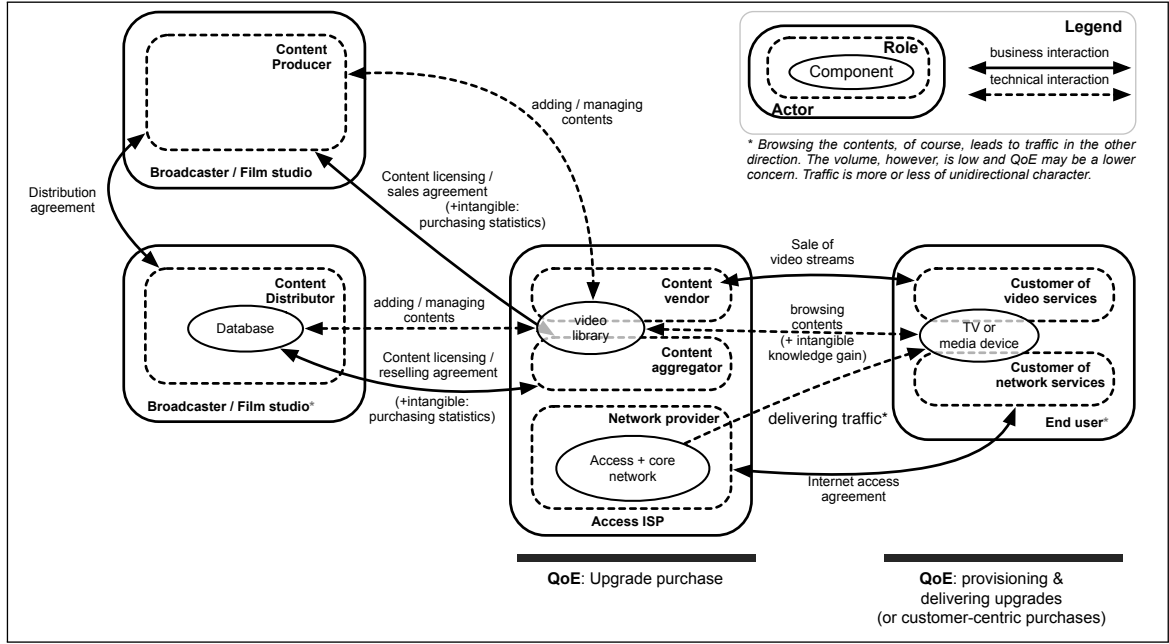


Figure 2.9: VoD VN centralised at the Access NSP [CEL14].

Pixar's revenues, *Disney* played an intermediary role by opening its broad distribution network and agreements to others like *Pixar*. In more modern scenarios, content may be directly sold via the platform, which is the common practice in the music business since the emergence of *iTunes* (and comparable offers).

While the content is collected by the “Content aggregator” and sold by the “Content vendor” role, the “Network provider” delivers the content. Hence, the vendor sells a promise to deliver the contents arranged by the aggregator via a network provider. Together these roles are handling the entire purchasing process with the end user. In this NSP-centric, all roles are played by the same actor, which represents a powerful concentration of business interactions in this network.

The central NSP entity handles the choice of appropriate QoS (network and video) and the transition to QoE. Apart from that, the users, as receivers of the network and video quality, are QoE-aware and -sensitive.

Contrary to the first case, the VoD platform-centric configuration—case 2)—does *not* integrate the network transmission in the entity that provides the platform services (cf. Fig. 2.10). Nevertheless, the content is still aggregated centrally at the VoD-platform. The QoE moderation competence is divided between the NSP in the access network and the VoD platform provider. This division may create coordination issues, as users may have difficulties in attributing deficiencies to the supplier of the network or the video service. Side payments (pink edge), where platform providers directly pay NSPs for their efforts, are a measure to regain control at a central entity, but may introduce NN concerns. Apart

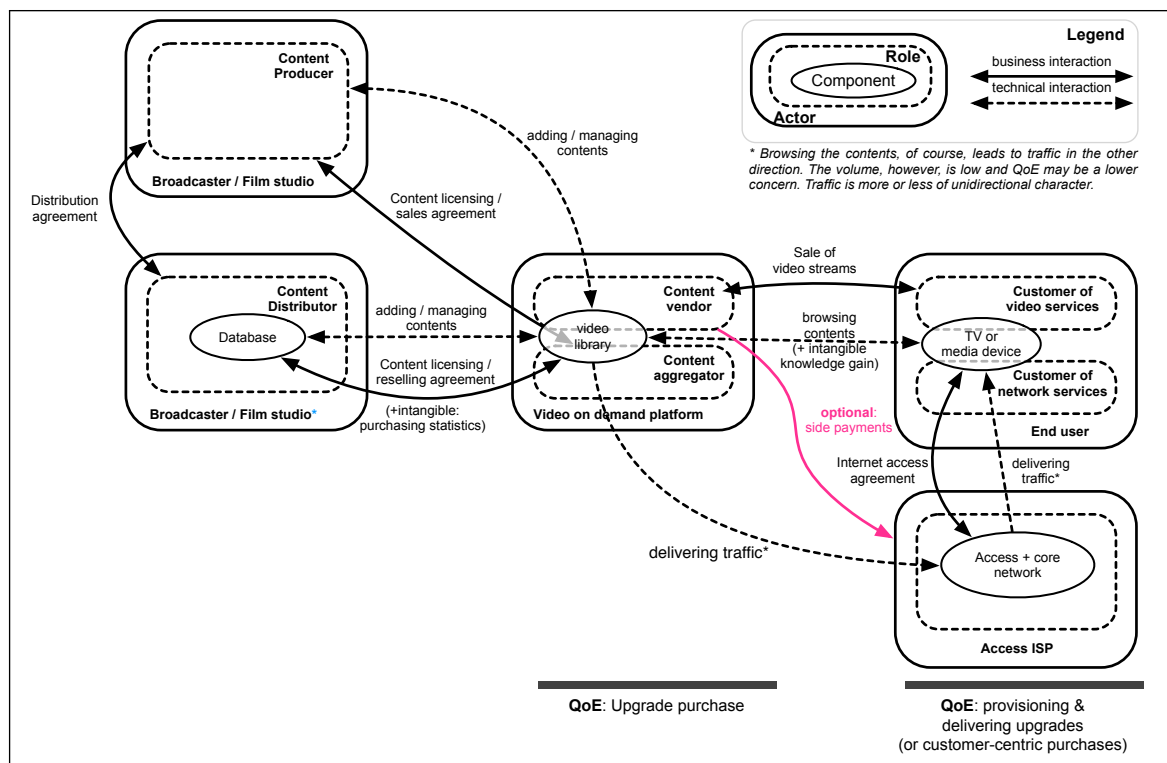


Figure 2.10: VoD VN centralised at the VoD platform [CEL14].

from access network NSPs, further NSPs provide transit or Internet access services in other regions, eg, when transmitting contents from the USA to the UK a provider in the USA, a Transit NSP are required in addition to the Access NSP in the UK.

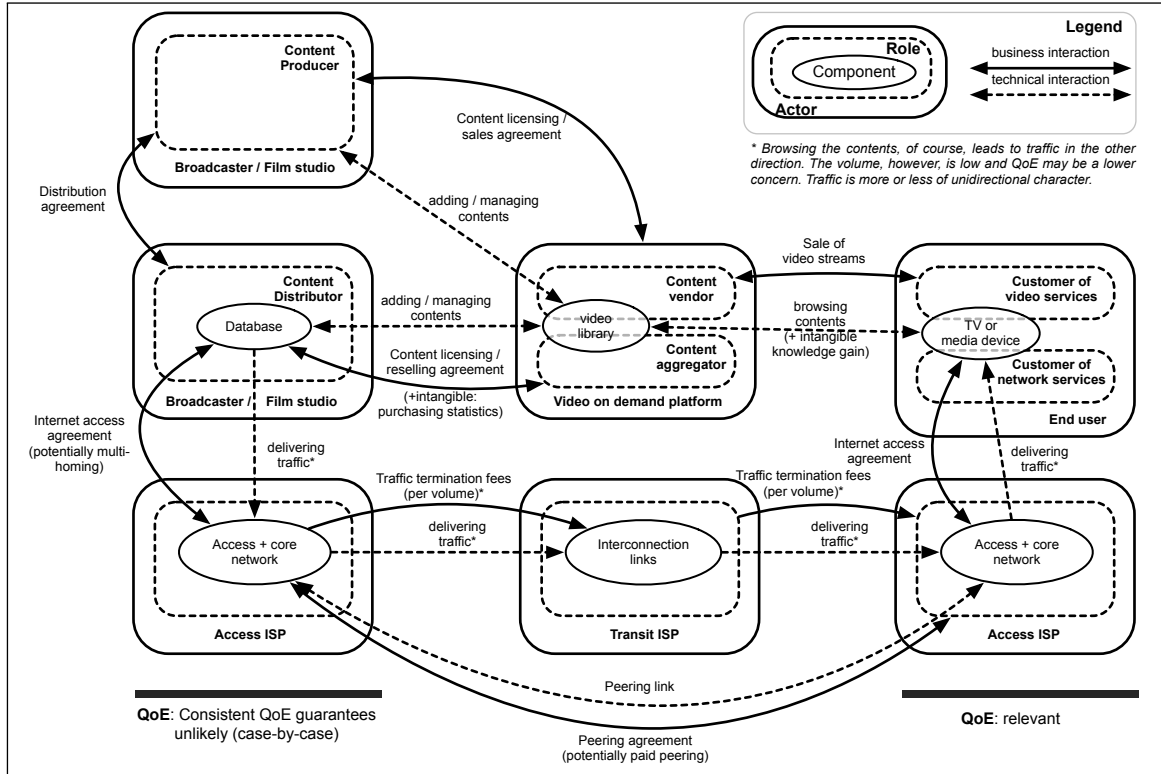


Figure 2.11: VoD VN in fully distributed form [CEL14].

When even more distributing the competencies in the network, the fully distributed case is obtained—case 3)—that is sketched in Fig. 2.11. In this case, the content is directly streamed from distributors to end users, while the platform only handles the sales of the content. In addition to those mentioned before, this case adds coordination issues regarding the content and its quality. While in the previous case a single entity could take control of the video QoS and the contents (probably taking care of encoding the raw video material to match the needs of the streaming platform), the service experience may substantially vary from content to content in the revised case. For example, some external video service may be down, the connection may be less reliably than usual or the desired quality level may not be available.

Hereinafter, the previously provided qualitative analysis template will be applied to the mentioned cases:

- **Value Network Parameters**

- **Combination of Assets / Heartbeat roles [All00] [Bal07]:**

- * Access to and control of contents, end customer, client devices and network services are well distributed among players in fully distributed case. With the degree of centralisation also the access to these assets is centralised. Thus, the NSP-centric configuration has the highest concentration of asset controls in a single entity, ie, the Access NSP. Hence, the NSP may “own” the customers.
- * **Heartbeat roles:** The VoD platform is the bridge between available contents and clients willing to purchase items from the store, ie, content lookup, consumption and publishing services. Thus, when removing the platform, the entire VN can be disrupted. Despite the criticality of removing the NSPs, alternative distribution channels like Digital Versatile Discs (DVDs) or Blu-rays may mitigate the economic consequences for the network.

– **Role Contribution and Structure:**

- * **Central entity:** The VoD platform has the most ties in the content value flow (see “Sequence”), while each NSP has comparable ties in the network transmission value flow (see “Sequence”). However, the Access NSP holds the critical interface between both value flows, which may be materialised when overtaking the platform role as well.
- * **Value creation structure:** Value is mainly created in the relationship between the end customer and the VoD platform.
- **Vertical Integration** increases with the degree of centralisation—first, through concentration of the content business at the platform and second through the merger of platform and Access NSP entities.
- **Customer Ownership:** see *Combination of Assets*, as customers are assets in this network.

• **Functional Architecture Parameters**

– **Modularity**

- * Both CPs and platforms may reciprocally be exchanged by competing offers. The number of large content distributors and VoD platform providers may, however, be low in practice. Hence, content distributors and VoD platforms need each other. The customer can choose from the best current offer.
- * Customers may be replaced by new customers (effort!) in the same market or the providers may extend their business to new markets (other video segments, geographical regions, etc.).

– **Distribution of Intelligence / Exchange of Value [All00]:**

- * Content intelligence is held by the VoD platform. The platform possesses critical information on their customers liking and habits, on the one side,

and the available content, on the other side—ie, the platform is an intermediary between end customers and content distributors.

* QoS intelligence refers to the decision on the network transmission (bandwidth, loss, delay, jitter, etc.), but also on the video encoding side (eg, bitrate, codecs). Depending on the chosen VN configuration, QoS intelligence may be distributed among Access NSP and the content distributors (fully distributed case) or the VoD platform (VoD-platform-centric), or entirely concentrated on the NSP (NSP-centric).

* QoE intelligence, however, is likely only possessed by the platform or the Access NSP, as content distributors do not have a direct client contact, ie, the platform provider has to forward feedback on the quality perceptions to content distributors. Side payments from VoD platforms to Access NSPs may also hand over the network QoE control to the platform providers.

- **Interoperability:** Both network services and video content can be regarded to be more or less interoperable due to the existence of broadly accepted standards. However, the interface between VoD platforms and the end customer will most likely be proprietary, eg, dedicated applications, devices or platform credentials.

- **Financial Model Parameters:** In analogy with other digital content stores such as Apple’s iTunes⁵¹ or Amazon Kindle⁵², the revenue sharing modalities are dictated by the platform entity. The high bargaining powers of the platform may further be increased when side payments are used, as the platform will overtake the role of the central distributor of monetary flows. In classical Internet access regimes, the Access NSP is the distributor of revenue to other NSPs in the transmission chain.

- **Value Proposition Parameters**

- **Positioning:** VoD platforms are positioned as a substitute to video rentals, live TV channels, DVDs/Blu-Rays and (illegal) file sharing platforms.

- **Customer Involvement:**

* The VoD platform as an intermediary entity is a two-sided market, which requires active business partners on both sides, ie, customers and providers of contents. Hence, subsidies of one side may play a major role in order to bootstrap the market.

* Besides editorial feedback, reviews and purchasing behaviours, the user involvement is limited.

- **Intended Value & Perceived Value [All00] (Network Indicator):**

⁵¹ <http://www.apple.com/itunes/>, last accessed: 20th May 2016

⁵² <http://www.amazon.com/kindle-store-ebooks-newspapers-blogs/b?node=133141011>, last accessed: 20th May 2016

- * The VoD platform provides a unified interface, which bridges the interests of customers (finding and consuming content) and distributors (finding interest customers). In other words, the platform provides simultaneous access to a critical mass of customers [EH95] and relevant contents.
 - * The professional moderation of QoE may be offered as additional unique selling proposition, ie, a clear product leadership.
 - * Past purchases may create lock-in effects for customers (ie, own content cannot be migrated).
 - * The perceived value will be attributed to the VoD platform, as experience-based, unconcrete and implicit network services are hard to brand.
- **Network indicators (complementary subset only) [All00]**
 - **Sequence:**
 - * **Content:** The value stream from the creation and distribution of content, via the aggregation at the platform, to the purchase and consumption by the end customers.
 - * **Transmission:** The provisioning of network streams in the network.
 - **Reciprocity:** Non-monetary (eg, video streams) and monetary exchanges create reciprocity in the entire VN. However, critical assets are not evenly distributed (see “Combination of Assets”) and the fungibility of resources varies, ie, a reciprocity of bargaining powers does not exist. Content for example may be hard to be sold to the right end customers without the help of powerful intermediaries. On the other hand, platforms may easily replace specific content by other material from this genre and at this quality. Thus, the monetary payment from platform providers may easily be redirected to other distributors. Despite the highly fungible monetary payments of end customers, they may still face difficulties replacing a particular VoD platform due to the absence of comparable streaming conditions, compatibility to clients and applications, content choice and access to prior purchases.
 - **Stability:** The degree of meshing in the VN is low, which is further reduced with increasing centralisation in the network (especially under an NSP-centric regime). In turn, the reliance on each role and entity, and especially on central entities is high.

Conclusion #2.2. *The VoD scenario aggregates at least two separate value flows in one VN, ie, the network transmission and the actual content sales. While both value streams have to coexist, their business models are only loosely coupled. Whenever a platform controls both value flows, a central bottleneck role with high bargaining powers results.*

Conclusion #2.3. *The VoD platform, as probable bottleneck role, represents an intermediary two-sided market for the service sales. While two-sided markets can be attractive, the market entrance phase is more challenging.*

Scalability effects may reduce the operations costs of the VoD platform. Thus, the access to a broad customer base is essential to staying competitive. In the case of an NSP-centric provisioning, the customer base is limited if insufficient external clients (no Internet service customers) can be attracted. A collaboration among NSPs or the service offering to customers of NSPs, will provide scalability gains (reduction of unit costs).

Conclusion #2.4. *If the NSP is also operating the VoD platform, strategic advantages can only be utilised, if target customers also include the Internet service customers from other NSPs, or if no cross-NSP competitor exists.*

2.5.2 Quantitative Assessment

Due to the difficulty of estimating all factors, the present analysis concentrates on the base case with and without side payments. Values for all available CDN offers as substitutes for transit network services, see the Transit NSP role in Fig. 2.12 and Fig. 4.8, are included in the list of alternatives for a particular business interaction.

As depicted in Fig. 2.12 (following a simplified visual notion w.r.t. [ZRG11]), the analysis starts with classical cascading payments for network services from the CP via supplying NSPs to the Access NSP (UK), which will deliver the video to the end user. The Internet access of end users is further arranged by separate contracts, which complies to current practice. Apart from the network transmission, CPs incur platform costs and compensation payments to holders of the copyrights. The entire service is ultimately purchased by end users who may directly consume the video on their TV. This scenario, thus, comes with distinct technical provisioning chains and business interactions. In other words, the CPs sell the video to the end user, but the actual transmission is performed via a chain of other entities.

For particular addresses in Manhattan (USA) and in Westminster (UK), 3 wholesale (fixed line) and 11 retail high-speed Internet offers have been identified, respectively. The high-speed Internet requirement is formulated in order to allow High Definition (HD) streams with high bitrates to be transmitted successfully to the customer. Likewise, 13 known Tier-1 providers and 23 CDN providers operating in the geographical area have been found.

Based on estimations⁵³, we have considered 1.7m VoD customers in the UK and have identified 10 VoD players still in service in the UK. Contrary to our first analysis of the VoD market in 2011 (cf. [ZRG11], only big players have remained in UK's VoD market, which has lowered the overall number of relevant entities.

We have set the market size of entities as follows:

⁵³ <http://point-topic.com/free-analysis/global-iptv-subscriber-numbers-q4-2013/>, last accessed: 20th May 2016

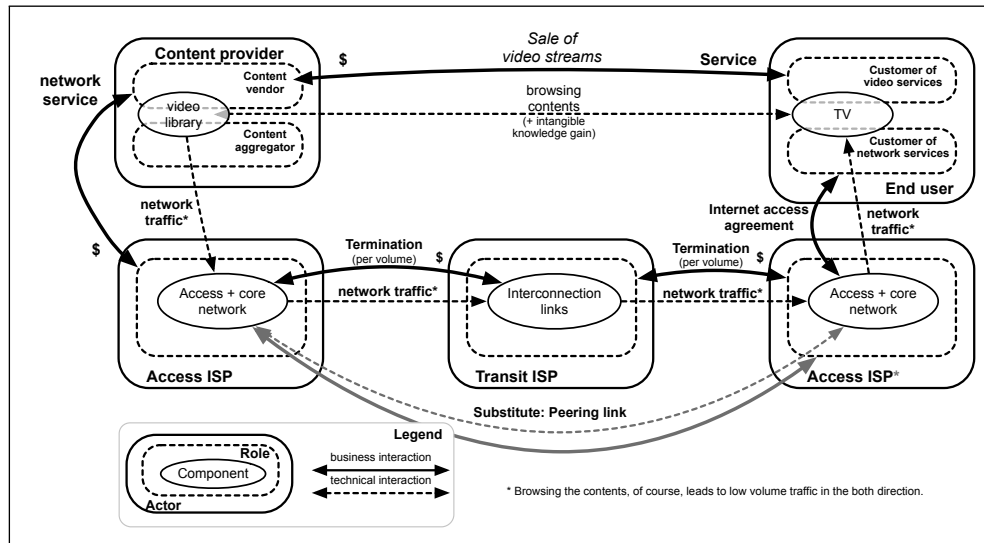


Figure 2.12: VoD VN without side payments

- £235m for VoD end users⁵⁴;
- \$21.08 bill. for VoD CPs⁵⁵;
- £32 bill. for Access NSPs in the USA and £4 bill. in the UK⁵⁶;
- Transit NSPs and CDNs at \$5.81 bill.

The service prices, eg, for streaming a single HD video or for transporting the associated network traffic volume, have been taken from public announcements of each provider, ie, their official web sites. Due to the inaccessibility of some information, eg, as some information may not be broadly published, we have approximated missing information by using the average price tag of known alternatives.

As transit pricing is contrary to CDN typically regarded to be a business secret, Dr. Peering's projection⁵⁷, of \$1.57 per Mbps per month for 2013 has been used.

Finally, the following static margin rates have been considered:

- 11.34% for all NSPs and CDNs⁵⁸;

⁵⁴ <http://www.screendaily.com/features/in-focus/uk-vod-all-eyes-are-on-this-market/5053562.article>, last accessed: 20th May 2016

⁵⁵ <http://www.marketsandmarkets.com/PressReleases/audio-video-on-demand-avod.asp>, last accessed: 20th May 2016

⁵⁶ Ofcom Communications Market Report (CMR): http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr13/icmr/ICMR_2013_final.pdf (p. 256), last accessed: 20th May 2016

⁵⁷ <http://drpeering.net/white-papers/Internet-Transit-Pricing-Historical-And-Projected.php>, last accessed: 20th May 2016

⁵⁸ Based on an average of BT's and Comcast's profit margins in March 2013, eg, http://ycharts.com/companies/CMCSA/profit_margin.

- 18.8% or 4.2% for VoD CPs with variable pricing⁵⁹ or with monthly subscriptions⁶⁰ resp.

Due to the absence of publicly available risk data or accessible expert ratings, we have assumed equal risk dependencies for all entities. For simplicity reasons, the risk value has been set to the maximum of 1, which adds $\frac{1}{3}$ to the overall dependency of every entity.

More details on the used market data and VNQ model parameterisation can be found in the source code or in the online spreadsheet⁶¹.

Table 2.3: VN dependencies in a VoD scenario.

End user	Content provider	Access NSP (USA)	Transit NSP	Access NSP (UK)
0.999	0.820	0.600	0.967	0.989

This input information has then been harmonised (eg, quantities, currencies) and properly coded as VNQ input file for our tool—see results⁶² in Table 2.3. The difference between the access NSPs in the USA and the UK results from their highly different entity sizes (ie, market volume) and degree of competition in their local markets. Apart from that, the access NSPs in the UK is only modelled as a retail provider, which does not offer wholesale solutions comparable to the entity in the USA. In practice, however, bigger NSPs will always try to target both customer segments. Apart from that, the result of the end user may be surprisingly high, which arises from the low degree of meshing in the VN, but especially from the strong competition among similar customers⁶³. In practice, end customers may collaboratively⁶⁴ protect their rights (and thus raise their market power) via consumer protection associations and may further be protected by some degree of regulation. The situation is different for the Transit NSP, which has limited potentials for integrating new kinds of business and has to face high competition from substitutes and market entrances⁶⁵. Thus, primarily the high dependency of the Transit NSP is concerning.

⁵⁹ Wuaki TV: <http://www.axonpartnersgroup.com/index.php/project/wuaki-tv/>, last accessed: 20th May 2016

⁶⁰ Netflix: http://ycharts.com/companies/NFLX/profit_margin, last accessed: 20th May 2016

⁶¹ “VNQ: VoD data” (spreadsheet): https://docs.google.com/spreadsheets/d/1Sf1A2IRmAHKuho_zE3hy13qwjrmIq4L4SsISQRH_EIc/edit?usp=sharing, last accessed: 20th May 2016

⁶² The results differ from the partial presentation in [ZR15] due to the update of currency conversion rates and an increased precision of the toolset.

⁶³ In a mass market, a customer may easily be replaced by another one.

⁶⁴ The VNQ technique of this thesis is rooted in the idea of competition for resources among players. Collaboration may, thus, mitigate market pressures that have accumulated plainly from a structural point of view.

⁶⁵ Especially the CDNs has appeared to be prosperous in the last years, which has further challenged the classical transit business. Market entrance barriers for CDNs seem to be limited, which has positively affected the number of competing CDN offers.

Conclusion #2.5. *In a limited VoD scenario, the Transit NSP has the most notable dependency on the VN. If configured as a separate entity, this position may need to be reconsidered in the future. The consumer business of small NSP can also be regarded to be structurally challenging in competitive markets.*

Conclusion #2.6. *Entity sizes strongly affect the market powers in a VN.*

The presented use cases will further be followed up when elaborating on the side payments concepts (known from the NN literature) in Section 4.5.4.

2.6 Concluding Remarks

This chapter has evaluated the general market design for the commercialisation of QoE-aware network services. Contrary to selling QoS-levels, QoE is an experience good, which is subject to many factors such as the used service or the context. A notable experience interaction exists with the price, where the so-called fixed point problem has been described. This problem illustrates the difficulty of transitioning from classical QoE investigations to monetary representations such as WTP.

The market design has implications on the transitioning between QoE and QoS. While consumer-facing offers are QoE-centric, in the backend the requirement for profit optimisation sets a focus on QoS and cost aspects.

Based on a novel quantitative VN analysis technique (as a complement to the previously applied qualitative analysis), ie, VNQ, we could illustrate high dependencies for Transit NSPs in a VoD case study. The access to critical resources (including the so-called customer ownership) substantially improves the market power of players. In the case of the VoD scenario, the Access NSPs and the content provides profit from such effects. Intermediary positions can strengthen the strategic position of entities additionally, which has not been assessed quantitatively. Graph-theoretic investigations have further illustrated driving forces for market power accumulation in VN configurations and have highlighted the requirement to assess risks associated with ecosystems properly, such as the QoE ecosystem, or particular entities. The significant influence of entity sizes (ie, an orthogonal scaling factor representing the market volume of stakeholders) on the strategic position of firms has further highlighted the importance of market-specific analyses and proactive regulatory measures to keep the market in balance.

As a result, from the market design we can follow a need to parameterise QoE and WTP, as well as to remodel the role of supplying services such as by the Transit NSP. A further challenge arises when comparing the content-neutral market design with non-neutral alternatives. Content neutrality⁶⁶ aspects are studied from several perspectives in

⁶⁶ Content neutrality refers to the equal handling of Internet content of any kind. A violation of content neutrality typically refers to a throttling, blocking or prioritisation of specific content types such as video traffic or contents between specific end points (providers). Application neutrality is a specific form of neutrality referring to the used application itself.

Section 4.5.

DEMAND & SPENDING

While one may speculate about causes, the limited practical success of charging [Tuf04] and revenue optimisation [KA03] approaches for QoS (see Section 2.1.1) and the debatable prosperity of network quality markets have stressed the need for a substantial paradigm shift. In response to the notorious underestimation of tariff gradation, this thesis postulates that economic and user aspects need to be fundamentally integrated with network quality architectures, ie, the solution requires a multi-disciplinary approach.

While this thesis is after more holistic solutions involving market frameworks, cooperation, competition and technological considerations, also the microcosm of network quality pricing deserves a thorough rethought. Taking user perceptions into account, a concept and formalisation of charging for QoE has emerged [HB13], but otherwise endeavours for quantitatively detailing charging for QoE have been surprisingly limited so far. As claimed in [ZRS14], this seems to be problematic for two particular reasons: first, there is a clear mismatch between objective QoS and subjective QoE, which is not even a priori conscious to users themselves. Network quality and over the top application such as VoIP are best classified as “experience products” [Nel70] requiring a posteriori perceptual evaluation [KR10]. In other words, users know whether they like or dislike a network quality after they have already experienced it. Second, there is another substantial mismatch between QoE and purchasing decisions of customers. Some users may appreciate the current QoE level but may nevertheless prefer to go for a free (or cheaper) alternative. Such effects may be subject to various motives, customer segments, socio-economic factors, demography, and many more factors, and are, thus, notoriously difficult to be approximated from QoE (which itself is already highly context-specific)—eg, see [WZR12] for approximated purchasing behaviours. Although both mismatches are acknowledged only rarely, they may affect the decision-making process integrally when optimising the service offerings to customers.

The present chapter, thus, shifts attention towards empirically measuring and modelling monetary expenditures for network qualities to derive appropriate charging strategies. The measurements strongly focus on the WTP concept, which is detailed later on. WTP figures can clarify the economic value of applying third-degree price discrimination for

network quality, next to conventional first-degree discrimination practices. Apart from this, the interrelation of quality perceptions and prices is assessed with the help of QoE figures with and without purchasing situation.

The obtained data are applied to supply-side optimisations in Chapter 4, eg, targeting the market entrance phase for suppliers, and to technical solutions in Chapter 5. Particular emphasis is put on the sustainable configuration of networks in order to motivate future investments in Next Generation Networks (NGN) technologies by NSPs.

This chapter primarily targets the following objective:

Objective: Empirical modelling of user demand & spending behaviour for quality-enhanced network services, including the scrutinisation of phenomena like cognitive dissonance and underlying purchase motives.

This objective can be subdivided into the following three contribution blocks:

- Clarification of the economic value of charging for QoE rather than QoS enhancements and implications on QoE ratings
- Empirical collection of WTP figures for quality-enhanced network services, including an understanding of underlying purchase motives
- Demand-side recommendations for the successful market entrance

The remainder of this section starts with a background on CS in the network context in alignment with [HKH05] (cf. Section 3.1). It continues with an illustration of the revenue and social welfare prospects of proper application-specific differentiation (cf. Section 3.2). On that account, the QoE charging ecosystem can be formed and parameterised by relying on a series of empirical trials (cf. Section 3.3). While corresponding QoE results are given in Section 3.4, a detailed assessment of purchasing motives is conducted in Section 3.5. Market entrance pricing strategies are addressed in Section 3.6 to specifically target the challenge of kick-starting a novel market for experience goods.

3.1 Internet Pricing & Customer Satisfaction (CS)

♣¹ Subsequently some elementary considerations on (Internet) pricing and CS, in analogy with Homburg *et al.* [HKH05] and aligned to fundamental definitions in Section 3.1, are used in order to define the boundaries for quality-differentiated access pricing and the benefit of adding QoE-awareness further in this section.

Since NSPs have expected a massive profit, fairness, congestion moderation and peak shaping advantages [Sen+13], a manifold of Internet pricing strategies have emerged,

¹ *Parts of this section follow and augment the materials presented in [HKH05], and transfer them to the context of telecommunications and QoE in particular.*

see [Sen+13; Tuf04] (including static, usage-based, time-of-day, dynamic, auction-based [MV95; LS98], QoS-based [Sen+13; Tuf04] etc. variants). Especially in the case where QoS-differentiation is applied, representing a form of third-degree price discrimination, the impact of CS on the success of Internet pricing deserves special attention.

According to [NS11], Internet pricing strategies can roughly be separated in pre-2000 and post-2000 phases. Early works, such as the concept in [Ful+98] that considers congestion modelling, concentrate on classical data-centric transport solutions. The later phase stronger concentrates on service class differentiation that models Internet pricing as optimisation problem involving QoS and technical constraints. With this, user-friendly tariffing [RHS03] and multi-class services [Kim06] have emerged, which is presented as a measure to shape the CS as trade-off to naïve revenue optimisation. An optimal pricing mechanism for quality-differentiated services may on the one hand concentrate on optimal CS—quality is delivered to those requiring it most—or financial performance, ie, revenue and/or profit. As naïve revenue optimisation² neither profits from effective (third-degree) price discrimination nor meaningfully considers the danger of high customer churn rates, CS will play a critical role even for the optimisation of the financial performance. This viewpoint is complemented by the argumentation of [Jai06; LC12] highlighting that QoS provisioning shall be jointly addressed by technological and economical considerations. Laghari *et al.* [LKC12] in the same line stress the importance of targeting all dimensions of communication ecosystems, which according to their definition includes the human, contextual, business, and technological domains.

By comparing the consumer welfare with the Price Elasticity of Demand (PED) of priority and non-priority traffic, Neeklanta & Sardenberg [NS11] demonstrate that especially dynamic rather than static flat rate pricing is capable of minimising the consumer welfare, thus maximising the revenue, for priority traffic (low PED). In contrast, the incentive for NSPs to deploy dynamic pricing schemes for non-priority traffic seems to be limited. Thus, quality-differentiated network services may profit from redesigned pricing schemes that explicitly consider “premium” service offers.

Besides that, works like Anderson *et al.* [AFR97] have suggested a positive relationship between CS and financial performance. Applying a broad literature study, Homburg *et al.* [HKH05] have systematically linked customers’ individual WTP to CS, eg, [Fin93], [RS90]. Seeking for empirical validation, Homburg *et al.* [HKH05] have set up two laboratory-based experiments on deepening the systematic relationship between CS and WTP. Both a questionnaire and an incentive compatible experiment based on the Becker–DeGroot–Marschak Method (BDM)³ [BDM64] with users’ own money have confirmed a direct relationship between CS and WTP, where a non-linear cubic model seems to outperform linear relationships. While this relationship is characterised by a

² When applying a naïve optimisation strategy for the BE Internet where price curves are flat, the optimum converges to the cost minimum at the lowest possible QoS levels, ie, the Capacity Overprovisioning (CO) is limited. CS figures are, however, essential for moderating the customer churn rate.

³ The results from the BDM experiment may have to be treated with care due to the specifically constructed purchasing decision that is rendered subject to subtle conditions, ie, random numbers.

concave relationship for low satisfaction, the function is convex in the other cases. Thus, an S-curved-shaped relationship results (as confirmed by [Ngo99]), where the customers' WTP is insensitive to small CS changes around an inflection point, ie, a reference point due to expectations or past experience. In reference to multiple sources, Homburg *et al.* [HKH05] further conclude that "the relationship between CS and WTP" is substantially strengthened by cumulative experiences (rather than isolated "transactional" judgments on the CS). Ngobo concludes in [Ngo99] from a different angle that a linear relationship between loyalty and CS exists, which, however, does not linearly translate to the customer base. In turn, Ngobo recommends a differentiation among customers rather than a systematic "delighting" of customers through a systematic increase of CS for the entire customer base. Combining those viewpoints with evidence found in behavioural economics [ALP06] and prospect theory [Kah12], it is likely that WTP non-linearly relates to objective inputs (risks and chances, or in our case QoS). For any empirical experiments, it is further considered that WTP might be affected by relative reference points distorting the assessment, which potentially creates a bias.

Transferring these findings to a network service charging context, we claim that CS may be best represented by Quality of Experience (QoE) ratings. QoE ratings such as MOS values on ACR-5 scale (as recommended by [P.800]) reflect a subjective rating of the perceived network or service quality. Despite the existence of such empirical CS-literature in the telecommunications context, the QoE understanding has been notoriously subject to the chosen application, equipment, network conditions, and many more factors. Thus, CS in the network context has to be treated to be highly context-specific, which is implicitly passed on to WTP and corresponding revenue figures.

Section 3.2 investigates the question whether QoE should represent a charging-relevant dimension. It compares the associated financial performance to QoS alternatives.

3.2 Does QoE Pricing Pay?

♣⁴ While already in 2006 Reichl & Hammer [RH06] have motivated a paradigm shift towards pricing the Internet on the basis of human perceptions, ie, charging for QoE, rather than technical parameters or resources, ie, charging for QoS [Tuf04], QoE-based price differentiation has by far not receive enough reception in academia. For this reason, the present thesis works out quantitative measures for the economic and societal potential that is lying idle at present. It illustrates those potentials by briefly summarising the findings in [WZR12] where the prospects of simplistic QoE-based price differentiation are compared with QoS-based alternatives.

Paris Metro Pricing (PMP) [Odl99] is a well-studied and very rudimentary pricing mechanism, which is built in analogy with Paris's metro system of the past. PMP subdivides the resources into two or more classes, eg, wagons of a train. At this point, no

⁴ This presentation briefly reiterates the results given in [WZR12].

quality differentiation whatsoever exists as all classes make use of the same initial resources. The differentiation is then created by varying the price as a tool to moderate the congestion within each class. As a result, more premium and more economically-oriented resource classes will emerge, which will differ in their price and congestion rate, hence quality. The PMP mechanism has been applied to the pricing of communications services in several works [GMS00; Mar06], where this section builds upon the works of Shetty *et al.* [SSW08b; SSW10]. In 2012, this initial work has been extended by [WZR12] (also see [Wah12]). The resulting model is named Pricing and Regulating Quality of Experience [WZR12] (PARQUE) and is subsequently compared to Shetty *et al.*'s initial QoS-based PMP model.

3.2.1 Model

PARQUE follows a 2-class design with access prices $p_1 < p_2$ for the economy Q_1 and the premium traffic class Q_2 respectively. The capacity ratio of premium Q_2 traffic $\mathcal{X}$ relative to the total traffic is exogenously given by the regulator⁵, from which the premium traffic capacity $C(Q_2)$ is derived,

$$C(Q_2) := C \cdot \mathcal{X} \quad (3.1)$$

where C is the capacity of a physical or virtual network resource, $\mathcal{X}$ is static, and Q_1 correspondingly receives $C \cdot (1 - \mathcal{X})$ of the capacity. The QoS q is measured as linear congestion ratio K for the quality class Q_i with $i \in \{1, 2\}$ and the corresponding overall resource demand Z for class Q_i (as a result of all service requests d of users and the required resources to satisfy the requested QoS),

$$q := 1 - K = \frac{Z(Q_i)}{C(Q_i)} \quad , \quad (3.2)$$

where Z is the total resource demand. Each user u is assigned to traffic classes based on their utility $\mathcal{U}_{\theta_u}$ for their personal WTP parameter $\theta_u \in [0, 1]$,

$$\mathcal{U}_{\theta_u} = (\theta_u - p_i) \underbrace{I(q_i - \theta_u)}_{=\beta} \quad (3.3)$$

$$\text{where } I(\beta) = \begin{cases} 1 & \text{if } \beta \geq 0 \\ 0 & \text{if } \beta < 0 \end{cases} \quad .$$

Their WTP θ_u at the same time represents quality and price preferences of customers, ie, an individual equilibrium point where user u can only be served in the interval with $\theta_u \in [\theta_u^*, \theta_u^\dagger]$,

⁵ The role of the regulator characterises the transition from one to two class network services by substantially influencing the user utilities as shown in [SSW08a]. The corresponding ratio for Q_1 is $1 - \mathcal{X}$.

$$\theta_u = p < \theta_u \leq QoS = \theta_u^\lambda \quad . \quad (3.4)$$

In other words, the utility $\mathcal{U}$ represents the difference between the user's WTP θ_u and the charged price for this quality class p_i . The utility $\mathcal{U}$ drops to 0 whenever the minimum quality requirements are not satisfied ($\theta_u^\lambda < \theta_u$ where θ_u represents the actual WTP being matched against the offered quality), ie, users refrain from using this quality class (see step function around β). W.l.o.g. both prices p_i and quality levels q_i are defined to range in $[0, 1]$. Thus, whenever $\theta_u \geq p_2$, the user j is assigned to class Q_2 . They are assigned to class Q_1 if $p_1 \leq \theta_u < p_2$. Whenever $\theta_u^i > \theta_u^\lambda$ or $\theta_u < \theta_u^i$ for service class i , users refrain from using the service class i , ie, the price or quality decisions of the NSP exclude the user from using the service class i . Whenever not matching alternative offer is provided, the user u is excluded entirely from the system.

For this model, the optimal prices p_i^* can be found based on a chosen objective function, eg, resulting in profit/revenue, user utility or social welfare optimisers. Any optimisers will take the set of users with empirically given or randomly generated parameterisation of QoS-sensitivity θ_u for each user u as input and will output the price thresholds for both classes.

At this point, the model is still based on QoS and the work in [SSW08b]. The model can, thus, be used for a direct comparison with the extension presented hereinafter. Contrary to classical PMP approaches, PARQUE is QoE- and hence application-aware. For this reason, [WZR12] has defined two different service types λ , ie, web traffic and video traffic. The video traffic ratio $\alpha \in [0, 1]$ moderates the traffic mix between the services λ : whenever α is 0 only web traffic is considered, while 1 refers to video traffic only. Each user u is, thus, represented by a single service request (web or video) and individual QoS-sensitivity parameter θ_u . For the required QoS to QoE mapping, PARQUE relies on works in the literature, primarily around the IQX hypothesis.

Figure 3.1 provides a schematic comparison of classical PMP-based QoS-differentiation and PARQUE-based QoE differentiation. Based on externally provided capacity ratio X between standard Q_1 and premium Q_2 quality class and overall capacity constraints C of the NSP's network link, a PMP optimiser (left side) may set the access prices p_i , $i = \{1, 2\}$, in order to maximise user utilities, revenues, social welfare or other combinations of user utilities and revenues.

On the right side, the extension of PARQUE introduces a factor κ , which represents the WTP for a certain QoE rather than QoS level. The idea is that users are assigned to one of the two charging classes according to their κ -parameter, while their traffic is still assigned to traffic classes based on their application-specific QoS requirements $[\theta_u^\lambda, \theta_u^\lambda]$ derived from initial κ values and the traffic mix. Thus, in some cases web traffic with low QoS demands is able to be assigned to the higher congested traffic class Q_1 , although a user has signed a premium class contract. For the required relations between QoS and QoE, Wahlmueller *et al.* [WZR12] follow QoE results from the IQX hypothesis for a packet loss metric, ie, referring to QoE results for web traffic in [FHT10] and for video traffic in

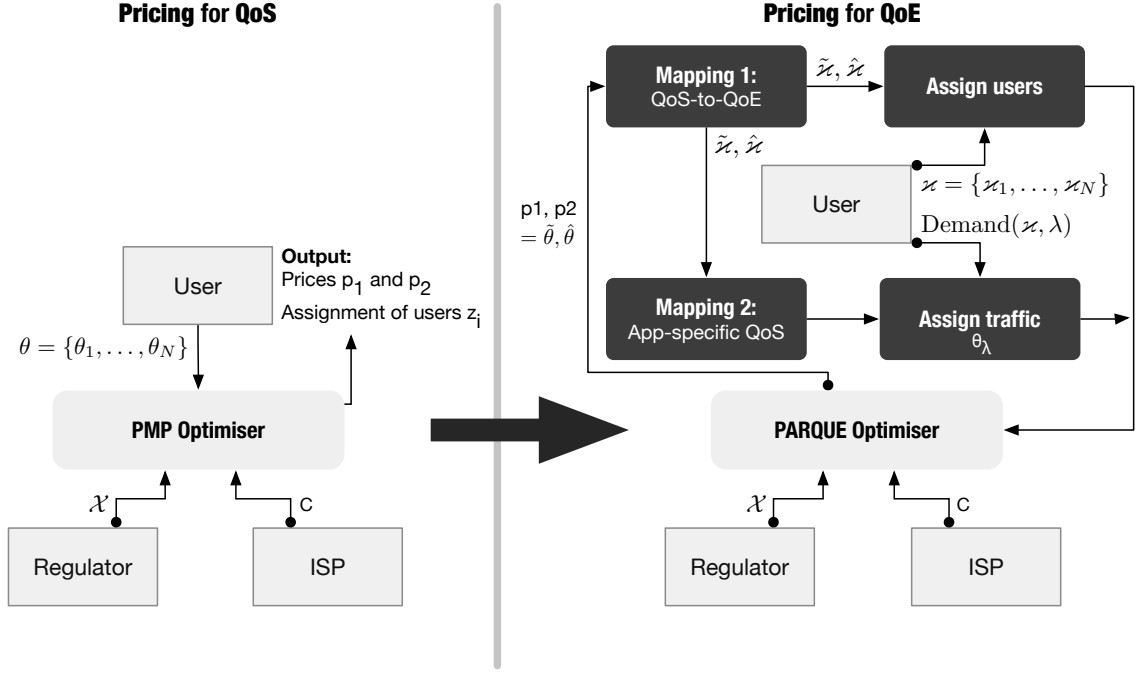


Figure 3.1: PMP versus PARQUE Optimisation

[Zin+10]. As expected, QoE ratings for video traffic are much more sensitive to increasing packet losses.

This optimisation process requires two additional mapping processes: *Mapping 1* maps an aggregate QoS value to a QoE representation, which is then translated back to app-specific QoS requirements in *Mapping 2*. These mappings allow the PARQUE optimiser module to be identical to its PMP counterpart where dependent on the used objective function and the available user population (and their preferences) p_1 and p_2 (equivalent to θ and $\hat{\theta}$ resp.) can be generated. Wahlmueller *et al.* have tested *Mapping 1* in two flavours, the α -sensitive $QoE(\alpha, \theta)$,

$$QoE(\alpha, \theta) := \alpha \cdot QoE_{video}(\theta) + (1 - \alpha) \cdot QoE_{web}(\theta) \quad , \quad (3.5)$$

and a naïve linear mapping $QoE(\theta)$,

$$QoE(\theta) := (QoE_{max}(p_{loss,max}) - QoE_{min}) \cdot \theta + QoE_{min} \quad . \quad (3.6)$$

Due to the surprisingly good performance of the naïve *Mapping 1*, as detailed later on, the analysis will maintain both flavours hereinafter.

Table 3.1: PMP versus PARQUE optimisation ($p_1 = 0.4, p_2 = 0.8$) [WZR12].

	Mapping 1	$\mathcal{R}$	$\mathbf{U}$	Users
PMP	none	~ 32	< 10	~ 60
PARQUE	α -sensitive	> 60	< 13	> 90
	linear	~ 40	~ 20	~ 65

3.2.2 Numeric Simulation

As previously mentioned, PARQUE can optimise the user utility, the NSP’s revenue, social welfare (as the sum of stakeholder utilities) or combinations thereof. For this purpose, 200 independent numeric simulation runs have been conducted with 100 generated users each, ie, $N = 100$ (constant), with random QoE-sensitivities. Each simulation run has been executed with both the PMP and PARQUE algorithm for which some interesting aggregate results are revisited below.

The PMP case has been able to demonstrate the self-stabilisation capability of a quality-differentiated market mechanisms which only manipulate prices of traffic classes.

Applying a revenue maximisation objective function results in an exclusion of up to 40% of the users by setting p_1 too high. Whenever p_1 is high, the number of users that are excluded from our service due to their unwillingness to pay the minimum price p_1 increases, which contributes to an increasing digital divide. The important role of the regulator has further been illustrated by varying the factor $\mathcal{X}$, which substantially influences the overall pricing, ie, p_1 increases with $\mathcal{X}$ while p_2 drops correspondingly. Another interesting effect concerns the social welfare optimisation where user utilities seem to dominate NSP revenues $\mathcal{R}$.

In both flavours, the PARQUE approach can satisfy the needs of more customers—in other words, fewer customers are excluded from Internet services. Apart from that, both user utilities and revenue levels can be increased when replacing PMP with PARQUE. In particular, the α -sensitive *Mapping 1* case results in substantially higher revenues for NSPs, while the linear mapping mainly substantially grows the user utilities but also slightly increases revenue levels. Exemplary results for $p_1 = 0.4$ and $p_2 = 0.8$ are depicted in Table 3.1.

Whenever the α video traffic ratio increases, the revenue grows but the user utilities drop. For more details on the comparison of PMP and PARQUE please refer to the original work presented in [WZR12].

Conclusion #3.1. *The inclusion of QoE in the charging process for Internet qualities can yield revenue and social welfare gains, and can be beneficial for customers when the scheme is properly configured.*

Although Conclusion #3.1 encourages future QoE-differentiated network services, ongoing NN discussions have to be considered on a strategic level (see Section 4.5). Nonethe-

less, the demand and spending behaviour for network qualities needs to be parametrised, which is targeted hereinafter.

3.3 WTP & Demand

Following the motivation and definition of WTP in Section 2.1.1, and its adequate alignment to CS in [HKH05], the present section studies WTP for the case of network video qualities (relating to QoS factors). In the context of network qualities, CS equals the commonly applied notion of QoE in the context of network qualities. Through the conduction of two empirical studies, we are not only able to illustrate a general disposition by users to purchase quality upgrades (ie, Readiness-to-Pay (RTP)), but also absolute and relative WTP values, and their interaction behaviour with our marketplace. Both trials make use of an ample group of subjects in accordance with ITU-T P9.10[P.910] (4 to 40 subjects) and Video Quality Experts Group (VQEG) methodologies (24 subjects), and make use of a test laboratory compliant to [BT.500-7] (eg, wrt distance between subjects and the screen). Particular attention is paid to extrapolating typically hidden (black box) processes of users when converging towards their inner equilibrium. An equilibrium exists at points where the quality and price beliefs of customers are in balance, ie, their WTP and overall demand figures can be retrieved. Such real data are rare, due to the difficult testability, but may provide substantial insight into the reasoning process of users when purchasing experience goods (contrary to search goods where an educated purchasing decision by customers can be done before actually purchasing a good) [Nel70].

While the following discussion mainly follows the publications in [Sac+12a; Sac+12b] (*2011 trial*) and [ZSR13] (*2012 trial*)—plus the partial wrap up in Deliverable D5.100 [CEL14] of the QuEEN project—, more specific details have also found their way into the subsequently categorised publications:

Table 3.2: Additional publications on WTP & demand for network quality.

Topic	Trial	References
RTP	<i>2011 trial</i>	see [SZR12a; SZR12b]
QoE ecosystem	both trials	see [Rei+12; Rei+13a]
Decision motives	<i>2012 trial</i>	see [Rei+13a] and Section 3.5
Market entrance pricing	<i>2012 trial</i>	see [ZRS14] and Section 3.6
QoE & price	<i>2012 trial</i>	see [SZR13b; SZR13a]

3.3.1 Related Work

Despite the existence of numerous empirical works on QoE, such as [LKC12; Sta+10], and its general importance, only limited attention has been paid to an (empirical) understanding of WTP or utilities for the context of telecommunications services so far. With today's

methodologies, many facets of the purchasing process still remain a mystery due to the notorious difficulty of approximating WTP from QoE or other CS-related figures. Such an approach discounts important decision aspects like the psychological phenomenon of cognitive dissonance [Fes57; Sac+12b]⁶, the domination of price or quality aspects within a customer or pseudo-irrational behaviours, eg, in analogy with risk-related estimation problems that have been observed in [FS48]. As argued in [Sac+12b], before research has targeted the economic utilisations, eg, in [Hir65], cognitive dissonance has been interpreted as “random deviation explaining irrational user behaviour” in early works [Bec62]. Some years later, early economic models, such as [AD82], have integrated user preferences and the control of user beliefs, and thus cognitive dissonance. Today, cognitive dissonance is commonly accepted to be a relevant factor for purchasing situations.

A formal description of the conceptual relationship between QoE and charging has recently appeared in [Rei+12], but empirical evidence for a sufficient parameterisation of the given model is still lacking. One of the available empirical works in the literature [RNR08] uses a survey-based technique to question the users’ theoretic WTP. It is, however, unclear whether those values could be materialised in practice, eg, referring to cognitive difficulties of estimating own behaviours [HD95], or biases due to intentional influencing of the public- or self-image.

A more practically oriented methodology has been applied in the M3I project’s experiment 3 in 2001 in the UK [FP502], where during video quality tests every participant received £10 to be freely used for quality enhancements. In contrast to [RNR08], this established a Standard Definition (SD) quality market positioning the user in a more realistic purchasing situation. The users interacted with the market by setting price or bitrate preferences that remained constant over time. Every 30 seconds the system selected one of the 6 available cost profiles (ranging from £0 at 64 kBit/s with the cheapest tariff to £4 at 2048 kBit/s with the most expensive tariff), which lead to fluctuations of the current quality (without preference) or price (without preference) levels. Despite the requirement for a thorough testing of tariff variations, the high complexity did not meet the demands of the end users according after test questionnaires. Users also expressed their preference for *constant quality levels* rather than pricing levels. Moreover, the setup has not been tested in close to reality situations in terms of “living room” atmosphere’s and test durations, ie, computer monitor in lieu of TVs, too many fluctuations and too short viewing periods. Apart from that, the interaction behaviour seems not to have been tracked by the system, which leaves some observations unexplained.

Nevertheless, the discussed experiment provided very valuable first indications on WTP in the context of network video quality. The average user spent £2.80 on quality upgrades in 2001. While users with a bitrate preference have ignored lower quality levels in general, no user has selected the two lowest levels (64 and 128 kBit/s). Information on observed user types, customer segments, interaction behaviour or motives have not been studied in

⁶ In our case, the discomfort from deciding upon positive quality and negative price cognitions, ie, forming an inner decision equilibrium balancing contradictory information.

this experiment. This limitation leaves substantial room for interpretation on what has caused these surprisingly high expenditures.

The subsequent parts discuss empirical laboratory-based studies that have been specifically designed for the VoD use case and have been tested in a controlled but realistic living room-like environment.

3.3.2 2011 Trial

♣⁷ In the light of the results of the M3I experiment, we had designed and conducted an own trial on streamed video quality in Vienna in 2011. This trial has aimed at creating an utmost realistic network video quality market with simple user interaction and home-like environment. In particular, the WTP for a VoD scenario had been studied when offering lowered network impairments for purchase—in our case packet loss for a User Datagram Protocol (UDP)-based video stream.

Subsequently, the details on our test group, our novel test design, as well as the technical and procedural details are given. This section concludes with the main results and some observed phenomena.

3.3.2.1 Demography

A sample size of 43 test subjects, who have completed all stages of the experiment, has been used. This can be regarded to be a large sample size, exceeding ITU-T's recommendation P9.10 [P.910] with 4 to 40 subjects and recommendation BT.500 [BT.500-13] with 15 non-expert subjects for comparable QoE trials, due to the enormous time⁸ and monetary efforts for conducting this kind of laboratory trial. Due to the usage of three identical measurement rounds, data on 129 purchases was obtained (including QoE ratings and several interim measurements during a single measurement). In 2011, the subject group was balanced among genders (22 male and 21 female) and mainly consisted of working age subjects (mean: 36.8 yrs) with completed higher education (eg, university degree). The subjects had limited experiences with VoD services.

3.3.2.2 Test Design

Building upon the initial results of the M3I project, an advanced design for “own money”⁹ WTP experiments for network video qualities is presented in this section. A trial following this design was conducted in FTW's *i:lab* laboratory in Vienna in 2011. But contrary to the M3I experiment, the revised design focused on simplicity for the subjects (when interacting

⁷ *The present section is primarily based on materials that have been presented in our prior publications [Sac+12a; Sac+12b] and also used in [CEL14]. The study was conducted in the context of the FP7 project Economics and Technologies for Inter-Carrier Services (ETICS).*

⁸ Only the conduction of the trial takes about two hours per test subject, while substantially more effort is required for in preparation and post-processing of each individual subject.

⁹ Subjects spend their money to purchase quality upgrades.

with our system), test modularity (variations of network impairments are testable), and a realistic living room-like atmosphere to obtain results with a clearer practical heading.

The design consisted of two dedicated but aligned experiments, not only measuring the WTP for VoD services but also assessing QoE ratings.

Experiment 1 consisted of three measurement rounds for measuring both QoE and WTP under packet loss impairments. Each measurement round, where users were watching a video clip of their choice, was scheduled for 20 minutes. Shorter measurement rounds may increase the test subject’s difficulty to evaluate the value of network video quality appropriately, and may thus create an additional bias. In all phases, users actively purchased quality upgrades with their own money; in the *2011 trial* they received €10 for this purpose.

In this experiment, users chose from the 4 quality levels (levels of packet loss impairments) for a UDP-based video stream, as depicted in Table 3.3. Packet loss rates by design logarithmically descend from Q_0 to Q_3 , while prices are only linearly reduced.

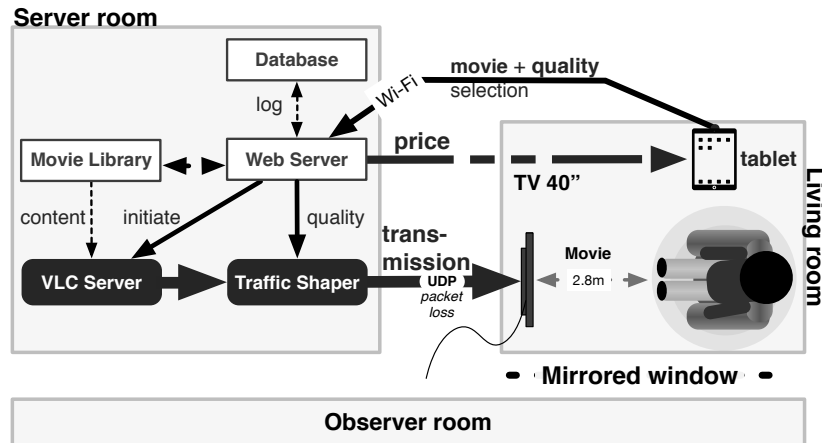
Table 3.3: Quality levels Q_0 to Q_3 .

	Q_0	Q_1	Q_2	Q_3
Packet loss	1%	0.25%	0.085%	0%
Price (20 minutes)	€0	€0.5	€1.0	€1.5

Experiment 2—conducted by the authors of [Sac+12b]—on the other hand, only assessed QoE ratings without any active user decision. Together with experiment 1, this design allows to extrapolate the impact of active user decisions on QoE ratings.

3.3.2.3 Test Routine & Technical Setup

The test started by adding €10 to the deposit of each test subject, which could be used for purchasing network quality upgrades or could be taken home in cash after the trial. After selecting a video of the user’s choice from our extensive library, the video stream was automatically started in the poorest quality Q_0 (highest packet loss). Immediately thereafter a Quality Selection Phase (QSP) was initiated, during which subjects could experiment with our quality levels for a few minutes free of charge. At the end of the QSP, the selected quality level was frozen for the subsequent No Interaction Phase (NIP). The price for the selected quality level was then subtracted from the user’s deposit. This process was repeated for three more QSP and NIP phases where the pricing was linearly adjusted to the remaining duration of the video. After an upgrade was applied (ie, a user reaches a NIP), no downgrade was possible in subsequent QSPs. This process was repeated for each of the three measurement rounds with a duration of 20 minutes each. After each measurement round, QoE ratings were collected.



Source(s): Ipad designed by Kelig Le Luron from the Noun Project

Figure 3.2: Technical setup (2011 Trial)

This design was realised as UDP-based video stream that was transmitted via Ethernet cables between a streaming server and a receiving device, ie, in our case a Mac mini that was attached to a 40-inch television set—see Fig. 3.2. The test subject was sitting on a couch in a distance of 2.8m to the television set as recommended by ITU-R BT.500-7 [BT.500-7]. A tablet, ie, an iPad, was used as input device for selecting a Standard Definition (SD) video of interest and for purchasing quality upgrades. For the purpose of increasing the realism of our experiment, a movie library was carefully designed with preview pictures, descriptions and informative details as a replication of known commercial offers. The video stream was impaired by packet loss using *netem tc* on a Linux server, which was influenced by the user with a click of a button, ie, quality upgrades.

3.3.2.4 Results

The results of the *2011 trial* have revealed a general WTP or RTP for purchasing network video quality enhancements in a VoD scenario, cf. the spending figures in Table 3.4. The subjects selected the second best quality level Q_2 on average (mean=2.04; median=2.25), which has resulted in a surprisingly high expenditure. While one may speculate about explanatory reasons, this may also support the view that some subjects would have been spending (substantially) higher amounts of money. In other words, the maximum WTP figures for the premium customer segment may still be unknown after the *2011 trial*.

Conclusion #3.2. *The 2011 trial demonstrates the existence of a substantial WTP or RTP for purchasing network video quality upgrades.*

In accordance with our a priori expectation, the QoE continuously and non-linearly grows in rising QoS—cf. Fig. 3.3. This is reflected in corresponding MOS figures (Fig. 3.3a)

Table 3.4: Average spending in each measurement round.

Mean	Median	Variance
€1.01	€1.03	$\sigma = 0.49$

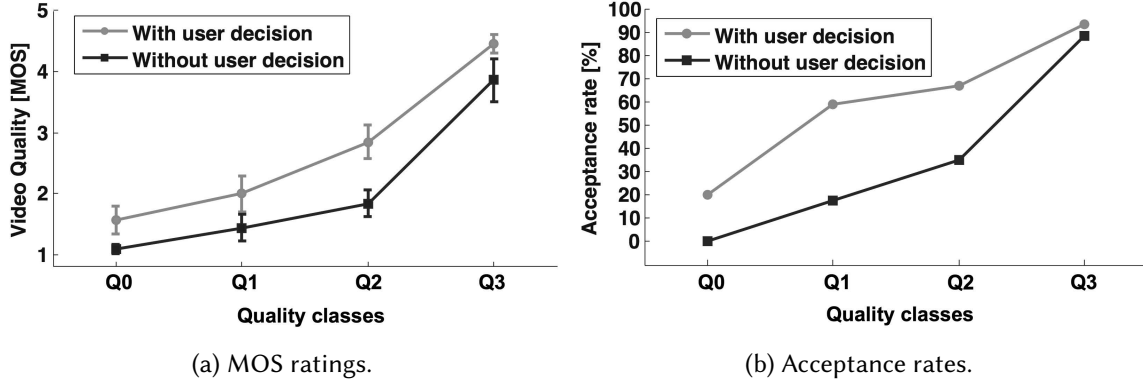


Figure 3.3: 2011 QoE results [Sac+12b].

and acceptance rates (Fig. 3.3b). Further, the results with and without active user decisions are disparate. The active decision to purchase a particular quality level positively affects the QoE ratings, when comparing those ratings to results from classical QoE trials (without active decisions)—see cognitive dissonance effects below. The influence of the price is more pronounced for lower qualities.

Some additional observations from the *2011 trial* are summarised hereinafter (following our contribution in [CEL14] and the original data in [Sac+12a; Sac+12b]):

- **Customer groups:** When inspecting the inhomogeneous spending behaviour of individual subjects, four high-level customer groups are revealed (ie, very high-level customer segments). This finding supports the view that a generic WTP approximation from QoE is notoriously difficult. A more elaborate customer classification is applied for a subsequent trial, which has been conducted in 2012 (*2012 trial*).
 - *Strategic players* (ie, free-riders) consume the best quality whenever it is free (or potentially whenever it is heavily discounted), but ultimately do not purchase equivalent quality upgrades.
 - *Quality players* (ie, maximal spending) are looking for the best available quality. Their purchasing decision is, therefore, dominated by quality considerations and only marginally influenced by price cognitions.
 - *Budget-minded players* (ie, minimal spending), in contrast, aim at minimising their spending, ie, quality demands are dominated by price considerations.

- *Average players* (ie, mass market) are price and quality sensitive. They considerably evaluate their options (and the associated tradeoffs) and eventually purchase a quality level in the middle of the provided range.
- **Cognitive dissonance** (see details in [Sac+12b]): It is complex for test subjects to assess their satisfaction with a price and quality tuple¹⁰. Contrary to classical QoE trials, our unique setup placed test subjects exactly in such a position where they had to actively assess and decide on purchasing network quality upgrades. This setup modification has triggered surprisingly high QoE ratings, which are well above our results from a separate passive QoE testing in a second experiment even for remarkably poor QoS settings—cf. MOS and acceptance rating (“Which share of users would accept this video quality at home?”) comparison in Fig. 3.3. For example, more than 60% of Q_1 buyers would accept the quality level at home (cf. Fig. 3.3b), despite the poor QoS. Such results could confirm the existence of cognitive dissonance effects sugarcoating QoE assessments in order to justify the own purchasing decision. In [Sac+12b] a separation of the passive from the active decision case is recommended and the exploration of niche business models such as low-cost and low-quality cases is advised.
- **General Trends:** As a noticeable group of subjects has purchased the top quality level, their maximum WTP may potentially exceed the bounds of the *2011 trial* setup. Users were very interactive throughout the entire QSP.

Conclusion #3.3. *Due to the existence of several customer groups in the 2011 trial, it can be concluded that a homogenous population of customers seems unrealistic in the market for network video quality.*

Conclusion #3.4. *QoE ratings, both acceptance and MOS ratings (on ACR-5 scale), are non-linearly affected by pricing. A WTP approximation from QoE without thorough characterisation of the relationship between QoE and WTP appears to be notoriously difficult.*

Conclusion #3.5. *Active user decisions may trigger cognitive dissonance effects that may be exploited by tailor-made business models for niches, eg, low-quality and low-price niches.*

Conclusion #3.6. *Classical QoE data cannot be assessed within a WTP trial, as the price cognitions will affect QoE figures.*

Conclusion #3.7. *The unclear absolute WTP bounds from the 2011 trial anticipate the testing of **higher priced tariffs**. This step should be paired with a test of **price variations** in order get clarity over purchasing motives (ie, satisfaction with quality level vs. monetary considerations).*

¹⁰ Product valuation is difficult and subject to a relative evaluation scope [ALP06]. Experience goods are even harder to assess without sufficient historical evidence.

3.3.3 2012 Trial

♣¹¹ Both the *2011 trial* (see Section 3.3.2) and the experiments in [FP502] have revealed a substantial disposition to pay for network video quality enhancements. However, due to some shortfalls in [FP502] (mainly the test complexity and the absence of HD content) and the absence of pricing variations in the *2011 trial* (as well as the apparently too low maximum prices), the absolute WTP and underlying purchasing motivations are still unknown. Especially variations of market entrance prices have remained completely untouched so far, which may hamper the successful bootstrapping of new products.

Due to the extensive data from the *2012 trial*, a more detailed analysis of market entrance pricing effects and purchase motives is postponed to a detailed elaboration in Section 3.5 and Section 3.6 respectively.

3.3.3.1 Demography

A sample size complying to and optimally exceeding the 4 to 40 test subjects recommended in ITU-T's P.910 [P.910] was targeted in 2012. For this purpose, a study with 44 test users ($> \frac{2}{3}$ females) was conducted, from which the results of one test subject had to be removed due to an erroneous recording of the data. 26 subjects were between 18 and 30 years old while the categories of between 31 and 45 years and older encountered for 10 and 7 subjects respectively. The majority of the test subjects were employed (16) or students (16) and had at least a Higher School Certificate. The subject group was unbiased, due to the limited experiences (only 3 subjects with unsystematic spending) with VoD services.

3.3.3.2 Test Design

The *2012 trial* was conducted at FTW's *i:lab* in October 2012 as an “own money experiment” with significant advancements over prior studies, including the *2011 trial*. On the one hand, the experiment targeted an improved extrapolation of factors motivating a purchase (ie, substantially increased number of quality levels), as well as the identification of the absolute maximum WTP of customers (ie, through a deliberate variation of tariffing). On the other hand, the new study intended to retain the usage simplicity for test subjects from the *2011 trial*. Apart from that, the *2012 trial* further has created a basis for a subsequent analysis of market entrance pricing, ie, the sequentiality of pricing network video qualities, which is presented in Section 3.6.

Contrary to prior studies, the revised experimental design deliberately integrated HD resolution video qualities being encoded in 17 *h.264*¹² [ITU-T H.264; ISO/IEC 14496-10] bitrates—see Table 3.5. Such an increased set of video qualities allows for tracking more

¹¹ The present section is primarily based on materials that have been presented in our prior publications [ZSR13; ZRS14; SZR13b] and also used in [CEL14].

¹² h.264 is also referred to as MPEG-4 Part 10, Advanced Video Coding (MPEG-4 AVC) is a video format that compresses the raw video. *h.264* has many application areas on the Internet and is the standard video format for Blu-ray discs.

fine-grained interactions of subjects with our system in order to better understand their motivation for their actions. Correspondingly, a very dynamically interacting system was required to apply price and quality selections quickly. A detailed investigation of user convergence and interaction behaviours is given in Section 3.5.

Table 3.5: Quality levels Q_0 to Q_{19} in kBit/s.

Q_0 128	Q_1 181	Q_2 256	Q_3 362	Q_4 512	Q_5 724	Q_6 1024	Q_7 1448
Q_8 2048	Q_9 2896	Q_{10} 4096	Q_{11} 5793	Q_{12} 8192	Q_{13} 11585	Q_{14} 16384	Q_{15} 23170
Q_{16} 32768	Q_{17} 32768	Q_{18} 32768	Q_{19} 32768	Quality class in kBit/s			

In addition to the 17 bitrates, ie, quality levels Q_0 (poorest) to Q_{16} (best), 3 additional *virtual quality classes* have been added that used the bitrate of Q_{16} with an additional price increase (Q_{17} to Q_{19}). The virtual quality classes serve cross-validation purposes and the extrapolation of price discrimination effects. In total 20 quality classes were available for purchase, which was not made explicit to test subjects.

Drawing from shortfalls in prior trials, the testing of tariff variations as well as the limitation of the complexity for test subjects has been rendered integral for the revised setup. For this purpose, three tariffs were designed that were only varied between measurement rounds (static within each round) and were not disclosed to test subjects (apart from the fee for the currently selected quality level). The three different tariffs (A , B , and C) used the following uniform price distributions:

$$\begin{aligned}
A &= [0, 2] \quad \dots \quad \{p_0 = 0, p_1 = 0.1053, \dots, p_{18} = 1.8947, p_{19} = 2\} \quad , \\
B &= [0, 3] \quad \dots \quad \{p_0 = 0, p_1 = 0.1579, \dots, p_{18} = 2.8421, p_{19} = 3\} \quad , \\
C &= [0, 4] \quad \dots \quad \{p_0 = 0, p_1 = 0.2105, \dots, p_{18} = 3.7895, p_{19} = 4\} \quad ,
\end{aligned}$$

where for any $Q_i \in Q_0, Q_1, \dots, Q_{18}, Q_{19}$ (ordered list) for $i = 0 \dots 19$ the price p_i was $\frac{i}{19} \cdot p_{max}$. The maximum prices p_{max} were €2, €3 and €4 for the tariffs A , B and C respectively.

3.3.3.3 Test Routine & Technical Setup

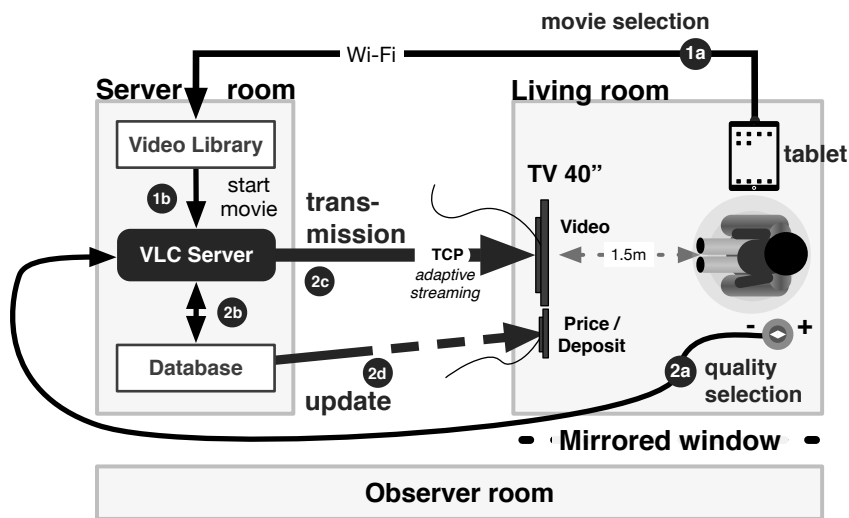
For the whole test routine, each test subject received €10 for their digital deposit. At the end of the trial, the deposit was automatically withdrawn and paid out in cash. During the test subjects could use their deposit freely to purchase quality upgrades. Thus, test subjects have always using their own money throughout the experiment, which is integral for the interpretation of the obtained results.

Each test subject was watching three 20 minutes videos of their choice (from an extensive HD video library) in a row. At each time the video was followed by a short break, ie, measurement rounds t_1 , t_2 and t_3 . In the background and unknown to the subjects, they were assigned to a price plan, ie, a sequence of tariffs. For this purpose, the following three user groups have been formed, which capture all tariff permutations in t_1 , use tariff B for comparison in t_2 and collect further data in t_3 ¹³:

Table 3.6: User Groups.

	Description	Tariff			Sample
		t_1	t_2	t_3	
Group I	<i>Increasing prices</i>	A	B	C	15 Users
Group II	<i>Decreasing prices</i>	C	B	A	15 Users
Control	<i>Constant prices</i>	B	B	C/A	13 Users

In the first two rounds, *group I* tested increasing prices over time, *group II* tested the opposite, and a *control group* was created for cross-validating results with static pricing. This test design has created the hook for analysing market entrance pricing effects in Section 3.6.



Source(s): Ipad designed by Kelig Le Luron from the Noun Project

Figure 3.4: Technical setup (2012 Trial)

¹³ This design step counteracts the limited sample size, which would not allow for a testing of all tariff permutations across measurement rounds. The test groups were further designed to compensate the unequal distribution of tariffs in t_1 and t_2 in the subsequent measurement t_3 .

For each measurement round, the user selected a video on the iPad from our HD movie library (1a in Fig. 3.4), which started the video stream in Q_1 and, thus, the experiment. The experimental setup used a customary 40" HD Liquid Crystal Display (LCD) television at a distance of 1.5m (as normed by [BT.710-4]¹⁴) as the main output device. Immediately thereafter, a QSP was triggered, which allowed the user to test any quality levels free of charge¹⁵ (2a, triggering 2b, 2c, and 2d in Fig. 3.4). During the QSP each selection by the user affected the currently streamed video quality and the price information shown on a second screen. Neither the current quality level nor any absolute bounds were indicated to the user. Paired with the numerous quality classes, an illusion of almost 'endless' quality levels was drawn which aimed at drawing the subject's attention to the current price and quality satisfaction evaluation. Infeasible inputs (deceeding Q_0 or exceeding Q_{19}) were ignored. After ≈ 5 minutes, where a timer always indicated the remaining time, the QSP closed and the last quality selection was taken as purchasing choice. The test has continued with the video in the selected quality for the remaining 15 minutes. This extended phase after the QSP is required to obtain realistic monetary results in the next measurement rounds.

The adaptive video streaming capabilities of *VLC* 2.1¹⁶ were used with Transmission Control Protocol (TCP) video streams. With this upgrade, the video bitrates were automatically adjusted to the current quality selection. Due to an adaptation of *VLC* 2.1, the quality switching time was reduced from ≈ 7 seconds to below 1 second. For the quick interaction with our system a jogwheel, eg, as known from controlling the sound volume, was used, where clockwise rotations referred to quality upgrades and vice versa. Every distinct movement of the jogwheel was logged.

3.3.3.4 Results

♣¹⁷ The WTP is inferred from the demand d for differently priced quality classes. Across all test groups (thus, tariffs) and measurement rounds t , ie, potentially representing two noteworthy biases, the median (mean) spending is €0.95 (€1.193) per 20 minutes movie. That is 31.7% (40.8%) of the respective p_{max} . While even the most expensive quality class, ie, Q_{19} for tariff C, was purchased at a price of €4 and some users opted for the free minimum at Q_0 , the majority of the customers was in the quality- and/or price-sensitive segment in the middle of the range. This view is confirmed when normalising the data in the interval $[0, 1]$ (where 0 represents €0 expenditure at Q_0 and 1 represents the respective p_{max} at Q_{19}): The spending of some subjects is between 0.85 and 1 (85% to 100%) of the p_{max} , more range between 0.40 to 0.60 and even the half of all subjects have values between

¹⁴ The recommended viewing distance is three times the height of the 16:9 TV. In the case of 40" this refers to approximately 1.5 meters.

¹⁵ Of course, all interactions have been recorded, as previously described, and are later on analysed.

¹⁶ Download and information: <http://www.videolan.org/index.html>, last accessed: 20th May 2016

¹⁷ The values presented in this section may partly differ from the results presented in [ZSR13], the original source, due to the correction of a limited number of assessment errors.

0.2 and 0.4. Local peaks occur around 0 and 1. Thus, not only a noticeable premium segment (with some users willing to pay almost any price apparently) but also a price-sensitive and quality-aware customer segment exists. A concentration around extreme values and the average quality level (at 0.50) cannot be observed. This circumstance supports our argumentation for the proper functioning of the used experience-based setup, where by design absolute price and quality bounds are hidden.

Considering variations along the time axis, ie, across measurements $t_{\{1,2,3\}}$, only marginal differences are observed despite the tariff variations (see group design in Table 3.4). The mean spending ranges between €1.15 to €1.24, which represents an insignificant fluctuation around 40% of the respective p_{max} values. The fluctuation of the median is higher, which ranges between 21% and 37%. Nonetheless, the purchasing behaviours are comparable over time, which highlights the consistency of the obtained data.

Slicing and clustering the **data per tariff** A, B and C, the median expenditure continuously grows with the rise of p_{max} (€0.74 at $p_{max} = 2$; €1.26 at $p_{max} = 4$)—see Table 3.7. In other words, the higher the price for a network quality service, the higher the revenue that is obtained.

Table 3.7: Spending per Tariff.

	Tariff		
	A: $p_{max} = 2$	B: $p_{max} = 3$	C: $p_{max} = 4$
Median	€0.74 (37%)	€0.95 (31.7%)	€1.26 (31.5%)
Std. deviation	€0.54 26% ¹⁸	€0.82 28%	€0.99 25%

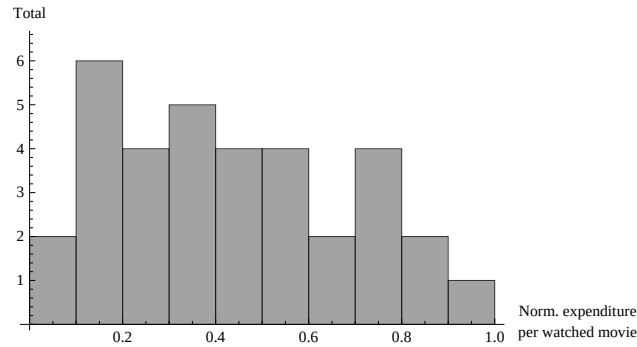


Figure 3.5: Spending histogram with normalised prices (x-axis)

By contrast, when the spending data are normalised in the interval $[0, 1]$ as described above, the opposite is observed, ie, a downwards slope of the normalised expenditure

¹⁸ Percental std. deviation on the basis of normalised price data in $[0, 1]$.

when raising the price—see Fig. 3.5. Explicitly, the normalised spending first decreases slightly before stabilising with a further increase. The first reaction may be caused by the test subjects’ unwillingness to pay a price above their initial equilibrium selection, ie, market entrance effects where prior pricing and purchasing decisions may affect subsequent iterations¹⁹, or because their absolute price beliefs have been broken. The second reaction may in turn relate to the situation where minimum quality expectations are underrun (or the new quality levels significantly differs from prior selection). All these considerations, confirm the anticipated price sensitivity of customers for network services and especially underline the functioning of the setup. Apart from that, these results also highlight the necessity for a specific analysis of tariff sequences (where a decision is conditional to prior tariff experiences), to which the dedicated Section 3.6 is devoted.

Conclusion #3.8. *A substantial maximum WTP exists for network quality enhancements.*

Conclusion #3.9. *The WTP monotonically increases in the charged price.*

While with the increase of p_{max} also the standard deviation grows, it fluctuates around 26% in the case of normalised data. Correspondingly, an Analysis of Variance (ANOVA) test illustrates the significant difference between absolute A and C expenditures at $\alpha := 0.01$, while especially normalised spending results are clearly insignificantly different. These results do not provide a clear view whether price increases may induce a higher diversity of choices within a user group. For this purpose, the Probability Density Function (PDF) per tariff, depicted in Fig. 3.6, is consulted. This view demonstrates a growing concentration between 0.2 and 0.4 when increasing p_{max} , which is paired with a very slight shift towards lower quality levels and an otherwise very similar density pattern.

The distribution of customers to interesting segments is further stressed and utilised in Section 4.

3.4 Quality of Experience (QoE)

♣²⁰ As depicted in Fig. 3.7, the obtained MOS values range from around 1 at Q_0 to above 4 for Q_{15} , for example. Comparably the acceptance rate starts at Q_0 with below 10% and quickly converges to a value close to 100%, which indicates that users at Q_7 or above have been satisfied with their purchase. Moreover, those ratings first drastically increase with the purchased quality level. Around Q_7 , however, a stagnation of the ratings kicks in for both values, although the video quality drastically increases towards Q_{16} . This rating behaviour may appear irrational and, thus, suggests the involvement of cognitive dissonance again.

¹⁹ Please, keep in mind that by aggregating the data per tariff, some users may receive high(er) prices at the beginning or at later measurement.

²⁰ This section briefly summarises the findings published in [SZR13a], which are based on the 2012 trial. The reconditioning of the results has been published in [Zwi+15] before.

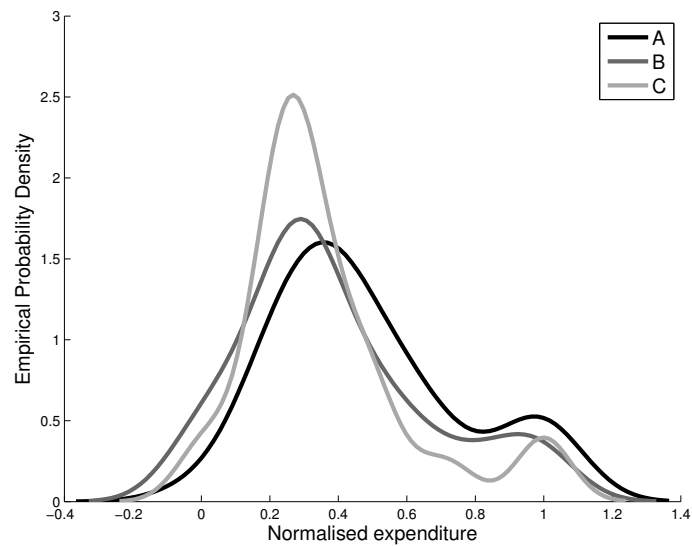


Figure 3.6: PDF for normalised expenditure data

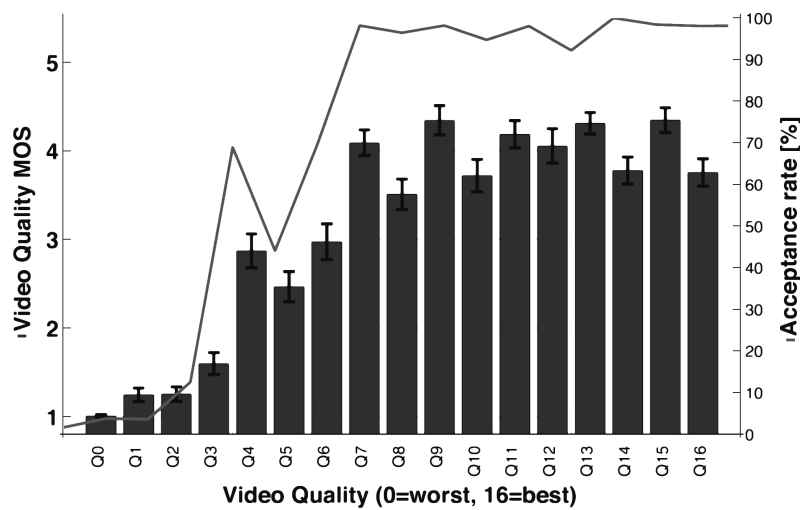


Figure 3.7: QoE ratings for the 2012 trial [SZR13a] (under price cognitions).

However, when comparing the obtained results with an identical testing without active decision (a passive selection), contrary to the *2011 trial* results [Sac+12b], Sackl *et al.* [SZR13a] argue that only inconclusive results are obtained. This may be caused by the test design where due to the high number of tested quality classes the sample sizes are too small and volatile. Thus, the aggregation of several quality classes has been necessary, which, however, introduces some fuzziness. Another explanation is provided by Sackl *et al.* [SZR13a], based on a retesting without content selection using an otherwise identical setup, where they claim that the content selection could be dominating price effects. Due to the statistical shortfalls of the *2012 trial* wrt to this data perspective, we recommend a retesting with fewer groups to verify or disprove price-induced cognitive dissonance.

For a practical usage of the data, eg, in profit optimisers or economic models, the data need to be smoothed and freed of inherent noise. Noise arises from the large number of quality classes and low sample sizes for each class, which may overweight an individual rating and quality class selection. Hence, when using the aggregate QoE results across all tariffs from [ZSR13; SZR13a] (with an estimated p_{max} around 3 in order to limit sample size issues), the following logarithmically shaped fit is obtained (cf. [Zwi+15]; also see [SS14] for other fits):

$$MOS_{p_{max} \approx 3}(bw) := 3.0143 + 0.3163 \cdot \log(bw) \quad (3.7)$$

where bw is the video bitrate in Mbit/s.

3.5 Motives

♣²¹ Unlike prior related work, each interaction of the user with the main input device, a jogwheel, has been logged with a granularity of one second and has later on been parsed. Such kind of data is only required for the purpose of studying the convergence of the test subjects to their inner equilibrium, ie, purchasing decision, and their interest in network quality markets²². With this data, underlying and by now hidden motives of users can be studied when deciding to purchase an experience product.

In the following, the aggregate interaction behaviour from the *2012 trial* is reviewed. Later on, this information serves as the basis for classifying users with the approach given in [Rei+12].

3.5.1 Interaction Behaviour

The subjects intensively interacted with our marketplace during each QSP—cf. Fig. 3.8 with one value per subject per measurement round. While the majority interacted between

²¹ The present section is oriented around our findings in [ZSR13] and specifically processes the interaction behaviour data from the *2012 trial*.

²² Whenever no alteration from Q_0 was recorded, the purchasing decision was made without considering the offered quality levels. The motive for such behaviour is substantially different to cases where subjects finally converge to Q_0 after having tested some quality levels.

10 and 50 times in the QSP of each measurement round, peaking between 10 and 20, some subjects applied even more than 80 inputs. With the help of Fig. 3.9a, it can be observed that the test subjects are most interactive at the beginning of the trial. Despite a continuously decreasing information gain when experimenting with our system, users retain a high interaction level throughout the trial. The final choice on a quality level, ie, purchase, typically takes place at the very end of each QSP. In particular, when inspecting the histogram of the latest inputs within all QSPs (cf. Fig. 3.9a), the peak occurs around the 300s mark of the QSP. Only a few subjects converge to their equilibrium quickly²³. According to our post-test interviews and an inspection of the data, a small percentage of such input patterns refers to *strategic players* that have watched a video in the highest quality during the QSP, but have eventually purchased a lower quality class in the end.

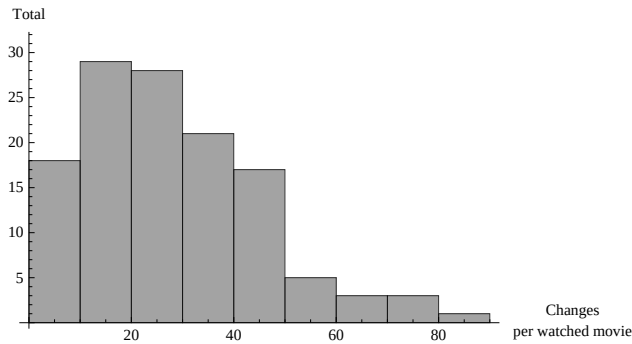
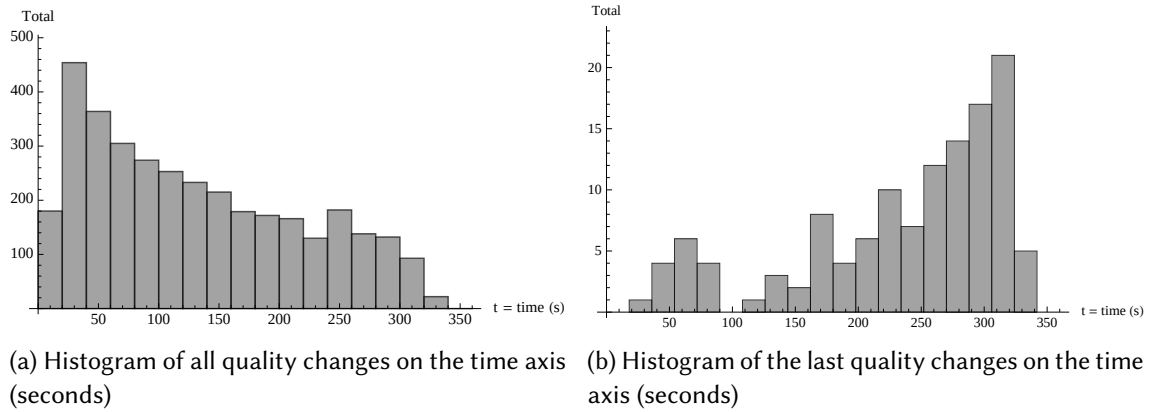


Figure 3.8: Histogram of total quality changes per measurement



(a) Histogram of all quality changes on the time axis (seconds) (b) Histogram of the last quality changes on the time axis (seconds)

Figure 3.9: Histograms of the recorded quality changes (inputs).

For retrieving any turning points²⁴ while users are interacting with our system, we have

²³ Please take note that due to a technical failure a subset of the user group has been able to test quality levels longer than the intended 300s.

²⁴ In analogy with the Critical Incident Technique (CIT), as used for empirical testings [Sal+13], some elements of

defined the *trend changes metric* as a sum of recorded input pattern changes, ie, trends. In other words, whenever a user has recently been increasing the quality class or kept it static, lowering the quality again results in a trend change and vice versa.

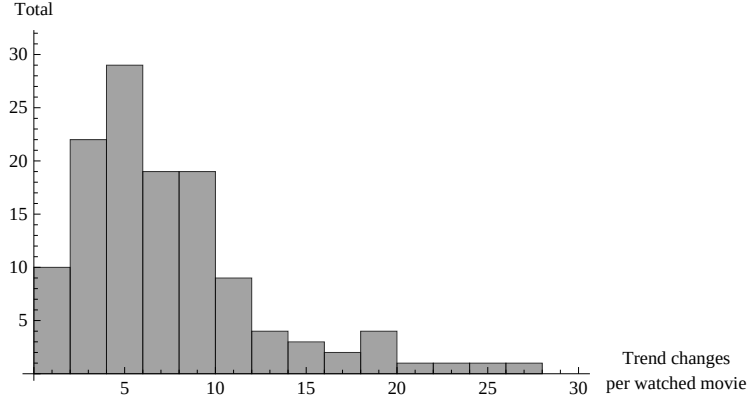


Figure 3.10: Histogram of total trend changes per watched movie

The trend changes metric, thus, yields substantially cleaned up interaction values with high information value. Our data, visualised in Fig. 3.10, illustrate that the majority of test subjects creates up to 10 trend changes with a peak around 5 while outliers even record for 25 of such changes per measurement iteration. The latter behaviour may be subject to a very inconclusive experimentation of the users with the provided quality levels.

When aggregating quality classes selected by all users over time, see Fig. 3.11, interesting alternations of descriptive statistics can be observed: By design originating from Q_0 , a steep increase of the mean quality level towards a flat maximum point takes place in the first few QSP seconds. In this phase, the dissatisfaction with the initial quality class is the driving motive for the triggered inputs. Users try to get rid of the most negative cognition by accepting a negative price cognition. Thereafter, a slight downwards correction of the quality is applied on average, which is dominated by price considerations as driving motive. All additional oscillations (depending on the aggregation of the data, eg, through the classification of similar kind of users) represent convergence adaptations following the same two motives on a smaller scale. Towards the end of the QSP, the development appears more consistent on average, which may hide the fact that test subjects were very interactive also in these phases of the *2012 trial*, cf. Fig. 3.9a.

Characterising the mean overall interaction behaviour of test subjects, it can be concluded that users are seeking for an *acceptance quality* just below their maximum WTP with the biggest distance to Q_0 . This process typically requires several adjustments iterations to find the own inner equilibrium. The corresponding function can be expressed formally by the polynomial approximation function $\mathcal{M}(t)$ (depicted in Fig. 3.11) with a goodness of the fit above 95%,

the recorded data provide a higher information gain for the analysis than others.

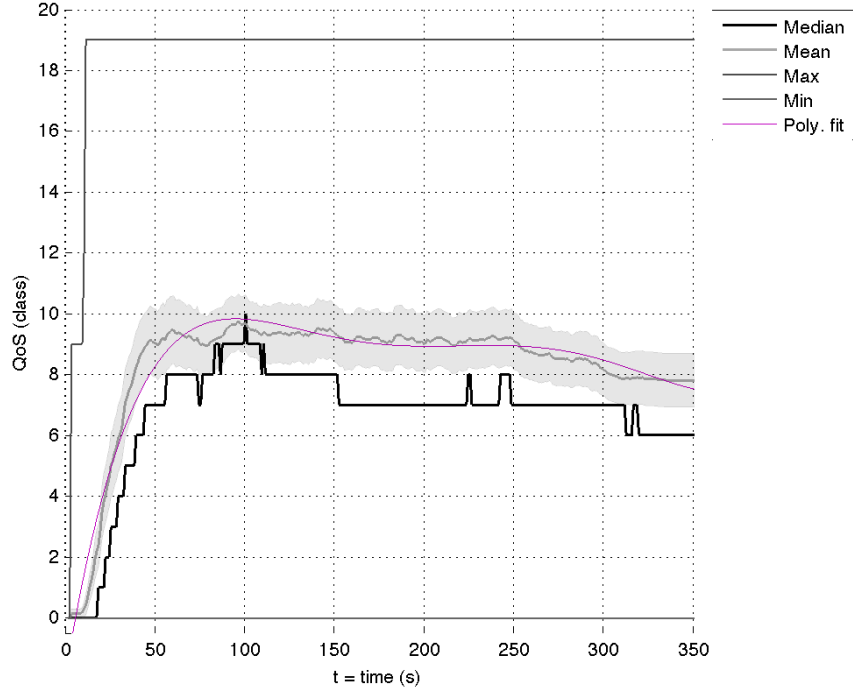


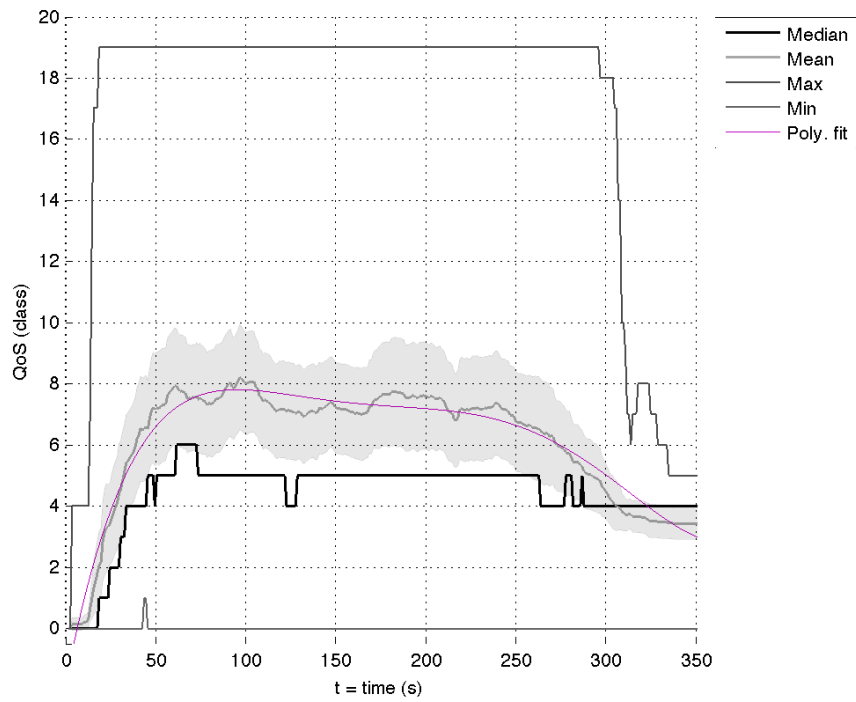
Figure 3.11: Overall results

$$\begin{aligned} \mathcal{M}(t) = & 4.3145 \cdot 10^{-11}t^5 - 4.8441 \cdot 10^{-8}t^4 + 2.0425 \cdot \\ & \cdot 10^{-5}t^3 - 0.004t^2 + 0.3562t - 1.7564 \quad , \end{aligned} \quad (3.8)$$

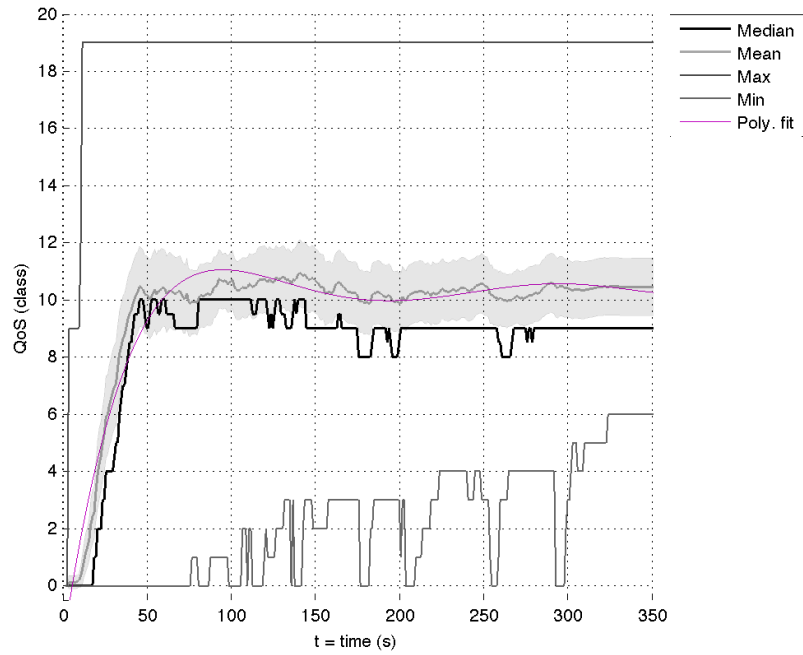
where $\mathcal{M}$ represents the mean quality selection over time t .

In the next step, customers with a spending above the median expenditure are defined to be “generous” (see Fig. 3.12b) and those with a spending below the median are called “parsimonious” (see Fig. 3.12a). Although the curves of both groups share many characteristics with the general case (steep quality increase first; cf. Fig. 3.11), in the second phase noteworthy deviations are noticed: On an aggregate level, the test subjects interacted, volatile, without a significant adjustment of the expenditure in this phase (ie, data “noise” is produced). The other group, however, significantly lowers the expected quality selection, especially in the very final phase. This further confirms the price domination in the second phase but also underlines the quality motive at the beginning.

When organising the data per tariff, it is revealed that the downwards trend in the second phase becomes more pronounced with the increase of p_{max} . While in absolute numbers the global peak for tariff A is reached after a local spike (non-significant difference between peak values; cf. Fig. 3.13a), a slight downwards trend (as in the overall figures) exists after an initial peak for tariff B—cf. Fig. 3.13b. A steep and significant reduction of the selected quality occurs for tariff C in the second phase—cf. Fig. 3.13c. Thus, it can

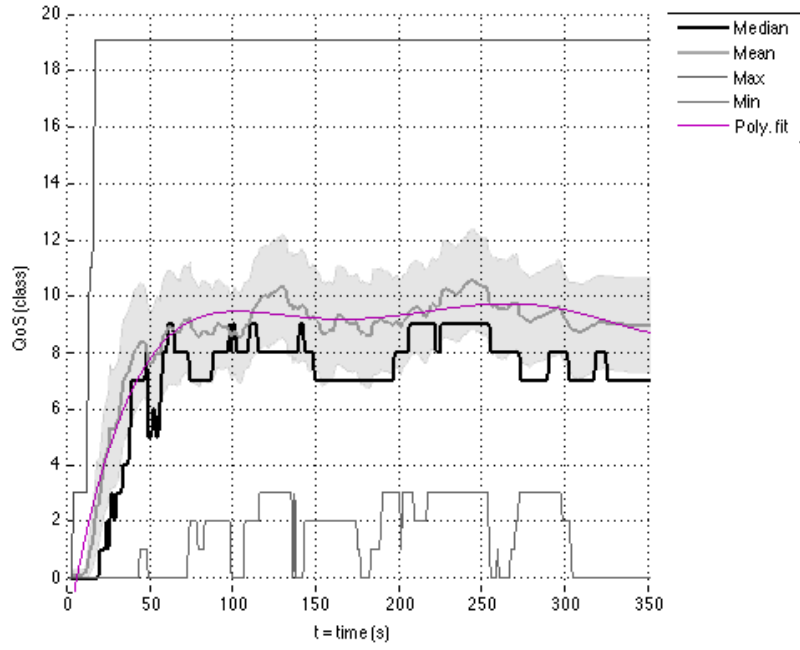


(a) Parsimonious subjects (below median)

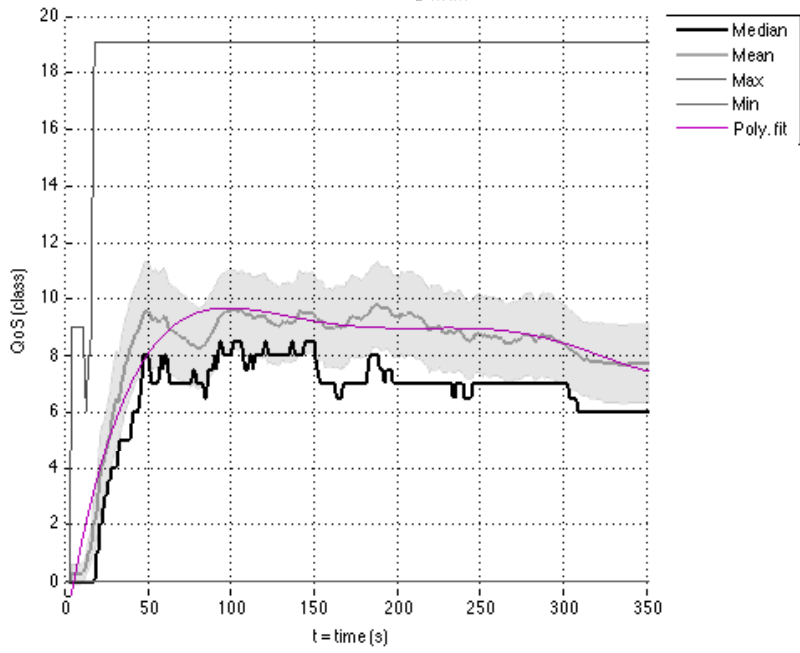


(b) Generous subjects (above median)

Figure 3.12: Uncovering the test subjects' convergence towards a purchase

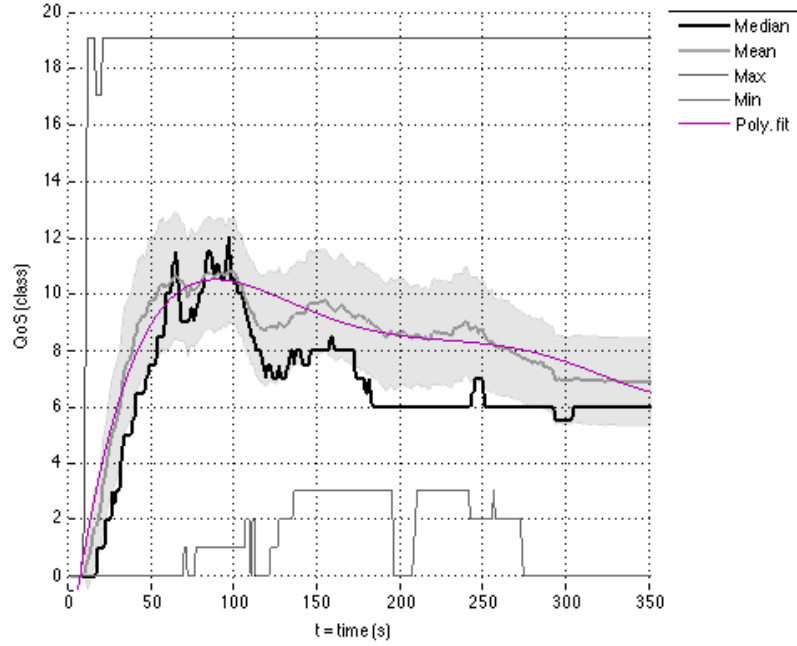


(a) Tariff A ($p_{max} = 2$)



(b) Tariff B ($p_{max} = 3$)

Figure 3.13: Purchasing decision convergence process aggregated per tariff (tariff B and C).



(c) Tariff C ($p_{max} = 4$)

Figure 3.13: Purchasing decision convergence process aggregated per tariff (tariff C).

be concluded that tariffing, as expected, significantly influences both the mean purchasing decision—and, hence, the overall expenditure—and the convergence patterns towards the equilibrium point.

Conclusion #3.10. *While the first interaction phase during the purchase of a network service is dominated by quality cognitions (ie, identification of the desired quality level), the few final seconds are dominated by price decisions (ie, reductions to a point where the price-quality balance is acceptable).*

Conclusion #3.11. *Tariff levels affect the interaction behaviour of customers with a market system.*

3.5.2 User Classification

♣²⁵ Using the data that was gathered in our study, a concept for classifying users according to their interaction behaviour with our marketplace has been proposed in [Rei+12]. With this concept, which is briefly revisited in the following, the users are classified along the 1) *convergence time*, 2) *interaction amplitude* dimensions.

²⁵ This section briefly revisits and revamps the material that has been presented in [Rei+12]. For a more formal representation, this thesis kindly refers to [Rei+13a].

For the convergence time cv , Reichl *et al.* [Rei+12] define the classes “*fast convergence*” (F), “*regular convergence*” (R) and “*slow convergence*” (S), plus an additional “*irregular*” (X) class where no clear assignment to any of the classes F , R or S is possible.

The typical case for a class X kind of user would be the free-riding case where the revelation of the real purchasing decision (and thus, WTP) is postponed intentionally to the very end of the QSP. The amplitude am is defined to drop with an increasing speed of convergence, ie, $am(S) > am(R) > am(F)$. In other words, slowly converging users are thoroughly experimenting with available purchasing options before finally arriving at a decision.

Following this idea, users from the *2012 trial* are assigned to classes on the basis of their interaction behaviour and a Root Square Deviation (RSD) $\tilde{\sigma}(t)$ over time t metric. Reichl *et al.* specify the following reference curves for the defined classes:

$$\tilde{\sigma}^F := 3 \cdot \exp(-0.01t) \quad , \quad (3.9)$$

$$\tilde{\sigma}^R := 6 \cdot \exp(-0.007t) \quad , \quad (3.10)$$

$$\tilde{\sigma}^S := 9 \cdot \left(1 - \frac{t}{300}\right). \quad (3.11)$$

Since the square difference $\mathcal{D}_i^k$ between a user i 's $\tilde{\sigma}_i(t)$ and any reference curve k by design decreases between the F and S classes, each user i can be assigned to the most suitable class as follows:

$$i \in \begin{cases} F, & \text{if } \mathcal{D}_i^F \leq 1. \\ R, & \text{else if } \mathcal{D}_i^R \leq 2. \\ S, & \text{else if } \mathcal{D}_i^S \leq 3. \\ X, & \text{otherwise.} \end{cases} \quad (3.12)$$

The interaction data of 40 out of 44 users (from the *2012 trial*) have been correctly tracked, of which 39 subjects are properly classified according to (3.12) in [Rei+12]. One of the users cannot be classified due to a minimalistic but fast convergence behaviour to the best quality level, ie, exceeding the $\mathcal{D}_i^F$ threshold. The 39 users are distributed to the classes as follows: Eight (8) users in F ($\mathcal{D}_i^F$: mean = 0.69; std. deviation = 0.16), 11 in R (1.58; 0.28), 15 in S (2.58; 0.36) and five (5) in X (8.30; 2.91).

Thus, the data of 34 users (85% of the total) show a convergence behaviour relating to the defined reference curves. Only five subjects have to be attached to class X and a single user cannot be classified (ie, 15% insufficiently classified users). Besides that, no demographic differences are observed whatsoever.

3.6 Market Entrance Pricing

♣²⁶ When facing the challenge of kick-starting a previously undeveloped market, as in the case of quality-differentiated network services, the successful handling of the market entrance phase is crucial. Following the literature review from Section 2.1.2, it appears realistic to assume that the customer assessment is set relative to reference points (reference values). The experiments around [ALP06] further illustrate the customers' difficulty of assessing the worth of services when being confronted with (arbitrary) reference points, which will especially hold in the case of experience goods. Depending on historical experiences of customers, the subsequent quality and price expectations may be influenced, ie, an unbiased view of products is replaced by a more realistic real world perspective. This is supported by Bolton and Lemon's claim [BL99] that customers "seek to maintain payment equity in the service relationship over time, adjusting items under their control (usage levels) in response to changes made by the firm (price changes, changes in service quality).", where they define payment equity as "...customers' perception of the fairness of the exchange of payment for service usage". As a consequence, a tuple $rf := \langle time, price, quality \rangle$ may serve as reference point rf . Thus, the empirical moderation and testing of reference points has to encompass the three dimensions of time (historical experience vs. present), price to be paid, and the delivered and experienced quality. Each dimension needs to be individually tested. Subsequently, the present section concentrates on the assessment of reference point effects, ie, biases, induced by price variations (with multiple network video quality alternatives) over time.

A proper understanding of customer behaviour when being confronted with price changes has not yet been studied adequately according to [HKH05]. This circumstance relates to an insufficient understanding of how price variations (eg, prices increases over several tests iterations) affect the WTP for network video services.

Subsequently, the *2012 trial* data are used in order to shed light on effects arising from the sequentiality of pricing setting for the context of network video qualities, ie, a study on market entrance pricing. For this reason, the present section is oriented mainly along the results given in [ZRS14] building on the trial results presented in [ZSR13].

3.6.1 Monopolistic Market: A User Perspective

Comparing the median spending of the *2012 trial* across all measurements (t_1, t_2, t_3) on absolute and normalised scale—bearing it mind that the values in $[0, 1]$ directly refer to the chosen quality level—the different characteristics for *group I* (increasing prices; $A \rightarrow B \rightarrow C$) and *group II* (decreasing prices; $C \rightarrow B \rightarrow A$) stand out: on normalised scale, where each price $p_i^{\{A,B,C\}}$ and the corresponding user payments by a user is rescaled in $[0, 1]$ relative to the respective $p_{max}^{\{A,B,C\}}$ (for tariff A, B or C) by

²⁶ This section summarises the findings presented in [ZRS14].

$$p_i(0, 1) = \frac{p_i^{\{A,B,C\}}}{p_{max}^{\{A,B,C\}}} \quad , \quad (3.13)$$

the spending in *group I* is stable whereas it drops in *group II* from measurement t_1 to t_2 . Thus, while price stimuli appear to be ineffective, price increases lead to an immediate compensation reaction since they break the chosen equilibrium²⁷. With a large enough price increase, however, test subjects again adjust their purchasing behaviour, ie, the purchasing behaviour is stable between t_2 and t_3 , as their minimum quality beliefs may be broken or endangered.

By contrast, *group II*'s sharply falling revenues when decreasing prices bottom out with the extent of the adjustment on the absolute scale. The revenues of the *group I* also drop in the first phase, but then strongly increase.

The *control group* remains static on both scales, which confirms the functioning of the setup and test group design.

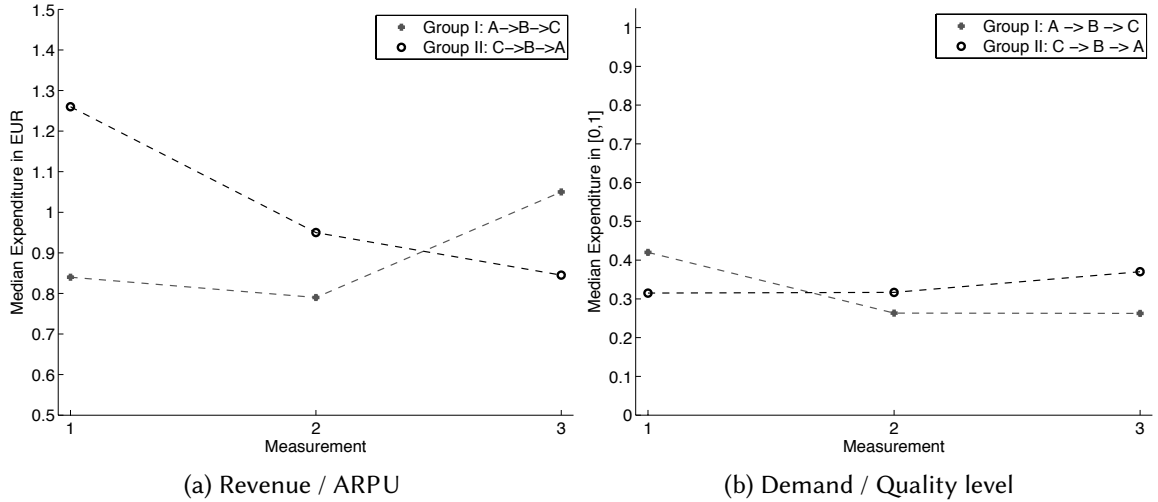


Figure 3.14: Revenue and quality level demand (normalised in $[0, 1]$) for measurements t_1 , t_2 , and t_3 .

Further elaborating the statistical analysis of per measurement and test group effects on absolute scale, the following elaboration concentrates on t_1 and t_2 : the correlations are highest for $C \rightarrow B$ (*group II*) with 0.909, which is greater than for the *control group* $B \rightarrow B$ with 0.844, and lowest for $A \rightarrow B$ (*group I*) is lowest (0.670). The higher the correlation, the closer is the similarity between the data in t_1 and t_2 . Despite the unfeasibility of correlations to provide causality or plausibility argumentation, it supports the general line

²⁷ Contrary to price increases, price reductions provide an additional positive stimulus on top of an acceptable equilibrium point. The current position is, thus, improved over the initially chosen equilibrium point, therefore no fundamental belief, such as an absolute maximal WTP, is broken.

of results. Thus, there are good indications that a price increase enforces an immediate behavioural adaptation in order to stay below the absolute WTP of a test subject (unless more important beliefs are not broken with a corresponding behavioural adaption).

Due to the limited sample sizes, a repeated measures analysis of variance (ANOVA with Repeated Measures (ANOVA RM) [Dav02]) of several cross-cuts of the data has yielded only a limited number of significant outcomes: a significant subject-matching at $\alpha = 0.05$ is achieved²⁸ for all tested cases. The time difference (measurement round) is only significant when comparing t_1 and t_2 for *group II* ($C \rightarrow B$) and the control group ($B \rightarrow B$). Group effects are insignificant in a similar comparison. Time & group interaction is closest to significance²⁹ in the general case, which perfectly fits the intuition that time and group effects need to coexist to yield distinguishable results. By design, the test groups only differ in tariffing at one measurement (t_1) in this comparison.

Specifically looking at the ANOVA of *group I* and *group II* in t_2 we still have to recognise insignificant results, despite a massive absolute difference in variance (0.762 in *group I*; 0.2810 in *group II*) and median (0.79; 0.950) values, ie, high effect sizes. The large difference, nevertheless, indicates the behavioural and perceptual difference between attributing losses (higher price; same quality) and gains (lower price; same quality)—as anticipated from factors of human behaviours known in literature (cf. Section 2.1.2)—, and further mirrors the intentional minimisation of active decisions by users (avoiding cognitive dissonance effects).

Concluding on this representation of our trial data, a reasonable likelihood exists that current prices can also affect future purchases. Price increases seem to be intensively perceived by test subjects while price cuts can substantially lower a provider's revenues. In particular, the conducted experiments indicate that revenue losses are reversed with growing price increases. Contrary, the revenue downturn is bottomed out when increasing from €3 to €4.

Conclusion #3.12. *A reasonable likelihood exists that historic prices for network services may affect future purchases with unrelated price figures.*

Thus, subsequently a series of characteristics and conclusions can be derived for the case of **price increases**:

- Very elastic demand curve, which flattens with increasing price
- Subjects compensate moderate price increases by purchasing lower quality levels, ie, leading to a comparable or lower ARPU
- Bigger price increases are not (fully) compensated by behavioural adaption, ie, revenue growth tendency

Hence, the subsequent conclusions can be derived:

²⁸ A cross-validation of ANOVA RM's functioning

²⁹ $\approx 70\%$ – 80% significance, which is way off the desired α value of 0.05.

Conclusion #3.13. *Monopolistic price increases in large steps are preferable over small modifications for NSPs³⁰.*

Looking at **price reductions**, the following characteristics can be observed:

- Small price reductions do not stimulate demand, ie, no behavioural adaptation / no redecision
- Clearer price, eg, +50%, incentives are recognised
- Steep ARPU drop flattens out with the magnitude of price reductions steeply decreases

Conclusion #3.14. *From a per-user perspective, NSPs are **not** recommended to offer unsustainable low teaser prices at market entrances that cannot be retained over time.*

Conclusion #3.15. *In monopolistic network markets, stable pricing³¹ always provides the best revenue figures. However, if necessary, price reductions from high initial prices are preferred over price increases from low teaser prices.*

Conclusion #3.16. *In a competitive setting, larger but sustainable price cuts may be more useful to attract new customers, which at the same time reduce the marginal cost, ie, positive scaling effects.*

Whenever price reductions are applied to profit from economies of scale and the prospects of higher spending (see #3.16), competitors may be forced to respond accordingly. Such strategies may culminate in a low-revenue equilibrium where demand sensitivity is (nearly) flat.

3.6.2 Market Challenges

Complementary to the per-user view on market entrance pricing issues, ie, a monopolistic view, this section briefly revisits the standard market perspective. This review prepares the discussion for antagonistic tendencies between monopolistic optima and their competitive counterparts.

- *Lack of demand & two-sided market issues [RT06].* New services, especially experience products such as network qualities, may require a stimulation of the consumer demand. In analogy with UMTS, for instance, in the past, stalemates may iteratively occur whenever non-iterative upgrades do not universally provide credible market perspectives. Thus, progressive pricing or intensive marketing may get the market up and running, and may motivate investments in quality-differentiated networks.

³⁰ It may be tested whether small price steps could recreate this effect.

³¹ Prices that do not require modifications after their initial announcement are referred to as “stable” in this context.

- *Market power.* When entering a new market, price dumping³² (eg, through unsustainable teaser prices) may be applied, if not against societal interest, to succeed in already settled markets.
- *Economies of scale.* Telecommunications immanently requires massive investments in infrastructure, eg, macro-cells, in order to satisfy customer demands, ie, network quality, and legal constraints (eg, wrt universal access). Thus, through a high utilisation of the infrastructure unit price gains may boost the market share.

In turn, NSPs should seek for stimulating customer demands through the exploitation of unit price gains, eg, through teaser prices. They should further increase their market share in order to attain a critical mass of customers [EH95]. Hence, this yields the following conclusion on NSPs' pricing strategies:

Conclusion #3.17. *Immature markets, ie, market entrance with or without high competition, may be best served by progressively penetrating the market, eg, through aggressive pricing, in order to decrease unit costs.*

When pairing the non-competitive conclusion #3.14 with the competitive view in #3.17, the first antagonism is yielded that calls for alleviative measures: Introductory prices, which entice customers to try novel services, may be an essential tool to bootstrap a market for experience goods, as in our case. But this strategy simultaneously creates severe WTP limitations that may disrupt the market sustainability in the long run. As argued in [ZRS14], the absence of a proper resolution drives the market towards lower prices and, thus, towards lower price segments, ie, in the ultimate case the market may converge towards *low profitable* and less *sustainable* BE-like market conditions. In other words, the short-term success may be traded in for long term prospects. This antagonism is, thus, placed in direct analogy to debatable “filling the pipe” strategies of the past, which have apparently caused a decline in revenues in recent years³³—see conclusion #3.18.

Conclusion #3.18. *Teaser price strategies for the market entrances phases, eg, “filling the pipe” strategies, may trade short-term interest for the long-term sustainability of the market. Such pricing strategies should, hence, be only used carefully and optimally as measure of last resort, eg, when marketing strategies insufficiently meet the demands of an immature market.*

3.6.3 From Monopoly to Competition

Subsequently, market stimulation methods omitting debatable teaser prices are discussed in order to regain some control and prospects on future revenue: First of all, business partners and customers may be attracted by clear outlines of attractive service or market opportunities.

³² Definition by the European Commission: http://ec.europa.eu/trade/tackling-unfair-trade/trade-defence/anti-dumping/index_en.htm, last accessed: 20th May 2016

³³ Wireless intelligence: <https://wirelessintelligence.com/files/analysis/?file=2011-03-10-european-mobile-arpu-falls-20.pdf>, last accessed: 20th May 2016

Conclusion #3.19. *Increased marketing efforts (demand stimulation due to blandishing pricing effects as suggested by [Sac+12b]) and ante-dated communication of investments (attracting interest of business partners) may be preferred over unsustainable introductory prices for Internet services in the long run.*

The degree of active user decisioning, ie, reconsideration of purchases may essentially influence optimal strategies:

Conclusion #3.20. *Active decisions of users should be avoided when applying moderate price increases. Subtle price increase automatisms may be used in order to circumvent situations requiring a decision by the customer and having the potential to lead to a loss of revenue (see Section 3.3 and 3.6.1).*

The following conclusion applies to active purchasing situations and various extents of competition:

Conclusion #3.21. *For **low competition** in a telecom market, significant price increases may be preferred over small changes to avoid revenue loss or stagnation. Small, but adequate, price reductions are advised (immediately revenue-effective), if necessary. In contrast, under **high competition**, more significant price cuts are advised due to market penetration effects and flattening revenue curves.*

Services may be tailored towards customer groups with varying price sensitivity of demand:

Conclusion #3.22. *First-degree (eg, through the functioning of virtual quality classes, see Section 3.3) and third-degree price discrimination (see customer segments in [Sac+12a]) is recommended for network services. Teaser prices should be applied only to price-sensitive customers.*

3.7 Evaluation & Concluding Remarks

In both our *2011* and our *2012 trial*, the mean spending behaviour (equal to the ARPU) for a 20 minutes video has been lower than in the M3I experiment 3 [FP502]—ie, €1.01 for SD content in the *2011 trial* and €1.19 for HD content of the test subject's own money in the *2012 trial*. The discrepancy may be linked to cultural differences or declining prices in the 10 years between the tests. This may, on the other hand, also result from our more advanced testing methodology with an improved practical orientation and lowered test user complexity. However, the obtained expenditure has exceeded our prior expectations prior, which also deserves further attention: in [ZSR13], we argue that the high expenditure may on the one hand be linked to absence of cheap and high-quality VoD offers at the location of testing. On the other hand, our high maximum prices p_{max} may have raised the spending to unexpected heights. The used prices, as high anchor points, may have revealed the customers' real appreciation for great network video qualities as an aggregate of the appreciation for the video and the network itself. The successful application of

price discrimination, ie, the actual purchasing of *virtual quality classes* (Q_{17} - Q_{19}) and its appreciation by customers, and the existence of an upmarket customer segment, where customers additionally pay for almost undistinguishable quality gains³⁴, further boosts the mean spending levels. Some users seem to favour excellence over price, which represents an interesting customer segment. All three experiments, however, conclude with the existence of a substantial RTP and interesting variations over tariffs. A more direct reconciliation of concrete results is, moreover, impossible due to the absence of detailed figures for the M3I experiment.

In both the 2011 and the 2012 trial, the acceptance rates and MOS values non-linearly increase with the provided QoS, which mirror the non-linear relationship between QoE and QoS. In comparison to classical ACR-5 MOS ratings, the acceptance rates seem to rise with a steeper slope when improving the QoS from the poorest quality levels, eg, see Fig. 3.3.

From their empirical experiments, Homburg *et al.* [HKH05] have expected a growing WTP with the rise of CS, in particular in a symmetric S-curve shape. While the basic growth trend can also be confirmed by our trials, the effect of cognitive dissonance [Sac+12b] (*2011 trial*) seems to blandish CS ratings. Referring to Fig. 3.7 where the quality classes (x-axis) correlate to linearly increasing prices (and thus WTP as only the ratings of buyers have been collected), the QoE ratings (both MOS and acceptance rates) start to saturate very soon without forming an S-curve shape. However, when comparing the QoE data from purchasing situations, see presented in Section 3.4, with classical QoE data (without pricing), eg, see [KRD12] and [SZR13a], the prices induce a flattening of the QoE towards both ends, as also described in [SS14]. This flattening could reflect the S-curve-shaped relationship between CS (here QoE) and WTP as characterised in [HKH05]. Future works may need to clarify the details and detailed shape of this relative relationship between QoE and QoE under price cognitions to verify the S-curve relationship hypothesis. Due to the high number of quality classes and the resulting small sample per quality class, we have neither been able to clearly reject nor verify the cognitive dissonance claims from the *2011 trial* in the *2012 trial*.

Furthermore, different kinds of users exist whether on the axis of overall expenditure level (see *2011 trial*) or their interaction behaviour towards their inner equilibrium (see *2012 trial*). The user interaction data of the trial in 2012 have revealed two phases in which users interact with our marketplace (up to 90 changes and 27 trend changes per measurement round): first, users have increased the quality levels until their price beliefs were broken. Second, they have applied a fine-grained correction to decline below their absolute price maximum at a point where quality and price beliefs are in balance. With our studies on purchasing interactions, the test subjects' handling of contradictory quality and price cognitions has been extrapolated during their convergence towards their inner equilibrium, ie, the active purchasing decision.

Through a dedicated assessment of market entrance pricing (in analogy cumulative

³⁴ The difference was scarcely perceptible even for trained personnel between Q_{14} and Q_{16} .

testings of CS in [HKH05]) effects in the *2012 trial*, this chapter has been able to illustrate that price reductions do not necessarily lead to an active reevaluation, ie, test subjects tendentially stick to their quality decisions from prior test iterations, which represents a notable reference point bias. On the contrary, price increases immediately trigger a reduction quality upgrades, ie, the price is focal in this phase for test subjects. In other words, we have observed a high demand sensitivity to price increases and low sensitivity to price decreases around a reference point. These trends, however, fade out or revert with a higher magnitudes of changes, which requires for strategic handling by supplying NSPs.

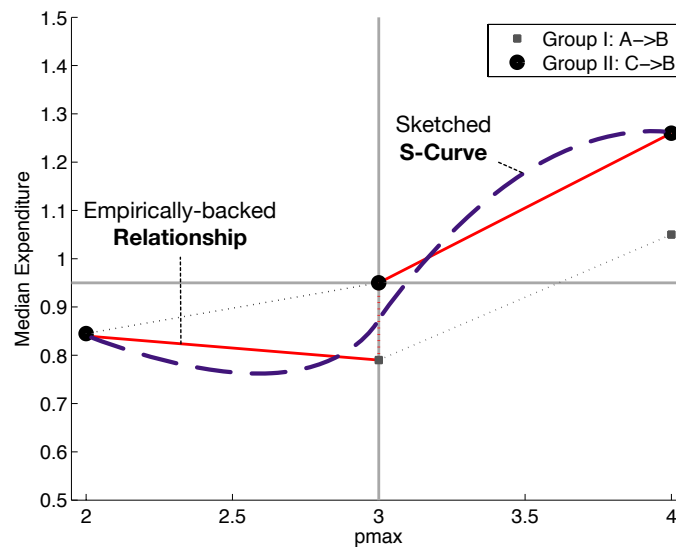


Figure 3.15: S-curve between median expenditure and p_{max} values (*2012 trial*)—red line: empirical data; purple line: smoothed result.

On this basis we construct another kind of non-symmetric S-curve describing a specific perspective on the relationship between WTP and the pricing in analogy (cf. Fig. 3.15). Due to the low number of three aggregate data points, the empirically-backed relationship (red line) is transferred to an additional sketched S-curve (dashed purple line). The resulting S-curve describes whether a reference point, in the current case the *control group*'s expenditure data (shown as grey axis), has been reached due to price increases or decreases. Under decreasing prices this point is exactly met for tariff *B*, while a gap (dotted red line) exists for corresponding value under price increases.

This represents an unexplored revenue potential, which may be lost due to market entrance pricing effects. This figure further illustrates that the discrepancy due to the sequentiality of tariffs increases continuously with p_{max} .

Due to relative small sample sizes, although exceeding the requirements of most im-

portant standards and despite the high difference in absolute terms³⁵, we have to admit the absence of statistical significance conforming our market entrance pricing effects. We, thus, recommend to eliminate one of the three testing groups in future trials, as implemented in the validation study, *2015 trial*, which was run both in Vienna (Austria; different laboratory) and Oulu (Finland)—see Chapter 6. The validation study further serves the purpose of verifying or falsifying the effects and phenomena identified in the *2011* and *2012 trial*.

³⁵ The high absolute differences are a measure of the effect strength, which is, despite the interdependence, non-identical to significance figures.

ECONOMICS OF SUPPLY

This chapter concentrates on the supplier’s problem in direct response to the given empirical demand parameterisation. Before introducing the contributions of this section, the supplier’s problem is briefly motivated: Due to the handling of an experience good [Nel70] in the form of network (video) qualities, the phenomenon of the *lemon market* problem [Ake70] is of particular relevance. Originally “lemons” refer to used cars that are a posteriori regarded to be mispurchases. As it cannot be determined a priori which of the cars could be a lemon, the associated risk needs to be integrated into the valuation of *any* considered car. This, unfortunately, lowers the reselling price for high-quality products below their actual worth. In other words, the a posteriori WTP would be higher than the a priori value, which undermines the market for above-average quality products. In the context of network services, any market aside BE products are potentially rendered unattractive or less valuable by lemon market effects¹. This section will, subsequently, concentrate on service guarantees eliminating the risk of a lemon market problem, while many of our results may also be transferable to other kinds of “premium” Internet service markets. The associated challenge of conveying experience products to customers is discussed in the related work in [Var+15], which deserves a clear formalisation in the future and goes beyond the scope of this chapter.

In this context, this chapter concentrates on the following supply-side optimisations: In direct continuation of Chapter 3, Section 4.1 transfers the empirically collected data to a perspective relevant for suppliers—ie, demand, WTP, market segments and revenues. This section specifically focuses on mitigating sample-size related effects. The outcomes then provide inputs to the core of the supplier’s problem: an optimisation problem to focus on the revenue- or profit-optimal usage of available resources, which is separately targeted by single-class (Section 4.2) and multi-class service provisioning optimisations (Section 4.3). The results illustratively connect the empirical work of this thesis with common business problems of suppliers, such as relating to resource scarcity, or resource

¹ Those effects may be alleviated by marketing instruments such as trial periods offering previews that are offered to customers free of charge for a representative period.

efficiency and effectivity. By this time the investments in network capacity are treated as exogenous input, Section 4.4 thereafter in contrast investigates the potential of the currently trending network offloading strategies costly cellular to free Wi-Fi spectrum is reviewed. Finally, when introducing quality-differentiated Internet services, NN concerns may arise, which are channeled into a discussion of societal constraints along the axis of fairness and regulation (Section 4.5).

Hence, this chapter primarily targets the following objective:

Objective: Economically efficient and effective utilisation of available network capacity when considering realistic customer populations (based on empirical data) and minding social constraints

This objective can be subdivided into the following three contribution blocks:

- Empirically-backed single- and multi-class revenue optimisation
- Analysis of the suitability of cost reductions around the hot topic of network traffic offloading
- Techno-economic strategies around societal constraints such as fairness or NN

4.1 Market Segmentation

♣² In direct response to the demand data collected in Chapter 3, this section characterises network quality markets beyond the scope of a single trial—in our case the *2012 trial*—towards forming the supplier’s perspective. For this purpose, techniques to generalise the collected empirical results will be studied in order to eliminate sample-specific effects. The outcomes will serve as key input for dedicated simulations of network quality markets. This section, therefore, characterises and quantifies WTP for the network video quality case and conceptually distinguishes it from customer segments figures. Finally, indications to generalise our laboratory testing experiences to other prominent test fields such as QoE is given.

4.1.1 Demand & WTP

The demand function d , as empirically assessed in Section 3.3.3 and with price curves translated to WTP, is subject to quality q and price levels. The product of demand and price characterises the revenue for the supplier. Due to rather low sample sizes, this section aims at generalising and classifying the empirical data by consulting literature on related cases: As characteristic for third-degree price discrimination [Cow07], eg, quality

² *An orthogonal and more high-level perspective than presented in this section is provided in Section 3.3.2.4 on the basis of the 2011 trial.*

attributes or heterogeneity in general [IM94], the demand for quality-price combinations is supposed to follow a beta distribution B (cf. Fig. 4.1 and Fig. 4.2). To the best of our knowledge, no other distribution was suggested in literature for third-degree price discrimination or related problems. For this reason, we will subsequently test the B for its suitability to generalise the empirical data of previous sections in order to obtain more representative data source to be used in subsequently designed optimisers³.

Using the Ferrari & Cribari-Neo beta regression [Fer04] (via the implementation in the R Project⁴ [CZ09] building upon a refined beta density,

$$b(y, \mu, \Phi) := \frac{\Gamma(\Phi)}{\Gamma(\mu\Phi)\Gamma((1-\mu)\Phi)} y^{\mu\Phi-1} (1-y)^{(1-\mu)\Phi-1}, \quad 0 < y < 1, \quad (4.1)$$

with the gamma function Γ , $0 < \mu = p/(p+q) < 1$ with search parameters $p, q > 0$ (p, q relate to the initial Γ function parameters), $\Phi = p + q$, and $p, q > 0$, an approximation Φ in $(0, 1)$ ⁵ is retrieved. For the corresponding beta distribution we will use the notion of B . Using this beta regression technique with a logit link function⁶ for the demand data aggregated in 3 quality class bins (in order to reduce the noise around similar choices), the regression functions (where x as indices for the quality class Q_x in the *2012 trial* that ranges in $[0, 16]$ and is translated to the bitrate in Mbit/s using the QC function),

$$B(x)_A := \frac{e^{-1.32512-0.03140 \cdot QC(x+1)}}{1 + e^{-1.32512-0.03140 \cdot QC(x+1)}} \quad , \quad (4.2)$$

$$B(x)_B := \frac{e^{-1.10596-0.07594 \cdot QC(x+1)}}{1 + e^{-1.10596-0.07594 \cdot QC(x+1)}} \quad , \quad (4.3)$$

$$B(x)_C := \frac{e^{-1.00800-0.08073 \cdot QC(x+1)}}{1 + e^{-1.00800-0.08073 \cdot QC(x+1)}} \quad , \quad (4.4)$$

and very low precision parameters Φ of $\Phi_A \simeq 8.550$, $\Phi_B \simeq 16.539$ and $\Phi_C \simeq 13.99$ ⁷ are obtained. The results will be detailed hereinafter.

Cook diagrams for similar beta regression operations on non-aggregated data (cf. Fig. C.1 in the Appendix) have illustrated a limited number of outliers. Those outliers may be explained by the small sample sizes in comparison to the offered quality-price tuples.

³ The suitability of B is regarded to be of speculative nature for the current context and the available data. For this reason, this section applies and explorative assessment.

⁴ The R Project for Statistical Computing: <http://www.r-project.org>, last accessed: 20th May 2016

⁵ The obtained demand d is an additional step rescaled to any population size.

⁶ The beta distribution comes in several shapes, which is addressed in the used beta regression tool by a series of link functions. For the our case, no other link function has constantly outperformed the standard link function, ie, logit link function.

⁷ The precision parameter $\Phi > 0$ gives indications on the present variation of the prediction, eg, $B(x)_A$, from the initial data.

Apart from that, the data have been fitted to a normal distribution, which yielded $N_A(\mu = 6.839; \sigma = 3.234)$, $N_B(5.824, 3.735)$ and $N_C(5.086, 3.524)$ (where μ is the mean and σ the variance).

Evaluation. Despite the reasonable Cook’s distance diagrams, according to the coefficient of determination⁸ R^2 unfortunately only between 10% ($p_{max} = 2$) and 48% ($p_{max} = 4$) of the variation are explained by the beta regression model⁹. Thus, it can be argued that the low coefficient of determination values result from the beta regression’s inability to capture the variability of demand peaks. The promising Cumulative Distribution Function (CDF) match (R^2 of 0.872, 0.915 and 0.897 resp.) highlights the powerful WTP trend prediction of the beta distribution¹⁰. However, we can still conclude that while third-degree price discrimination is often supposed to relate to a beta distribution [Cow07], for our laboratory results this cannot hold on a general basis. The demand for particular combinations of price and quality needs to be separated from the overall purchasing trends.

In particular, this may be explained by the very low number of data points, ie, 17 considered quality classes, the highly limited sample size, and the kind of data. We, however, argue that the mismatch most likely results from the test question that confronted users with a series of quality-price combinations at the same time. Despite exceptional efforts to eliminate known test biases in the *2012 trial*, the remaining test bias induced by price tags and multiple quality-price offers (anchor points for relative assessments) has to be regarded to be of inherent nature. The test question itself induced an inevitable impracticality to eliminate the remaining bias due to the different pricing for different qualities, which seems to have affected the subjects’ decision process to some degree. Known other bias have been removed by hiding absolute quality and price ranges, quality class information and the behaviour of others from the subjects, ie., our design has focused on the “purchasing experience” itself.

However, as anticipated by [Cow07], the beta regression may well characterise demand levels for third-degree price discrimination as collective figure for a binary purchasing question, ie., “Do you want to purchase quality A in price level P” (without any other options or alternatives).

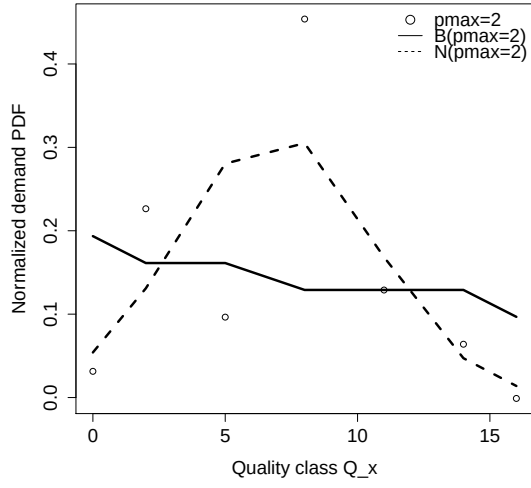
As a result, the CDF approximation (cf. Fig. 4.2) by beta regression appears to be of higher quality than its PDF counterpart. Nonetheless, the fitting quality appears insufficient to meet common scientific standards. In order to exclude the optimiser as potential cause for the insufficient curve fitting, Sage¹¹/Python, Matlab and other R curve fitting techniques were also put to test. None could, however, produce better results with the same data set where Matlab even failed to produce results at all. However, carefully

⁸ R ’s pseudo R^2 for its univariate standard unit beta fitting has been used due to the inapplicability of classical R^2 for such kind of input data.

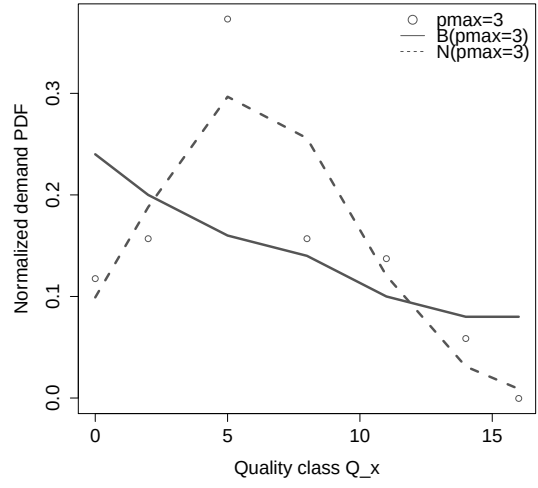
⁹ The coefficient of determination is 1 (=100%) whenever a perfect linear relationship exists, which is not the case for our model. A very strong relationship typically exists above 0.7.

¹⁰ This view very likely corresponds to the essence of WTP predictions in literature.

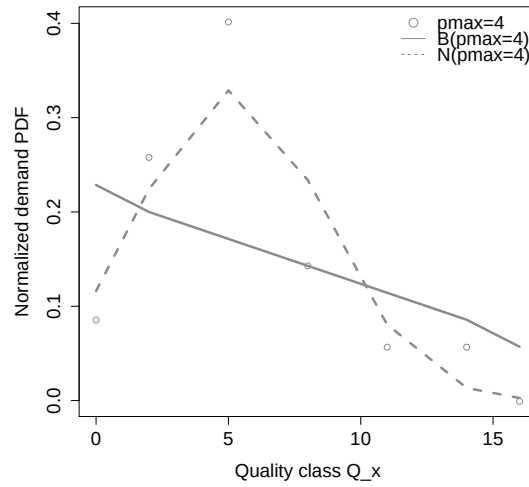
¹¹ <http://www.sagemath.org/>, last accessed: 20th May 2016



(a) Tariff A with $p_{max} = 2$

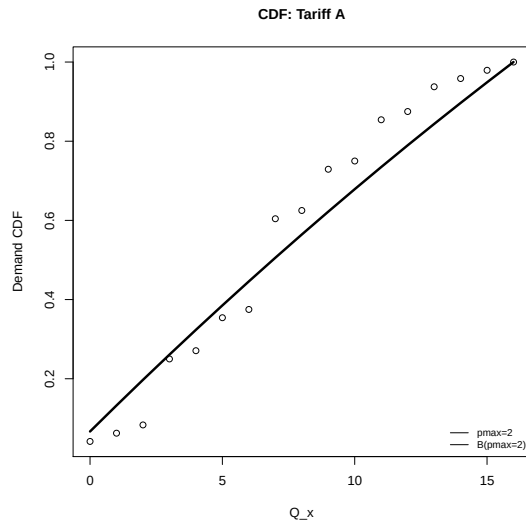


(b) Tariff B with $p_{max} = 3$

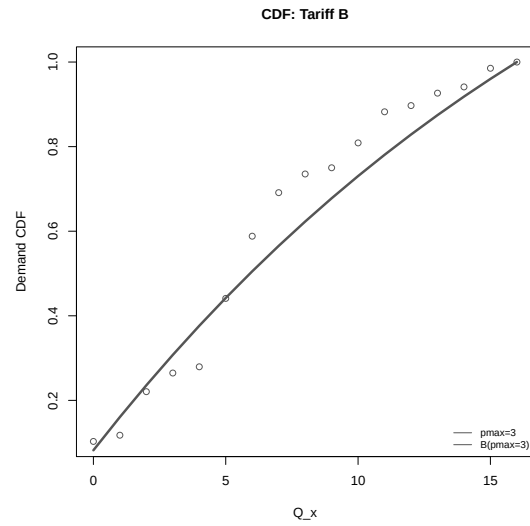


(c) Tariff C with $p_{max} = 4$

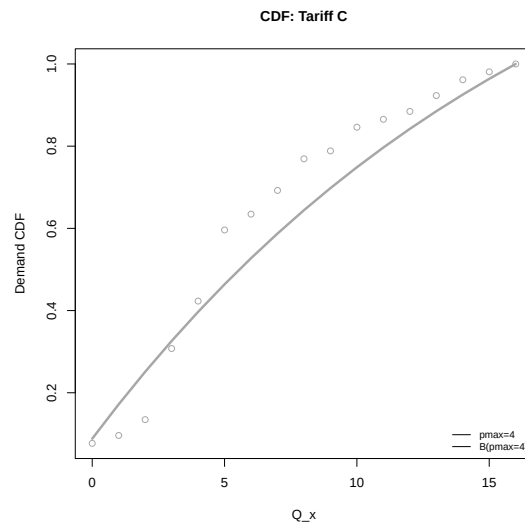
Figure 4.1: PDFs for d and $B(d)$ for varying p_{max} (bins of three classes).



(a) Tariff A with $p_{max} = 2$



(b) Tariff B with $p_{max} = 3$



(c) Tariff C with $p_{max} = 4$

Figure 4.2: CDFs for d and $B(d)$ for varying p_{max} .

inspecting Fig. 4.1, the quality class extremum Q_0 shines out. Telecommunications has a long history in offering so called BE-Internet, a non-usage-based flat tariffing scheme that suggests that low or acceptable quality classes are free or appear to be free in situational usage contexts. While the *2012 trial* setup convinced subjects to test different quality levels and many of them bought intermediary or even high-quality offerings, the stronger commitment to Q_0 seems aggravate the beta curve adaptation, which could represent an industry specific bias.

So, when excluding Q_0 , and for symmetry and adequate quality class bins with constant size of three also Q_{16} (a rarely purchased class due to the also offered purely first-degree price-discriminated quality classes Q_{17} to Q_{19}), Matlab identifies a substantially better fit (see Fig. 4.3. Contrary to the previous technique, Matlab refers to the beta distribution in the following notion:

$$B(a, b, p, q) = \int_a^b (u - a)^{p-1} (b - u)^{q-1} du = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} (b - a)^{p+q-1} \quad , \quad (4.5)$$

where the parameters a and b are optimised based on the provided data and Γ is again the gamma function with parameters p and q . The details of the computation are, however, less explicitly introduced for Matlab than for the case of R statistics. The Matlab optimiser returns the following distributions: $B(a = 1.881, b = 2.377)_A$, $B(a = 1.878, b = 2.636)_B$ and $B(a = 1.692, b = 3.002)_C$. Due to some distortions introduced by the optimisers, the results are not compatible with the empirical input data. For this reason, a rescaling is necessary to assure that the sums of assigned users to quality classes remains identical:

$$B(a, b)_{\{A, B\}}^\sigma := B(a, b)_{\{A, B\}} \cdot \sigma \quad , \quad (4.6)$$

where σ scales the summed output (summed shares of users) such that it becomes 0.968, 0.882, 0.914 for tariff A , B and C respectively. These values are the ratio between the users assigned to the subset of quality classes (Q_1 to Q_{15}) and the overall empirical data (Q_0 to Q_{16}).

The results provide a very high coefficient of determination: The new prediction explains 96%, 98% and 97% of the input data for tariff A , B and C respectively¹². On the other hand, only a few data points were studied, which entails a need to validate appropriateness of beta distributions in future trials.

When upscaling the results to the entire population (100% of the users attached to classes between Q_0 and Q_{16}), unfortunately similar kind of discrepancies arise as in the case of a beta regression solution—see Fig. 4.4.

As a result, we will use a softer interpretation around the initial beta regression solution with respect to the usage of beta distributions in this context: From this point of view, the beta regression values are best understood as aggregate metric for single-choice **global demand levels** (and transitively **WTP**) rather than as metric capturing the distribution

¹² The obtained results are scaled differently due to the needs of the different optimisers. The corresponding empirical data is scaled to an identical scale for a direct comparison.

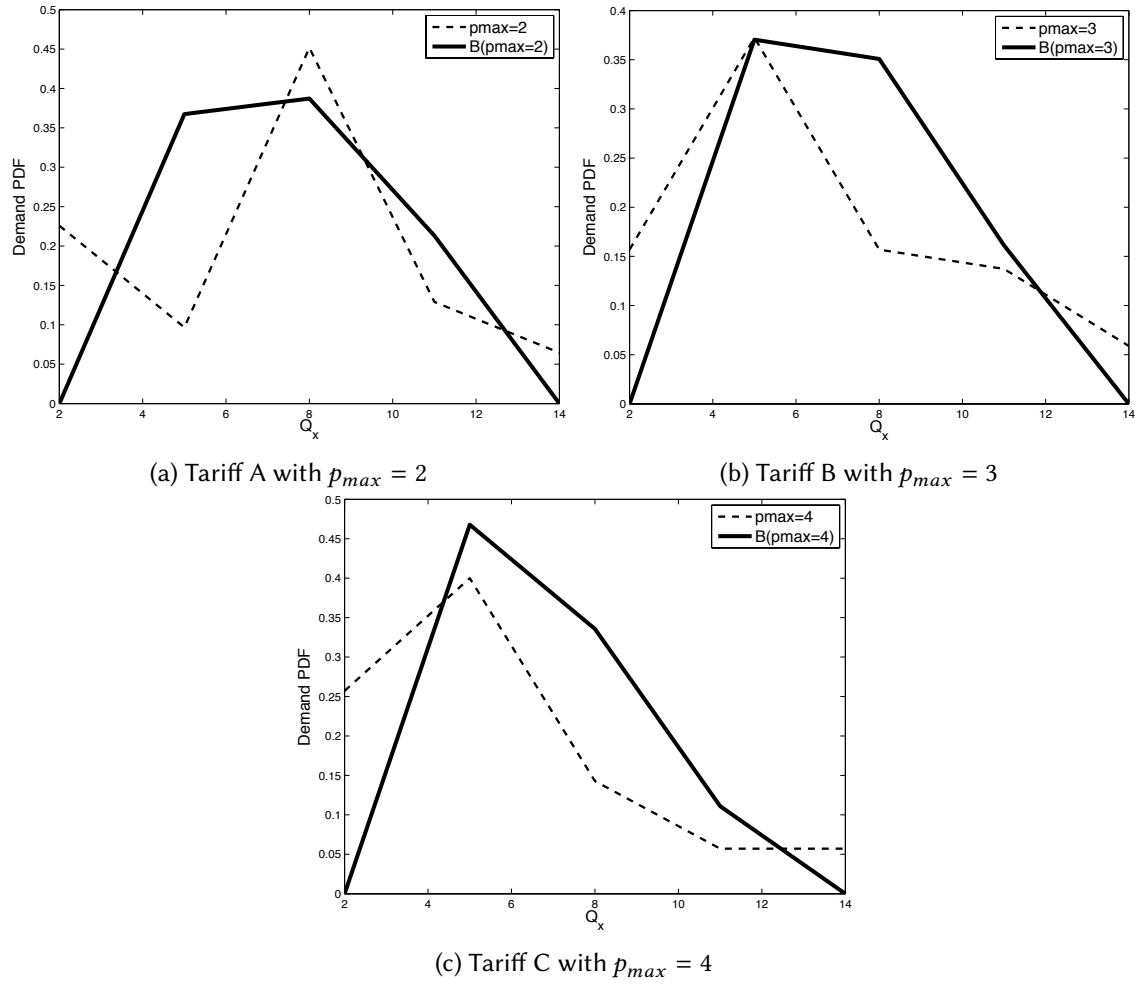


Figure 4.3: PDFs for d and $B(d)$ for varying p_{max} (excluding end points; bins of three classes).

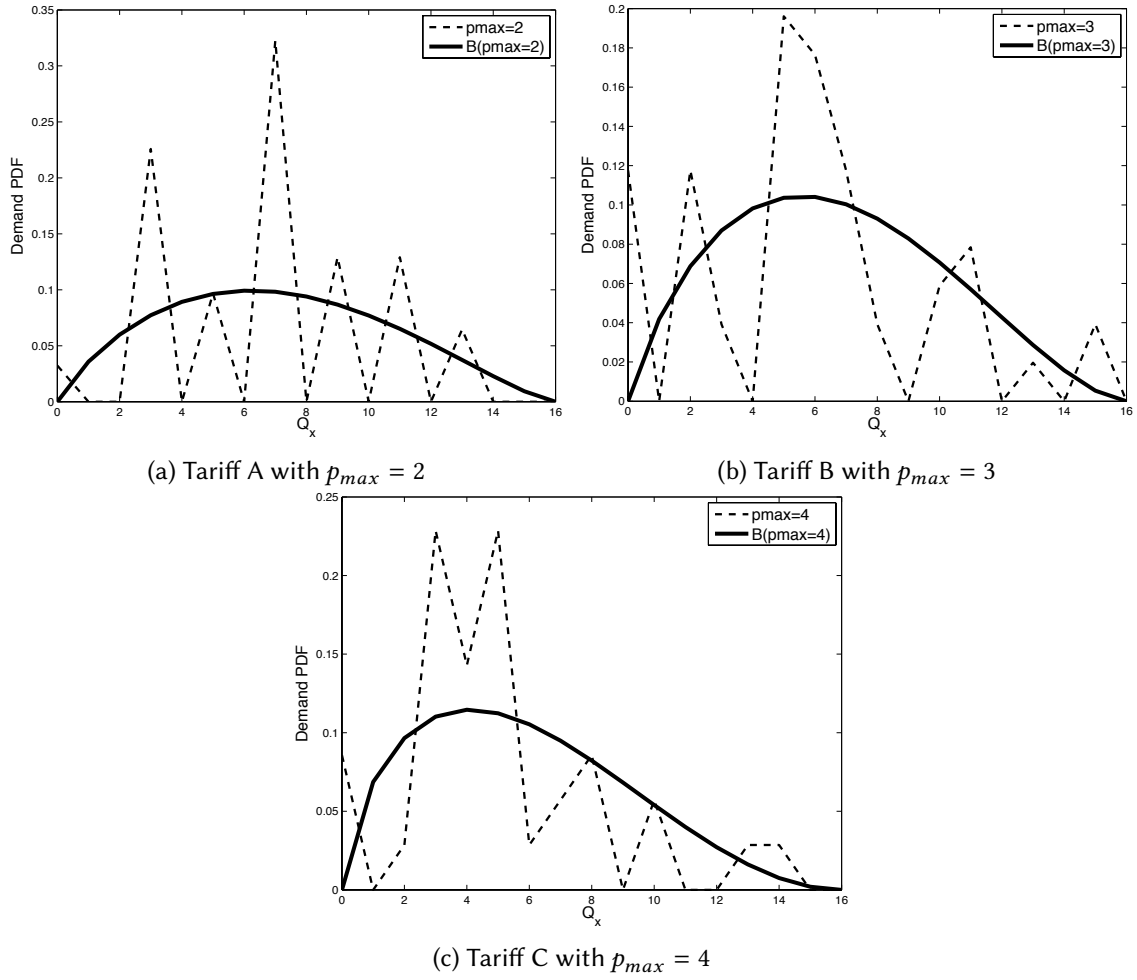


Figure 4.4: PDFs for d and $B(d)$ for varying p_{max} (excluding end points; bins of three classes).

of demand to **customer segments**¹³ when offering a set of alternatives (where the normal distribution is an acceptable fit, as detailed later on). In other words, when users are asked whether to buy a product at a certain quality and price (at given location, time, etc.), the demand (correlating to WTP values and revenues) can be approximated using the given beta regression functions, which is in line with suggestions in [Cow07].

Thus, in the case of the *2012 trial* (and in any comparable study) a sharp separation of WTP and customer segment views is necessary: Resorting to literature [Cow07] and the spot on CDF prediction, we can, however, reasonably predict WTP trends (ie, global view) from our empirical data (see Conclusion #4.1), ie, the essence of WTP predictions is beta distributed. The local view will be rather normally distributed (cf. Conclusion #4.2 and #4.3 as later on introduced), as discussed in Detail in Section 4.1.2.

Conclusion #4.1. *WTP is according to literature [Cow07] beta distributed. For the case of telecommunications products however, BE Internet (low cost, no quality guarantees) may represent a notable exception from an otherwise beta-distributed relationship. Further empirical testing is required.*

Conclusion #4.2. *Customer segments (attaching users to the best matching offer from an available set of choices) and typical demand-centric WTP figures (demand for an offer with a given price and quality without any choice), as addressed in [Cow07], are likely different concepts following different demand distribution curves. In other words, the demand for each combination of price and quality from a set of alternatives, ie, customer segments, needs to be separated from the overall purchasing trends, ie, WTP.*

In the future, more demand-oriented questions may be tested, eg, “Do you want to continue watching this video in the current quality for a price of €2.50?” (without any alternative, typical range or other bounds). This is comparable to the scenario where in 4 shops quality A is tested in *Shop 1* with one price, in *Shop 2* with an alternative price, while *Shop 3* and *4* test those two price levels for the same product in quality level B (quality $B > A$). Under such circumstances, the likelihood of retrieving a beta distribution will be much higher, exactly as envisioned in literature. In our case, however, regression to the mean biases may have been induced by offering a choice of different quality-price combinations at the same time.

4.1.2 Customer Segments

As customer segments in real world markets (where customers will most likely be confronted with a set of price and quality choices) do not follow a beta distribution in the *2012 trial*, we have undertaken a series of other tests. Among those efforts has been the normal curve fitting, as briefly sketched before, which has surprisingly provided the best fit to our empirical data, ie, see Fig. 4.1 and Fig. 4.2 for the high expression power of the normal

¹³ Customer segments are the distribution of demands to particular offers (ie, a choice of quality-price combinations).

distribution fit, which has outperformed other distributions and reasonable polynomial fits.

In the *2012 trial*, the customer segments, as the demand distribution to the offered quality-price alternatives (ie, local scope), follow a normal distribution. The associated coefficients of determination increase with the p_{max} from 0.534 for tariff A to above 0.848 for tariff C¹⁴. Despite the very expressive approximation result of beta distributions for WTP CDFs, strictly better results have been obtained for the normal distribution: 0.979 (tariff A), 0.992 (B) and 0.989 (C). This outcome is reflected in the general Conclusion #4.3.

Despite the relevance of the beta distributed demand and thus WTP levels, customer segments in a single market (*local scope*) are of high economic value alike.

Conclusion #4.3. *For the **normally-distributed customer segments** (in terms of demand being attached to offered quality and price combinations) the peak can be expected at medium quality and price levels (on a relative perspective between poorest and best offers).*

The general characterisation of the customer segments distribution is reflected in our Conclusion #4.4.

Conclusion #4.4. *The peak in the the normally distributed **customer segments** curve may, due to steeper or flatter price increase curves, slightly shift towards lower- or higher-quality classes respectively.*

Apart from the distinction between WTP and customer segments (cf. Conclusion #4.2), and the associated distributions, the *2012 trial* is considered to provide valuable indications for both perspectives on network video quality markets.

4.1.3 Generalisation

Despite the distinct shapes, distributions and values for WTP and customer segments, which see a difference between a more global and a more local view on network video quality purchases, this effect is not unique to this kind of trial, as will be shown in a specific though experiment¹⁵: Scientific laboratory and field tests, especially in the QoE context, often centrally focus on local views that are hard to be generalised. This effect is very pronounced in QoE testings, due to the often used (eg, see [Str12]) intentional bias of priming subjects in dedicated training sessions, ie, as recommended by [P.910; P.800; P.1501]. In training sessions, subjects are exposed to the used test ranges, services and scenarios prior to the main trial. Users are typically introduced to best and/or poorest quality levels at the beginning of the trial in order to eliminate external effects and to increase the controllability of the experiment. The consistency of QoS-QoE curves, empirically often of logarithmic shape [Rei+10] correlating to pure human perception of differences

¹⁴ The coefficients of determination for tariff A and B are still rather low due to outlier effects in the small sample sizes.

¹⁵ To the best of our knowledge such a thought experiment has not been discussed in literature so far. Such an experiment is, however, necessary to illustrate the role of QoE data in the commercialisation of network qualities.

(cf. Fechner scale [Hei04]), appears to prove this approach right. However, to the best of our knowledge, limited efforts have been dedicated to understanding the macro-pictures, ie, how to transition from locally normalised data bound to a specific context to a more global understanding of appreciation for QoS characteristics. In analogy to the previously discussed WTP trials, QoE similar to customer segments curves would also need a more global counterpart in analogy to WTP. For simplicity, we will denote this more general cross-trial QoE representation as *global QoE*.

In our thought experiment will further strictly differentiate between local, referring to the common QoE results in literature), and global QoE views. Both views have their purpose, but for different kinds of assessment: 1) The local QoE and customer segments view, following logarithmic and normal distributed relationships respectively, best represent the market environment where users are directly confronted with a set of choices. QoE may, in this case, serve as local CS metric while customer segment results may give indications on resource requests when offering a specific set of a priori parametrised products. 2) The global view for QoE and WTP, representing an unknown relationship or a beta distribution respectively, may provide highly relevant information for scientific and market optimisations. In the case of WTP, it may give indications for the optimal selection of appropriate quality and price combinations, which yields a maximum revenue for a given capacity. In the QoE case, it may represent a cross-testing and potentially even cross-service appreciation of QoS by users. To date, such data do not exist to the best of our knowledge.

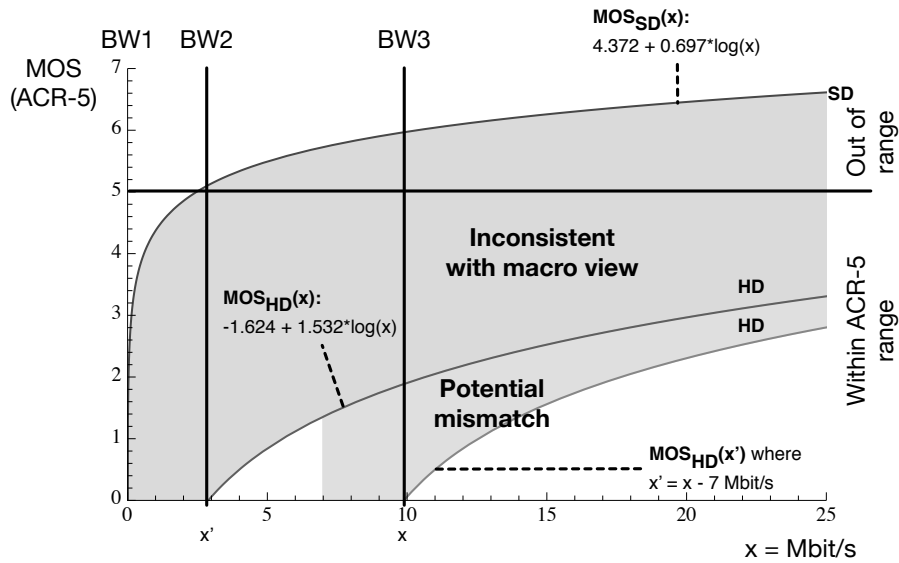


Figure 4.5: Mismatch between local QoE views (see curves) and macro scope assessments (excerpt of the entire test range).

For illustration purposes, consider the fictive scenario of three assessment cases that

study the effect of bandwidth variations on the QoE in the form of MOS ratings in ACR scale. The first assessment uses the SD video results from [Fu+13] (reconditions using a logarithmic approximation in this work; based on the recommendation in [Rei+10], which outperforms the exponential fitting as preferred in [JL10] for video),

$$MOS_{SD}(x) := 4.355 + 0.697 \cdot \log(x) \quad (4.7)$$

with x representing quality level in the form of video bitrates in Mbit/s.

The two other cases rely on the HD video results from [KRD12] from which the following curve shapes have been logarithmically approximated:

$$MOS_{HD}(x) := -1.624 + 1.532 \cdot \log(x) \quad (4.8)$$

The third case, however, replaces x by x' where $x' = x - 7\text{Mbit/s}$ (where x is the used bandwidth or bitrate in Mbit/s; x' shifts the test range towards higher QoS levels). The entire scenario is illustrated in Fig. 4.5, which obviously leads to some remarkable results:

- The local view of the SD and HD assessments are incompatible, ie, QoE ratings are inconsistent between service types. A user might rate a 10 Mbit/s SD video as perfect (in our approximation above the maximum of 5) while “poor” results would be obtained in the local HD test case. In reality, however, the HD video encoded in 10 Mbit/s may excel the perceived quality of the SD video in identical bitrate. Thus, the local and global views are incompatible.
- The case of the two HD videos illustrates the potential effects of testing quality levels within given ranges. In the fictive example of $MOS_{HD}(x')$, higher quality ranges are tested. In both cases $MOS_{HD}(x)$ and $MOS_{HD}(x')$, due to intentional test biases, curves will converge to a minimum and maximum QoE level within the tested range. Due to limitations of human perception, see Fechner scale [Hei04], a logarithmic shape will likely result in both cases. Thus, an inconsistent area between both curves will arise.

Both inconsistencies illustrate the local character of QoE ratings obtained from typical research trials. A cross-service or cross-trial perspective has not yet emerged. A solution approach for the specific case of controlled cross-service degradations will be introduced in Section 5.3.

Conclusion #4.5. *Within a single assessment context or trial, QoE perception may follow a logarithmic shape between the poorest and best quality level. Across trials, this perspective may, however, lead to inconsistencies that may hamper the meta analysis and commercialisation. A globally applicable QoE metric has not yet emerged, but would be required for practical utilisation of QoE information.*

The thought experiment and the presented conclusions are subject to the available empirical data. To verify, falsify or extend the assessment of this experiment we recommend further experimentation, eg, by conducting specific empirical trials.

4.2 Single-class Service Provisioning

♣¹⁶ Based on the gathered information on (non-)competitive demand patterns in Section 3.3.3 and the generalisations in Section 4.1, the core of the supplier’s problem is to optimise profits based on the expected spending behaviours. In this light, a single-class optimiser is formulated in this section, which is later on extended to a multi-class profit optimisation in Section 4.3. Due to the typically static capacity C of a network link, network resource of even entire network, and the thus almost static costs, both optimisers will focus on revenues. In the present context, revenue is primarily characterised by the demand and price figures provided in Section 4.1.

Such a single-class optimiser needs to run at a central point in the network where the session demands of all users are known or can be predicted for a period. The configuration of a traffic class, eg, w.r.t. QoS settings and access policies, is likely changed on a medium-term basis (eg, hourly or daily), as any update needs to be best on results of an optimiser and a change needs to be propagated through the affected network¹⁷. Due to the enormous scale of the Internet, even when restricting the perspective to a particular Domain, a highly efficient mechanisms is advised.

There exists a vast amount of literature on various auction designs such as Progressive Second Price (PSP) [LS98] and their properties [Tuf02]. Auctions can be seen as tools for assigning winners to a constrained amount of available goods, eg, a single-class network resource assignment. When considering experience goods, as in our case, subjects are however placed in an awkward position where they have to a priori evaluate an unknown kind of service in order to formulate their bid for the auction. This pitfall effectively reduces the practical value of an auction of any kind.

The complexity of running auctions (including the interaction with the customers) constitutes a substantial burden on a resource assignment system, which is subsequently circumvented. For this reason, the previously discussed empirical probing is used to estimate the underlying population’s demand for quality classes. These quality classes are then offered for purchase, ie, a mechanism is designed where customer segments are explicitly targeted based on empirical evidence and revenue optimisation considerations.

4.2.1 Base Model

The transferability of NSP’s utility from obtained revenues is assumed—ie, the NSP rationally assesses the revenue growth. However, neither quality perceptions, ie, QoE, nor the WTP linearly increases with the invested resources—QoE often follows a concave, eg, logarithmic [Rei+10], pattern. To avoid the difficulty of appropriately approximating the WTP from QoE, the supply model is constructed on the basis of the *2012 trial* data (cf.

¹⁶ *The present section is primarily based on materials that have been presented in Section 5 in our prior publication [ZRS14]. Since then, the precision of the results has been substantially improved.*

¹⁷ For example, traffic oscillations due to inconsistent states in the network have to be avoided, as it could affect the stability of the network.

Sections 3.3.3).

Due to economies of scale, investments (ie, CAPEX) will pay off more quickly under a *full utilisation of capacity*, ie, a broad user base. A more demanding customer who requires twice the resources of a discount customer (eg, through higher quality demands), may further likely require less than twice the costs in terms of customer support, for example.

Subsequently, a profit optimiser model π is constructed, which acts conversely to an auction where the best matching quality level q and its customers is served, while others are ignored (corresponding to an NSP offering a single revenue-optimal quality level). This system uses a strict access control, and active QoS and thus QoE management centred around empirical revenue expectations (WTP and demand). The system is further tailored to data obtained from *2012 trial* and the underlying QoS and price parameters. Classical QoS charging mechanism, as revisited in [Tuf04], often do not incorporate customer experiences influencing purchasing behaviours, while existing multi-class QoE charging approaches such as [WZR12] only approximate the WTP from QoE ratings. Despite the attractiveness of multi-class service offers, the lower technical complexity of access-controlled single-class systems may promote this straightforward design. For the context of this thesis, profit π can be formalised by:

$$\pi(Q, C, n) := \max_{Q_x \in Q} \mathcal{R}(Q_x, C) - \text{CAPEX}(C) - \text{OPEX}(C, n), \quad (4.9)$$

where Q are the quality classes from the *2012 trial* of the form $Q = \{Q_0, Q_1, \dots, Q_{15}, Q_{16}\}$, $\text{OPEX}(C, n)$, logarithmically flatten in n (the number of customers for the capacity C)¹⁸ due to efficiency yields, and $\mathcal{R}(Q_x, C)$ is the revenue generated from a single provisioned QoS class Q_x using the overall capacity C . In our case, the QoS depends on the user demand $d(Q_x)$ for video streams and the associated resource demands $\zeta(Q_x)$ for a single stream in the discrete quality class Q_x . The user demand is determined by the empirical findings of the *2012 trial* as generalised to beta distributions B in Section 4.1.1. The resource demand $\zeta(Q_x)$ for streaming a single video stream in the quality class Q_x is defined in kbit/s (bitrate of the video stream) as follows:

$$\zeta(Q_x) := 128 \cdot 2^{\frac{x}{2}} \quad (4.10)$$

with $x = \{0, \dots, 16\}$ as Q_{17} - Q_{19} have been introduced as virtual quality classes with requirements identical to class Q_{16} . For Pareto-efficiency¹⁹ $[d(Q_x) \cdot \zeta(Q_x)] = C$ for the only provisioned q_x holds. Otherwise a non-efficient state exists, where some of the capacity remains idle.

Due to the convexity of $\zeta(Q_x)$ in (4.10), an upper limit for the revenue $\mathcal{R}_i$ for user i (ARPU) follows by:

¹⁸ Total costs, thus, have a form of $C \cdot a \cdot \log(n) + C \cdot b$ with given a,b.

¹⁹ Pareto-efficiency may also be achieved for $Q_{>16}$, which is explicitly excluded from the model.

$$\mathcal{R}_i(Q_x) := p_{max} \cdot \frac{x}{19} \quad , \quad (4.11)$$

with linear pricing steps as described in Section 3.3.3. This is linked to the empirical expenditure distribution that captures the observed demand levels:

$$\mathcal{R}(Q_x, C) := \mathcal{R}_i(Q_x) \cdot \min\left(d(Q_x), C/\zeta(Q_x)\right) \quad , \quad (4.12)$$

where $d(Q_x)$ is demand for Q_x , which together with the capacity C and resource demands $\zeta(Q_x)$ limit the provisionable quantity (ie, video streams). With this, the profit given in (4.9) can be calculated. Due to the domination of $\mathcal{R}$ by demand growth, no interior extremum can be observed when optimising $\mathcal{R}$ for Q .

Conclusion #4.6. *For a single-class service provisioning, the revenue $\mathcal{R}$ for network services continuously increases in demand, and, thus, theoretically goes to ∞ for infinite capacity ($C = \infty$). Hence, the provisioning of the a QoS class Q yielding the individually highest demand is naïvely preferable.*

For CAPEX and OPEX directly increasing costs can be recognised with capacity, thus monotonically increasing in C . Hence, a minimal C to satisfy rewarding demand is, of course, beneficial. Based on our definitions, costs logarithmically increase in n (while bargaining powers of individual users will decrease). Thus, marginal costs are decreasing in n .

Conclusion #4.7. *Whenever no sufficient premium demand exists, ie, very few customers utilising the capacity C under acceptable bargaining powers, a substantial extension of the customer base may provide cost advantages over competitors due to scaling effects.*

When optimising $\mathcal{R}$ over the empirical data gathered in Section 3.3.3 (43 participants; 2012 trial) on a link with 10 Megabit/s, $x = \{7, 6, 5\}$ for Q_x for the respective p_{max} of 2, 3 and 4 yield the respective maxima (see revenue details in Table 4.1)—reflected in #4.8.

Conclusion #4.8. *For a single-class service provisioning, VoD services—under the given tariffing design and controlled conditions—should be provisioned with intermediate quality.*

This effect caused by the demand domination of revenue as specified in (4.12). As the demand d shifts towards slightly lower QoS levels (and the respective quality classes Q_x) when increasing p_{max} , ie, an increased concentration on lower qualities results, the revenue optimality shifts to intermediate quality levels.

When using $B(d)$ instead of d , the population is broader distributed over all available quality classes. This change lowers the demand for particular classes and, thus, decreases the revenue $\mathcal{R}$ when provisioning only a single class. This transitively also increases the optimal x^* .

We can further observe from Table 4.1 that the maximum price p_{max} can positively affect the obtained revenue within reasonable ranges.

Table 4.1: Revenue optimal quality provisioning.

	Tariff		
	$p_{max} = 2$	$p_{max} = 3$	$p_{max} = 4$
$d : Q_{x^*}$	Q_7	Q_6	Q_5
$d : \mathcal{R}^*$	5.16	8.52	8.42
$B(d) : Q_{x^*}$	Q_{12}	Q_9	Q_{10}
$B(d) : \mathcal{R}^*$	1.26	4.26	4.21

Conclusion #4.9. While price increases (raised maximum price p_{max} and price steps) first yield higher profits, this effect flattens with increasing price differences.

4.2.2 Extended Model

Under a game-theoretic reasoning, it is assumed that customers who are not being served their primary choice Q_{x^*} may be most interested in slightly higher or lower values than x^* (rendered optimal due to demand fluctuations and access control). Thus, the base model is extended by accommodating for a linearly decreasing disposition to accept deviations $\epsilon = \{-8, \dots, -2, -1, 0, +1, +2, \dots, +8\}$ from their Q_q choice,

$$d(x, \epsilon) = \max\left(0, d(x + \epsilon) \cdot \left(1 - \left|\frac{\epsilon}{8}\right|\right)\right) \quad , \quad (4.13)$$

where $d(Q_x, \epsilon)$ is the demand for the provisioned quality class Q , as linear relaxation around the original Q . The number 8 arises from the range of ϵ in $[-8, 8]$, which assures that at 0 the entire demand is captured and at the extrema (or whenever the demand is 0) of ± 8 the demand always drops to 0.

When utilising the beta smoothed demand function $B(d)$, the analysis can be upscaled to any population and link size without overweighting empirical outliers. By running the simulation with the extended model, the empirical data (original population and customer segments of the *2012 trial*) can be compared with an **unconstrained capacity** case. For example, as applied subsequently, the initial test population can be replaced by **100k** users. Obviously, the revenue substantially increases with the population size but also with an increasing capacity—see Table 4.2.

Table 4.2 illustrates that the increased demand over the initial model, due to the intake of nearby demands, drives the optimum towards lower quality levels when capacity restrictions are tight, eg, 10 Mbit/s. The opposite happens under unconstrained capacities, as, the revenue efficiency, ie, the fraction of charged price and the used resources, is dominated by the “filling the pipe” strategy. When also increasing the population to 100k, the optimal quality classes remain almost identical to the unconstrained case with the

Table 4.2: Revenue optimal quality provisioning with modified demand.

	Population					
	2012 trial (35) ²⁰				100k	
	Capacity					
Tariff	10 Mbit/s		Unconstrained			
	Q_{x^*}	$\mathcal{R}^*$	Q_{x^*}	$\mathcal{R}^*$	Q_{x^*}	$\mathcal{R}^*$
$p_{max} = 2$	Q_5	7.37	Q_9	17.05	Q_9	52.39k
$p_{max} = 3$	Q_3	12.79	Q_8	32.84	Q_8	62.98k
$p_{max} = 4$	Q_4	16.84	Q_8	28.63	Q_7	79.99k
	$B(Q_{x^*})$	$B(\mathcal{R}^*)$	$B(Q_{x^*})$	$B(\mathcal{R}^*)$	$B(Q_{x^*})$	$B(\mathcal{R}^*)$
$B(d); p_{max} = 2$	Q_7	5.16	Q_{11}	8.11	Q_{12}	47.93k
$B(d); p_{max} = 3$	Q_4	12.63	Q_{12}	30.32	Q_{11}	58.06k
$B(d); p_{max} = 4$	Q_4	15.16	Q_{10}	25.26	Q_{10}	68.99k

initial population. Hence, capacity constraints and demand levels are most decisive for the quality class optimisation.

Conclusion #4.10. *For a single-class service provisioning, high demands and low capacity shift profit optimum towards low-quality services (lower x values and the corresponding quality classes Q_x).*

Especially in the unconstrained case, $\mathcal{R}(d)$ outperforms $\mathcal{R}(B(d))$, and q^* again rises with the approximation $B(d)$. This outcome is likely related to the smoother distribution of demand levels to individual quality classes, which may shave demand peaks. Hence, in a multi-class model this effect will likely vanish—see Section 4.3.

We can further conclude that any $Q_{x>12}$ (or Q_9 when using unsmoothed data) is less profitable than optimal intermediate quality:

Conclusion #4.11. *For a single-class service provisioning, the premium segment (high quality, high price offers) is unattractive even under unconstrained capacity.*

When linking Conclusions #4.7 and #4.11, an extensive customer base (ie, highly demanded low/intermediate quality levels) is beneficial, which eventually represents another antagonism with respect to Conclusion #3.14. It can further be stated that “filling the pipes” strategies can, depending on load levels, affect the target QoS when offering services to customers, and may still be rational in cases of almost unconstrained capacity.

4.3 Multi-class Service Provisioning

Instead of targeting a single quality class to be provisioned on empirical evidence, the presented optimiser is extended to a k -class optimiser, ie, a multi-class QoS service with

exactly k QoS classes. The k -class design with a fixed number k is required in order to keep the technical complexity limited at later stages, as otherwise the number of classes could grow in an uncontrolled way (likely increasing towards infinity with a growing number of subjects). Such a system would be hard to realise, maintain and even market. While also techniques known from auction design are used in this section, an actual auction may be avoided whenever no clear communication strategy for conveying QoE-differentiated products to consumers is found (see [Var+15]).

4.3.1 Background

In literature many auction mechanisms with beneficial properties (eg, incentive compatibility) exist, in the network scene most notably the smart market approach [MV95], PSP [LS98], and MIDAS [CDS01]. Especially PSP is a well-engineered mechanism substantially reducing the messaging exchange complexity through packet-wise approaches over the smart market approach. PSP follows a classical Vickrey second price auction approach where after a convergence phase the highest bids are served at the price of the first unconsidered (‘losing’) bid. A multifaceted range of extensions and optimisation such as [Tuf02] or [JW08] exists. The latter approach replaces per-link auctions by per-path bids, which has a clearer end-to-end orientation. Although PSP and its various flavours are proven to be incentive compatible, leading to truthful bidding behaviours, and come with an acceptable complexity of $O(n^2)$, where n is the number of players, it assigns “arbitrary shares of the total quantity resources” [LS98] (in reference to the quantities requested in the bid by subjects). In our case, however, the QoS level has to be exactly matched and guaranteed where an equilibrium with high user demand and revenue efficiency exists. In other words, PSP auctions do not correspond to the view of empirically-backed QoS guarantees where customer requests are guaranteed or not provisioned at all due to the unclear demand or WTP for alternative quality classes. They further do not support a multi-class design with a fixed number of classes k , which could entail some complexity issues in the technical realisation. Apart from that, auction designs such as PSP not only require intensive interaction with subjects (eg, several bidding rounds), they also assume the existence of an often continuous utility function, ie, the more resources are provisioned, the higher the utility. This assumption is, however, not covered by our empirical findings. In particular, in our case we have been able to illustrate that QoS can only trigger further purchases, based on underlying utility evaluations, up to the point where the absolute WTP is reached. The simplification used in PSP is, hence, not valid in this context. Thus, a more discrete handling of demand and WTP levels is required. MIDAS in turn uses a Dutch auction design comparable to PSP, which, unfortunately, cannot substantially clear away or alleviate any of the obstacles mentioned above for our design. As a result, more complex winner determination mechanisms are necessary, ie, a k -class combinatorial winner determination.

The combinatorial winner determination (for auctions) is often approached by the primal-dual optimisation technique or specific auction designs [Nis+07]. A notable ex-

ample of a primal-dual-based solution is Kelly's resource allocation model in the case of TCP's congestion control [Kel03]. The subsequent elaboration will, however, concentrate on sketching the approximation for auction-oriented schemes.

The complexity of the winner determination for combinatorial problems, known to be NP-hard [CDS01], requires a special approach in our case²¹. According to Proposition 11.6 in [Nis+07] even the approximation of any combinatorial auction with single-minded bidders better than $m^{1/2-\epsilon}$ is NP-hard. Hence, Nisan *et al.* [Nis+07] recommend the theoretically best but greedy non-Vickrey-Clarke-Groves (VCG) mechanism. The VCG auction design is incentive-compatible²² and comes with a series of desired characteristics²³. In the polynomial approximation—with guaranteed $O(a\sqrt{m})$ for m resource items and resource allocation a among all players n —the incentive compatibility can be retained if an objective winner determination function optimises social welfare. Other objective functions may raise the complexity. The approximation²⁴ functions by descendingly ordering the bids as such that the list of bids starts with the bid i with the highest the highest bid value b_i ,

$$i = v_i / \sqrt{|S_i^*|} \quad , \quad (4.14)$$

where v_i is the bid value and $|S_i^*|$ is the resource request of bid i . The set of winners is now iteratively formed by adding the next best bid from the list until no more bid can be satisfied within the capacity limits C . The payment p_i of each winning bid $i \in W$ (where W is the set of winning bids) is determined by (see [Nis+07]):

$$p_i := \frac{v_j}{\sqrt{|S_j^*|/|S_i^*|}} \quad , \quad (4.15)$$

where v_j is the value of bid with list index j and j is the smallest index (in the ordered list of bids according to their bid value b_i) such that bid j is not included in the set of winning bids W . This means that the value v_j and the resource demands $|S_j|$ (relative to the demands of the winning bid $|S_i|$) for the first losing bid j determines the payment p_i for each winning bid $i \in W$.

Better results may be obtained by linear programming techniques (eg, as primal-dual) or more complex heuristics. These techniques go beyond the scope of this thesis.

4.3.2 Model

A winner determination phase comparable to combinatorial auctions can be applied where users request a certain quantity of resources, ie, their primary QoS class q^* , with

²¹ Due to the massive scale of the Internet, a non-polynomially solving algorithm may be infeasible. Thus, a computation of a sufficiently large populations or numbers of requests needs to be taken care of in practice

²² Incentive compatibility is satisfied when the truthful bidding is the rational strategy for bidders.

²³ Details on VCG are explained in [Nis+07].

²⁴ The presented approximation intentionally uses the notation used in the original source. Factors v_i should thus, for example, not be confused with the corresponding Gini coefficient parameter.

the preset price p . Due to the degrading interest of customers in other QoS classes than their primary choice, several other bids are automatically created with a price p' (where $p' < p$) decreasing with distance to q^* (see Section 4.2.2). The most rewarding set of QoS is then to be provisioned by running winner determination algorithms on top of the available set of information.

On a high level there are two kinds of optimisation axes for the approximation of the winner determination: 1) **frequency** of running the winner determination algorithm, and 2) **complexity** of winner determination algorithm (or heuristic).

On the frequency axis, the general setup of our auction allows for integrating the Bessler & Reichl optimisation [BR06] for comparable auctions, ie, in our case a decoupling of winner determination phase from the service provisioning itself. In practice, service quality levels may, for example, be defined per month for most relevant services and may later on correspondingly be provisioned and charged for a whole month.

On the complexity axis, bids can directly be aggregated by drawing on our histogram data from the *2012 trial* for each individual QoS class, ie, effectively reducing the number of players and, thus, bids to be considered (ie, the computational effort for a is reduced). Further, whenever price and quality classes follow a common characteristic, eg, a linear increase, the problem scope can be substantially simplified. This is not a feasible reduction in our case as it would neglect the existence of contrary empirical evidence. The details of this model are specified in the following.

Let us formulate a discrete objective function $\mathcal{R}^*(Q = \{Q_1, Q_2, Q_3, \dots\})$ for the three or more optimal quality classes²⁵ $K^* = \{Q_1, Q_2, Q_3, \dots\}$ ($|K| = k$; $K \subseteq Q$) that maximises the revenue²⁶ $\mathcal{R}$ per resource block. The perspective of resource blocks is chosen due to common practice of block-wise resource assignments in cooperative game theory (for example see [Nis+07]). The corresponding objective function $\mathcal{R}^*$ identifying the optimal set of quality classes $K^* \subseteq Q$ can be formalised as follows:

$$\begin{aligned} \mathcal{R}^*(Q) &:= \max_{K^*} \sum_{Q_i \in K^*} \frac{p'_i}{\sqrt{\zeta(Q_i)}} \cdot d_i \\ &s.t. \sum_{Q_i \in K^*} \zeta(Q_i)' \cdot d_i \leq C / \zeta_{gcv} \end{aligned} \tag{4.16}$$

where $p'_i := \frac{p_i}{\zeta_{gcv}}$
and $\zeta'_i := \frac{\zeta_i}{\zeta_{gcv}}$

with ζ_{gcv} representing the greatest common divisor for the resource demand ζ of all qualities classes, ie, $\zeta(Q_i)$, K is the set of winning quality classes, C is the available capacity, p' is the payment for a resource block of Q_i and ζ'_i is the number of requested resources for

²⁵ A quality class Q refers to a discrete q level.

²⁶ In the presented case, the optimisation of the revenue correlated with the maximisation of social welfare. This results from the bids (originating from WTP testings) that truthfully describe the serve valuation of the bidder (in our case the customer segment).

Q_i . The price for a block of size ζ_{gcv} is p'_i for any quality class $Q_i \in K^*$. In similar fashion, ζ'_i represents the number of standard resource blocks of size ζ_{gcv} .

While the stated objective function may also be used for a *primal-dual* optimisation model (constructing the dual as opposite of the designed objective function), the process is subsequently only illustrated for the theoretically best *polynomial time approximation* [Nis+07], as indicated before. This process will ignore revenues less than or equal to 0²⁷. The bid value b_i will be

$$b_i := \frac{p'_i}{\sqrt{\zeta'(q_i)}} \quad , \quad (4.17)$$

where p_i is the price corresponding to the term v_i as used in [Nis+07]. The payments of all winning bids will be set equal to the corresponding bid value b_i , as only the winner determination process of combinatorial auctions is used to parameterise the quality class portfolio. Hence, no direct bidder involvement exists and incentive compatibility is of minor relevance at this stage.

4.3.3 Results

When rerunning the extended model²⁸ simulation for the multi-class case and using an efficient version of the described polynomial-time approximation, a list of winning (naïvely best fitting) classes is obtained. This list contains the revenue values in descending order (revenues are decreasing from the first to the last winning quality class)—cf. Table 4.3 for various population and link sizes as in Table 4.2. Primed quality classes, eg, $\{6'\}$, refer to partially provisioned classes due to capacity limitations. User bids are, as per definition of combinatorial auctions, entirely satisfied (a winning bid) or rejected (losing bid). In comparison to the the single-class optimisation, the best entry in the list of optimal Q_{q^*} slightly shifts towards higher q^* values, which becomes especially apparent under tight capacity constraints. Whenever capacity constraints are loose, the revenue gains due to multi-class optimisations become significant. Apart from the second case with unconstrained capacity and the population of the *2012 trial*, the highest p_{max} always leads to the highest revenue figures.

Conclusion #4.12. *Multi-class service provisioning yields higher revenues and, thus, revenues for capacity C than a corresponding single-class service provisioning. The effect becomes more distinctive under tight capacity constraints, where low quality offers are rendered more attractive in the single-class case, but not with the multi-class design.*

²⁷ Under the assumption that the maintenance and/or advertisement of quality classes will both increase the system complexity and costs, the utility would be negative and is hence eliminated from the optimal solution.

²⁸ The extended work works effectively whenever noisy data are received (eg, high peaks at class q and $q + 2$, but no demand for $q + 1$). Whenever the data are less noisy, it can be recommended to revert to the simple model due to overlapping demand figures that can negatively affect the approximation process.

Table 4.3: Revenue optimal multi-class quality provisioning with modified demand.

Tariff	Q_x^*	$\mathcal{R}^*$
Population: 2012 trial (35, beta regression smoothed using B), $C = 10\text{Mbit/s}$		
$p_{max} = 2$	$\{6'\}$	6.53
$p_{max} = 3$	$\{6'\}$	9.47
$p_{max} = 4$	$\{6'\}$	12.63
Population: 2012 trial (35, beta regressed smoothed using B), $C = \text{Unconstrained}$		
$p_{max} = 2$	$\{6, 5, 7, 4, 8, 9, 3, 10, 11, 12, 2, 13, 14, 15, 16, 1, (0)\}(all)$	82.11
$p_{max} = 3$	$\{6, 5, 7, 4, 8, 9, 3, 10, 11, 12, 2, 13, 14, 15, 16, 1, (0)\}(all)$	351.79
$p_{max} = 4$	$\{6, 5, 7, 4, 8, 9, 3, 10, 11, 12, 2, 13, 14, 15, 16, 1, (0)\}(all)$	293.68
Population: 100k, $C = \text{Unconstrained}$		
$p_{max} = 2$	(identical to 2012 trial pop. case)	496.573k
$p_{max} = 3$	(identical to 2012 trial pop. case)	633.095k
$p_{max} = 4$	(identical to 2012 trial pop. case)	746.112k

Conclusion #4.13. For multi-class service provisioning, more expensive tariffs typically yield higher revenues for a given capacity C .

Note that the highest value (first) in the ordered lists of bid values $\mathbf{b}_i$ is not identical to Q_q^* in Table 4.2. The difference results from the square root in the optimisation algorithms (as specified by [Nis+07]). When removing the square root not only identical quality class are returned, but also the revenue grows due to the (partial) intake of additional quality classes. While there exist greedy strategies to limit the number of returned quality classes, in future work the limitation of quality classes should be built into the optimisation process.

4.4 Offloading Network Traffic

♣²⁹ The previous sections have focused on the revenue-side of the supplier's problem. This section complements this perspective with efficiency considerations by investigating the potentials of offloading network traffic to reduce costs or potentially even increase revenues of NSPs.

Wi-Fi coverage currently experiences an intensive densification: Wi-Fi solutions have recently started to penetrate urban areas through a direct integration of Wi-Fi functionality in fixed line CPE, the direct marketisation of Wi-Fi offerings, and the formation of

²⁹ This section briefly revisits our work on network offloading strategies presented in [Zwi+13].

“crowd-sourced” global Wi-Fi networks³⁰ such as *fon*³¹ or *UPC Wi-Free*³². A broad Wi-Fi coverage with spatial overlaps of CPEs already exists especially in densely populated areas, which seems to initiate a convergence of fixed line and cellular network usage and commercialisation scenarios.

Paired with recent technical advancements reducing access barriers to Wi-Fi networks, forwardmost Subscriber Identity Module (SIM)-based authentication, IEEE 802.11u and Hotspot 2.0 [HSP2.0] plus client-side network selection, attractive conditions for large-scale traffic offloading from cellular to Wi-Fi networks are further rendered for combined fixed and cellular operators. For such operators, network offloading has an interesting cost dimension, as investments in costly cellular infrastructures may be reduced and scarce capacity may be utilised more efficiently. Considering a world of an almost ubiquity of Wi-Fi (in acceptable quality), a disruptive ecosystem change may be triggered where expensive cellular technologies such as LTE may be questioned. In [Zwi+13], we have concluded that this new environment may best target the growth market of static and nomadic services while cellular services, due to their seamless handover capabilities, may complement this view with mobile offers.

The associated network offloading topic has already hit the European and North American markets with, for example, SFR and Free in France, Comcast in the USA, Deutsche Telekom in Germany or Belgacom in Belgium substantially offloading traffic or even providing novel services around this paradigm. Network offloading, in this case, refers to transferring traffic from licensed and, thus, costly cellular spectrum (eg, using LTE) to the cheap unlicensed spectrum (eg, Wi-Fi).

4.4.1 Model

♣³³ Considering a static network snapshot $\langle loc, t \rangle$ where loc is the location and t the point in time, a competitive model is formed between a fixed line only operator i and a cellular plus fixed line competitor j . Both i and j provide Wi-Fi (cheap unlicensed spectrum u backed by a fixed line backbone) to their fixed line customers. Provider j can also consider offloading traffic from the cellular (expensive licensed spectrum l) segment to the unlicensed spectrum. The unlicensed spectrum is shared among i and j .

The QoS $q(d, C)$ is further modelled sensitive to traffic demands d and the available capacity C after deducting the intra-customer site traffic—comparable to [Rei+13a]. Only for the unlicensed spectrum case the joint capacity C^u (static for $\langle loc, t \rangle$) is shared among i and j . Thus, in analogy with the model in [SSW08b] the $q \in [0, 1]$ can be described as

$$q(d_i^u + d_j^u, C^u) = 1 - \frac{d_i^u + d_j^u}{C^u} \quad (4.18)$$

³⁰ Fixed line or other operators that deploy CPEs with Wi-Fi functionality, which shares the connectivity with other customers in proximity.

³¹ <https://www.fon.com/>, last accessed: 20th May 2016

³² <http://www.upc.at/internet/wi-free/> (German), last accessed: 20th May 2016

³³ This subsection is strongly and intentionally aligned to the presentation format in [Zwi+13] due to the model character.

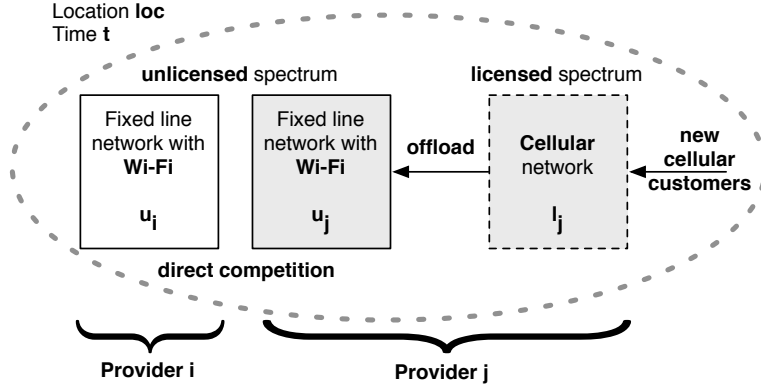


Figure 4.6: Competition between a fixed line only and fixed line plus cellular operator [Zwi+13].

where $d_i^u + d_j^u \leq C^u$. Conversely, $d_j^l \leq C_j^l$ is defined for the licensed spectrum.

The demand d is modelled as function $d^{(u,l)}(\Delta p, \Delta q)$ in relation to the customers' current contracts, which serve as reference points. Δp and Δq are defined as relative price and QoS tradeoff around the reference points, ie, the QoS and price gain or loss from status quo resp. As a result d decreases with increasing (positive) Δp , while it increases in Δq . Loyalty factors may blandish the assessment of the reference point's price and quality levels. For the offloading case from l_j to u_j , Δp and Δq can, thus, be formulated as

$$\Delta p = p_j^l - p_j^u \quad , \quad \Delta q = q_j^l - q_j^u \quad . \quad (4.19)$$

The revenue $\mathcal{R}$ of provider j for demand d can be described by

$$\mathcal{R}_j^n(d_j^n) = d_j^n p_j^n = d_j^n(\Delta p, \Delta q) p_j^n \quad , \quad (4.20)$$

where p_j^n describes the price charged by provider j for the network in j 's spectrum n (with $n \in \{u_j, l_j\}$), and by the sum of revenues in networks u_j and l_j ,

$$\mathcal{R}_j^u(d_j^u) + \mathcal{R}_j^l(d_j^l) \quad , \quad (4.21)$$

and conversely for i 's unlicensed spectrum network.

Whenever provider j offloads traffic, the appeal of the cellular network with q^l relatively rises at the expense of the Wi-Fi network u —cf. (4.18) and (4.19). Inferred from d , the prospective demand of j 's cellular network increases with its improved QoS or decreasing QoS of i 's network.

We further assume that the exogenously given E is the cost per bit per Hertz—besides OPEX they also encompass the CAPEX, eg, from spectrum fees or backhauling investments, in the form of depreciation costs—that satisfies $E_j^l > E_j^u$ due to inherent resource scarcity in cellular networks (limited excess resources due to limited available spectrum) in addition to high license fees. As a consequence, this negatively affects the price p^l for

cellular customers in practice ($p^l > p^u$), and, thus, also lowers the demand $\mathcal{R}^l(d') > \mathcal{R}^u(d')$ for any given d' .

4.4.2 Intra-provider Equilibrium

Under these conditions, offloading cellular traffic is **dominant** for j for the following cases:

- (i) **Quality gains** ($q(d_j^u) > q(d_j^l)$, ie, $\Delta q_j < 0$): This state allows the reselling of freed capacity in l_j (additional revenue), lowers the transport cost for the offloaded traffic (due to $E_j^l > E_j^u$) and may even increase the demand d^l due improved QoS whenever affected users remain unaware of the offloading procedure, ie, no disutility is created. All these benefits are obtained at the expense of negative externalities³⁴ in the equally shared spectrum (lower QoS; intensified competition for demand; profit loss is shared).
- (ii) **More rewarding traffic**: Beyond that point, network offloading is still dominant whenever cost reductions, $E_j^l - E_j^u > 0$, together with revenues from reselling freed capacity still exceeds the demand and revenue loss in u_j , ie, $\mathcal{R}_j^l$ or C yields outperform $\mathcal{R}_j^u$ reductions.
- (iii) **Bulk traffic offload**: When keeping sensitive traffic in the protected cellular network, the large-scale offloading of bulk traffic to the unlicensed spectrum is strictly dominant (selectively even for $q(d_j^u) \leq q(d_j^l)$) as long as the customers are not necessarily aware of this practice.

Around these considerations an **intra-provider offloading equilibrium** for two application types (extrema) $z = \{b, v\}$ is formed hereinafter, where the bulk traffic b has the lowest QoS demands while the video conferencing traffic v is assumed to have the highest QoS (ie, “premium” traffic). In analogy with the formulation in [GMS00], an equilibrium with demand d_j^{u*} exists whenever for any marginal piece of bulk traffic j is indifferent whether to offload to u_j or to keep it in l_j . For clarity this equilibrium is formalised as follows:

$$\begin{aligned} l_j \rightarrow u_j : & \left(q_j^u(d_j^u + \epsilon) < \underline{q_{z=b}} \right) \wedge \left(q_j^u(d_j^u) > \underline{q_{z=b}} \right) \\ \implies & q(d_j^{u*}) = \underline{q_{z=b}} \end{aligned} \quad (4.22)$$

for some $\epsilon > 0$. Thus, a convergence around the minimum quality requirements $\underline{q_{z=b}}$ for bulk traffic, ie, high utilisation of u_j , is inferred, which reflects the conditions (i) to (iii).

Conclusion #4.14. *Combined cellular and fixed line NSPs have a clear **incentive to offload** traffic from cellular to Wi-Fi networks using their fixed line backbone networks.*

³⁴ A definition and model representation for network externalities in the telecommunications context is given in [EH95].

Conclusion #4.15. *Combined cellular and fixed line NSPs have **no incentive** to protect a QoS level above bulk traffic requirements in the unlicensed spectrum, ie, an equilibrium at a high utilisation of unlicensed spectrum exists.*

4.4.3 Inter-provider Equilibrium

The shared Wi-Fi spectrum represents a common good that implies a resource competition between i and j . Comparable to the tragedy of commons [Har68], a “greedy” high utilisation by both players appears to be dominant, as lower utilisations by i would shift revenues from i to j as q^u rises and can be traded for q_j^l . In the opposite case, j ’s offloading dominance is already characterised by the intra-provider equilibrium.

Apart from that, the price p for any premium video traffic v is assumed to be greater than for bulk traffic, thus both i and j strive for carrying more v traffic (ie, the highest possible share of the overall v traffic), for which a high q_j^l will prove a strategic advantage. Around this strategic advantage for j a second stage *inter-provider equilibrium* exists, whenever j becomes indifferent between hosting the traffic in u_j or l_j :

$$\begin{aligned} & \left(q_j^l(d_j^{l''}) < \underline{q_{z=v}} \right) \wedge \left(q_j^l(d_j^l) > \underline{q_{z=v}} \right) \\ & \implies q(d_j^{l*}) = \underline{q_{z=v}} \quad , \end{aligned} \tag{4.23}$$

where $d_j^{l''} = d_j^{l''}(z=b) + d_j^{l''}(z=v)$ and

$$\begin{aligned} d_j^{l''}(z=v) &= d_j^l(z=v) + \epsilon \quad , \\ d_j^{l''}(z=b) &= d_j^l(z=b) - \epsilon' \end{aligned}$$

for some $\epsilon > 0$ and $\epsilon' \geq 0$ with sufficient premium demand $d_j^l(z=b)$ under the satisfaction of capacity constraints and a Pareto-optimal utilisation³⁵. Thus, by offloading traffic to u_j provider j will secure the maintenance of $\underline{q_{z=p}}$ in the l_j . In other words, j has no incentive to provision any $q(d_j^{l*}) > \underline{q_{z=v}}$ (upper bound) or to protect any $q(d_j^{l*}) > \underline{q_{z=b}}$. Thus, i ’s dominate strategy is to materialise all possible bulk traffic, as premium traffic under resource scarcity in $\langle loc, t \rangle$ will be captured by l_j .

Conclusion #4.16. *Combined cellular and fixed line NSPs can cannibalise the unlicensed spectrum to materialise on their **strategic advantage** of controlling a portion of the available licensed spectrum.*

Conclusion #4.17. *Combined cellular and fixed line NSPs have **no incentive** of protecting a QoS level above the most challenging service requirements in the licensed spectrum.*

³⁵ The capacity is fully utilised, thus resources can only shift between kinds of traffic, which transitively shifts revenues between $u_{i/j}$ and l_j .

4.4.4 Interpretation

Apart from the drawn conclusions, Zwickl *et al.* [Zwi+13] further argue that Λq may also be altered in the favour of provider j by overloading the shared spectrum l_j , ie, a relative QoS gain for l_j . In the extreme case, a complete blockage of the shared spectrum may be considered to display j 's strategic advantage. As neither the shared spectrum may be fully synergistically controllable, due to physical reasons and market competition, nor regulating authorities may accept the intentional blockage of valuable spectrum, the probability of this special case may still be regarded to be at most moderate for mature markets.

Paired with the current trend of global Wi-Fi access agreements and the drive of NSPs deploy and support seamlessly accessible Wi-Fi hotspots (for their customers), the parallel existence of fixed and cellular Internet access contracts for each customer may be questionable. Due to cellular customers' access to broad CPE-based Wi-Fi deployments, their interest in an own fixed line contract may wane. This represents another form of the tragedy of commons as free resources tend to be overutilised while contributions to a healthy and sustainable Wi-Fi environment are irrational. This issue could potentially lead to market oscillations where customers terminate or sign fixed line contracts based on the current acceptability of the Wi-Fi network in their area. For detail on such effects, we kindly refer to [Zwi+13].

4.5 Fairness & Regulation

As complement to the revenue- and cost-centric perspective of previous sections, the present section discusses conflicting interests referring to the Internet usage, the underlying resource distribution, relating revenue demands, and social values around the Internet (eg, NN)—as initially motivated in the Preface of this monograph. This thesis refers to these trade-off considerations between societal, individual and business interests as *societal constraints*. Such matters are addressed in this section through a presentation of a fairness & regulation dimension associated to the supplier's problem, which has the potential to disrupt optimisations for the network quality market.

Section 4.5.1 and Section 4.5.2, particular, start with fundamental considerations on club goods and fairness among users. The developed definitions will set the scene for network service provisioning strategies later on but will also substantially affect our view on NN. In Section 4.5.3, known definitions of NN and the legal situation primarily in Europe and North America will be revisited. This is followed by structural (see Section 4.5.4) and game-theoretic (see Section 4.5.5) assessments of the economic impacts of breaking content or application neutrality, as a specific form of NN.

4.5.1 Club Goods

Today's BE access to the Internet is sold as a *club good* [ST97]. Contrary to pure *public goods* (eg, parks, air or sunlight), NSPs have the power to restrict the resource access (ie, access fees, access policies and means of access) within defined boundaries. Inside the club, ie, the network of an NSP, intensive steering and discrimination of traffic is typically limited or even disallowed (also see NN in Section 4.5), ie, specific service degradations may not be used or may be difficult. Hence, once the access to the resources has been granted, a greedy utilisation of resources is dominant for the customers. The BE customers enter a competition for resources where collaboration is not incentivised. Whenever NSPs want to maintain acceptable quality levels, the pressure is relayed to themselves. In quality-differentiated networks, we can further assume that rational users will, irrespective of the actual service demands, always request the best QoS level unless the choice affects the service fee.

As a result, the resources of a publicly accessible network or a network open to user groups (Internet access as a *club good*) have to face rivalry among customers. In practice no inherent incentive is given for respecting the demands of other users, automatically implying competitive advantages of defecting, which transitively leads to a degradation of trust within the network. For example, a user can double-send packets to minimise the probability of packet losses or to mitigate or even equalise latency variances. Another strategy could be the masquerading of traffic as a service type that receives better QoS prioritisation (based on a constant MOS level for each service), eg, an audio stream is injected into a video container file for the transmission, and extracted at the receiving end point again.

Such kinds of selfish user optimisations, ie, defective strategies, may lead to an inadequate or non-optimal resource usage. A better quality level than required may automatically be provisioned, eg, the packet delay is kept low, although a download service may not be able to utilise this advantage. This strategy is inefficient, which may affect the economic stability of the industry. Some defective strategies may be eliminated by an **intelligent pricing scheme** as discussed in Section 2.1.3. Nonetheless, it is assumed that a complete protection against malicious practices may not be possible by purely adjusting the pricing modalities.

Defective strategies may, however, play a negligible role in **private networks**. Private networks may serve national security, military services, emergency handling, disaster recovery or enterprise means. Such networks may, thus, automatically come with a high degree of cooperation and collaboration among users (or hierarchies clearer regulating and moderating the usage of services). Commonly accepted strategic goals may exist, which may be respected by (or enforced for) a vast majority of all participants. Consequently, a trusted zone is established allowing the focus on their strategic goals rather than on circumvention strategies³⁶. From a humane and human behavioural perspective, positive

³⁶ This may, hence, come with a limitation of accessible or allowed services/applications to a set being compliant to

incentives and conventions are preferable to forces. The restriction to private networks is in most cases not applicable.

The per-user fairness, especially relevant for club goods, is investigated from a pricing perspective in Section 4.5.2 in order to mitigate some fairness issues in public networks. These considerations will be followed up by a discussion on NN, which has recently attracted a lot of interest.

4.5.2 Per-user Fairness & QoE

As many concerns about the fairness among users or usage scenarios may arise, an investigation of strategic alternatives (ie, the design of clever pricing regimes) is inevitable.

Fair Quantity-centric Pricing. Optimally, the prices of requested network resource disproportionately increase with the requested bandwidth or quality level in order to protect the satisfaction of the most essential communication requirements. However, these matters are more complex when closely examining the set of strategic responses of clients (players) requesting bandwidth under a quantity-centric pricing regime. Whenever technically applicable, the splitting of traffic in smaller portions or via separately signed contracts is individually rational whenever the steep segments of the pricing curve can be circumvented.

Fair Quality-centric Pricing. By contrast, options of customers may be more limited when requiring preferential treatment involving other quality characteristics. Such a case exists, for example, when assuring low latency figures or providing guarantees going clearly beyond BE Internet, for which the aggregation of small portions of traffic or the usage of multiple contracts for the Internet access is only hardly applicable. Thus, a market segmentation induced by **quality-centric pricing** may be more realistic than for pure quantity-aware counterparts. Hence, disproportionately increasing prices with quality gains can be used for protecting the provisioning of lower quality classes. The utility increase and demand for higher quality classes is linked inherently to the congestion rate of the network³⁷.

Price Curves. The resulting payments for additional QoS/QoE imply that the price is monotonically growing with the raise of QoS / QoE levels. Beyond that, the satisfaction of fundamental connectivity needs enforces a curve that has to become steeper along with growing QoS/QoE levels. A logarithmic price increase in quality would, on the contrary, lead to limited protection of basic connectivity demands, as the marginal price for a quality

the collaborative standards of the organisation.

³⁷ The QoE difference between BE Internet and “premium” offers may determine the upgrade value for customers. This difference may be moderated by the upgrading premium service or by introducing artificial scarcities in the BE segment unless avoided by regulators or by the competitive situation in the considered market. The practice of artificial scarcity is ethically dubious.

increase in the “premium” segment would become lower and lower. Consequently, the low quality offers would have to bear a high ratio of the overall costs, which could render basic connectivity more expensive. For example, with an exponentially growing price curve in quality the basic connectivity demands are protected while the premium segment is configured in a way to provide additional service offerings that support the functioning of the entire ecosystem.

Consequently, the QoS/QoE–price relationship is non-linear when aiming at enforcing some degree of fairness—cf. Figure 4.7a. Complemented by Homburg *et al.*’s [HKH05] observation of an S-curve-shaped relation between CS—in the current context corresponding to QoE—and WTP, plus assuming the QoE demands of the population to be normally distributed³⁸, for example, then an interior equilibrium around the more elastic inflection point will be obtained.

On the other hand, it would be useful to protect against users that request too high quantities of resources as insurance against packet losses or insufficient QoS levels. This would, however, imply a non-linear pricing (eg, logarithmic price increase in the quantities supplied) for the requested quantities, which could be circumvented (for users with the technical capabilities and multi-homing agreements) by splitting the sent traffic into several logical end points (see above). Thus, only a linear pricing scheme may work on a general basis.

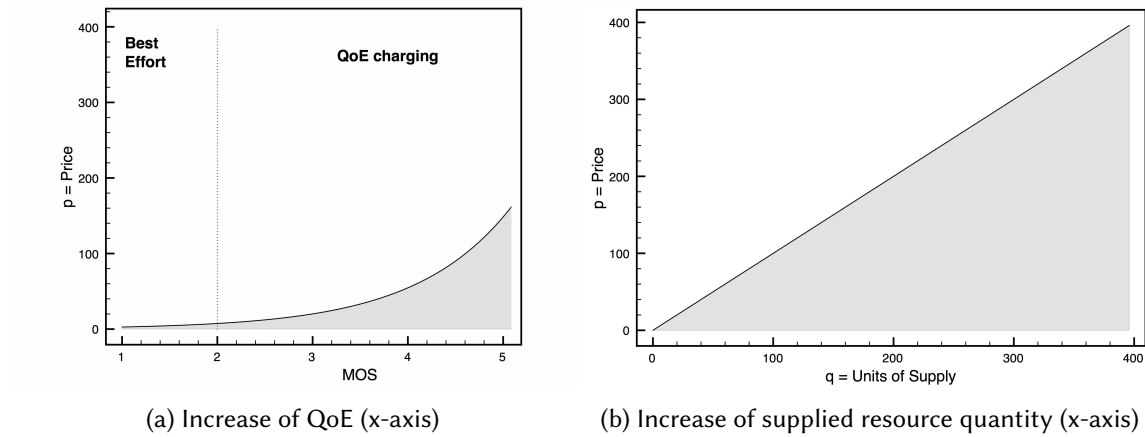


Figure 4.7: Generic price curves in relationship to QoE and supplied quantities.

These considerations can later be applied to empirical parameterisation efforts and revenue optimisations.

³⁸ A non-heavy-tailed distribution towards high QoE demands is at least assumed, whenever an appropriate granularity level for the price to QoE is chosen.

4.5.3 Net Neutrality (NN)

The satisfaction of social values, eg, NN, is treated with care when reconsidering pricing schemes for network quality differentiation (wrt third-degree price discrimination). The QoE- or QoS-aware pricing mechanism is placed in an area of tensions with NN [Wu03], for which a manifold of highly disparate definitions and regulations exist. These definitions are a result of highly intensive and very emotional discussions in media³⁹, industry and research, and include statements like:

“ The neutral communications medium is essential to our society. It is the basis of a fair competitive market economy. It is the basis of democracy, by which a community should decide what to do. It is the basis of science, by which humankind should decide what is true. Let us protect the neutrality of the net. ”

*Tim Berners-Lee, 2006*⁴⁰

“ A neutral network might be designed without legal prodding – as in the original internet. In an ideal world, either competition or enlightened self-interest might drive carriers to design neutral networks.

However, when that isn’t the case—when carriers are interested in discriminating for one of various reasons – matters get more difficult, and a law may be necessary. ”

*Tim Wu*⁴¹

“ Today the Internet is an information highway where anybody – no matter how large or small, how traditional or unconventional – has equal access. But the phone and cable monopolies, who control almost all Internet access, want the power to choose who gets access to high-speed lanes and whose content gets seen first and fastest. They want to build a two-tiered system and block the on-ramps for those who can’t pay.

Creativity, innovation and a free and open marketplace are all at stake in this fight. Please call your representative (202-224-3121) and let your voice be heard. ”

*Eric Schmidt (Google Inc.), 2006*⁴²

Such statements adduce good reasons to revisit the emergence of NN notion and its semantics by working through the manifold of available definitions, which follows hereinafter.

³⁹ “The Biggest Net Neutrality Lie of All”: <http://www.forbes.com/sites/larrydownes/2014/07/17/the-biggest-net-neutrality-lie-of-all/>, last accessed: 20th May 2016

“ The End-to-End argument says “don’t force any service, feature, or restriction on the customer; his application knows best what features it needs, and whether or not to provide those features itself.” ”

*Jerome H. Saltzer, 1999*⁴³

Saltzer *et al.*’s end-to-end design principle [SRC84] is often regarded to be a key principle to the Internet, which requires communication protocol operations to be placed close to the end points. It is, therefore, violated by any kind of data discrimination action. This principle has also significantly influenced the emergence of NN around [Wu03]. Later, discriminating actions have been further specified by Ofcom [Cro08], which distinguishes between *foreclosure* (eg, exclusion of rivals, content providers) and *sabotage* (referring to quality deteriorations). Around 2000, this argumentation provided the fundamental basis for the emergence of the NN discussion in the USA [LL01]. The NN notion itself is today attributed to Tim Wu’s work in 2003 [Wu03], which included one of the most prominent NN definitions still today: “maximally useful public information network aspires to treat all content, sites, and platforms equally”. Today, NN not only represents a “network design principle”⁴⁴, but especially a multi-disciplinary discussion with legal, economic, cultural and technological aspects [WAR11]. Wong *et al.* [WAR11] see the discussion related to a typical public good problem (such as cultural contents), which is rooted in an equality discussion, eg, referring to the definition of human rights. However, due to the charged access fees, Internet access services may be better classified as a club good (ie, a more specific form of public goods). For club goods, free-riding, for example, is possible once the access has been granted. The economic and legal problems relate to the moderation of fairness, which is in [WAR11] approached with alternating public good provisioning strategies.

More technical flavours of the discussion, eg, in [Cro07], have further detailed the openness of NN: In 2007, Hogendorn [Hog07] distinguished between the openness to the access network (“conduit”) and the openness of accessing contents. To his understanding, the latter represents a more accurate definition of such kind of openness that characterises NN. Around this idea of undiscriminated contents and applications, the sharp terms of *content neutrality* and *application neutrality* have emerged, eg, referred to in [CKA10; ACK10].

According to [Cro08] the Internet is a two-sided market around the consumer and content provider business where market failures are possible whenever “competitive bottlenecks” are created. In other words, no consumer is interested in the Internet, if they cannot use any services while the same holds for content providers who require a business case. With their analysis on two-sided market effects, Economides & Tåg [ET12] have been able to illustrate the possibility of positive and negative consequences arising from NN regulation.

⁴⁴ Network Neutrality FAQ: http://www.timwu.org/network_neutrality.html, last accessed: 20th May 2016

In 2014, AT&T, like other NSPs, started its sponsored data program,⁴⁵ which discriminates services through billing practices—further emphasising the market and business dimension of NN and discriminatory actions. Since the program does not account the data consumption from partnering services, it explicitly prefers those services in billing processes and raises their appeal to end customers.

The US American case further illustrates the legal and political dimension of NN. While the Federal Communications Commission (FCC), the responsible regulator in the USA, formulated a more ambitious “Open Internet Order” [FCCOI], only its transparency rule passed the federal courts⁴⁶. These rules contain an enforcement to disclose performance figures, eg, “network management practices”—details are presented below.

The following subsections will illustrate the current state of NN regulations in the European Union (EU) and the North America.

4.5.3.1 European Union (EU)

Following the EU’s Universal Service Directive (USD) 2002/22/EC [Eur02], minimum quality levels are required:

“ In order to prevent the degradation of service and the hindering or slowing down of traffic over networks, Member States shall ensure that national regulatory authorities are able to set minimum quality of service requirements on an undertaking or undertakings providing public communications networks. ”

Article 22(3) USD 2002/22/EC [Eur02]

In 2009, with the Telecoms Package an EU-wide regulating body, the Body of European Regulators of Electronic Communications (BEREC), was created. Following the USD, a common but vague set of common regulatory rules for the country-specific regulators was additionally defined.

In 2011, Neelie Kroes, European commissioner for digital affairs, called for a BE Internet “to which everyone has access” under “effective competition” with “transparency to allow consumer choice” and “ease of switching”. In 2014, almost at the same time as FCC’s frequent adaptations to the current legal situation in the USA (see below), the European Parliament⁴⁷ has voted for the equal treatment of services in the Internet. It has issued a ban for blocking or reducing the quality of specific services, by passing Amendment 237⁴⁸. Despite the broad support by society, the parliament’s ruling has been opposed

⁴⁵ <http://www.att.com/att/sponsoreddata/en/index.html>, last: accessed: 20th May 2016

⁴⁶ Open Internet: <http://www.fcc.gov/openinternet>, last accessed: 20th May 2016

⁴⁷ Press release by the European Parliament (EP): http://www.europarl.europa.eu/pdfs/news/expert/infopress/20140331IPR41232/20140331IPR41232_en.pdf, last accessed: 20th May 2016

⁴⁸ Trautmann, C., Andersdotter, A., Ernst, C. and Kari, R.R., “Amendment 237”: <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+AMD+A7-2014-0190+237-244+DOC+PDF+V0//EN>, last accessed: 20th May 2016

and subsequently undermined by efforts of member states to soften the NN definition. In November 2014, the Council of Europe (CoE), representing the member states, has issued the *Draft Recommendation CM/Rec (2014)* [COE15]. This recommendation supports the protection of equality rights on the Internet, but also suggests to allow traffic management practices as long as necessary. In 2015, the EP and Council of the EU (CoEU) have made an agreement, which provides more details on the softer interpretation of the existing NN regulations [CoEU 10409]. This position has raised many concerns in the media⁴⁹ immediately. The effects and detailed interpretations have remained unclear, however.

Austria: As reaction to the actions of BEREC, the Austrian parliament has passed an adaptation of the telecommunications law in 2011 (Telekommunikationsgesetz Novelle 2011⁵⁰), which extends the competencies of the regulatory authorities, ie, Rundfunk- und Telekom Regulierungs-GmbH (RTR), the Austrian regulator. While, before 2011, the RTR has mainly regulated the competition in the telecommunications market, after that it gained responsibilities towards protecting customer interests, defining minimum quality levels, and assuring transparency of services for customers. Due to the high market competition at this time, no definite actions towards NN regulation had been considered. In 2013, a RTR's position paper [RTR] on NN promoted the free choice of contents, service and devices, and has argued that any deviation from NN has to be transparent and justified and cannot entail a loss of quality. Adaptations of the telecommunications law have been initiated in 2015, which are still under review.

Netherlands: In 2012, the Netherlands, as the first country in the EU and the second worldwide, have passed a very progressive NN law⁵¹—Article 7.4a of the Telecommunications Act. This new law protects end users by disallowing wiretapping practices and strictly regulating conditions under which NSPs can (temporarily) disconnect their customers. Apart from that, Article 7.4a enforces the equal treatment of services and disallows any slowdown that is not caused by congestion itself. Notable exceptions refer to conditions of integrity/security issues, unwanted traffic or compliance with other laws.

Ofcom (UK): In [OFC11], Ofcom notes the importance of unrestricted access to BE effort as a driver for innovation. Improperly configured managed services (“premium” Internet services) hampering a proper coexistence with BE Internet (too low QoS) may endanger future innovation.

⁴⁹ Issie, Lapowsky (The Wired), “The EU Could Kill Net Neutrality With a Loophole”<http://www.wired.com/2015/06/eu-net-neutrality-loophole/>

⁵⁰ “Telekommunikationsgesetz Novelle 2011” (german): http://www.ris.bka.gv.at/Dokumente/BgblAuth/BGBLA_2011_I_102/BGBLA_2011_I_102.pdf, last accessed: 20th May 2016

⁵¹ “The Netherlands Passes Net Neutrality Legislation”: <https://www.eff.org/deeplinks/2012/05/netherlands-passes-net-neutrality-legislation>, last accessed: 20th May 2016

- An upfront definition by Ofcom [OFC10]: “there should be no prioritisation of any type of traffic by network operators; and those providing content, applications and services via the open internet should not be charged by network operators/ISPs for the distribution of that content to the network operator/ISPs’ customer base.”
- The major NSPs in the UK have voluntarily launched and signed the “Internet Code of Practice” in 2012⁵², which disallows the discrimination of competitors and blocking or degradation of selected services, as well as it assures transparency on all traffic management practices. To the best of our knowledge, this practice remains in full force until today.
- In Ofcom’s yearly plan for 2013/2014, traffic management practices (including those concerning NN) are regarded to be a non-issue for the UK market due to the absence of “substantive concerns in relation to the traffic management practices used by fixed ISPs”. In the mobile market, the transparency of traffic management practices are considered to be crucial. Thus, their position from 2011 [OFC11] is entirely retained.
- Ofcom [OFC11] further characterises active competition with **information transparency**, which allows customers to make an informed purchasing decision.

Other Countries: Some European countries like France (only recommendations based on a consultation analysed in [ARC10]), Germany or Italy have not issued any specific NN regulations. For more details on the European NN landscape, the dedicated section in [FP712a] is recommended.

Conclusion #4.18. *For the common Internet QoE-differentiation or any paid prioritisation is legally disallowed in the EU. However, “special services” can make use of such practices, based on the interpretation of the EP and CoEU that uses legal loopholes for a softer handling of NN in practice.*

4.5.3.2 North America

“ The “Open Internet” is the Internet as we know it. It’s open because it uses free, publicly available standards that anyone can access and build to, and it treats all traffic that flows across the network in roughly the same way. The principle of the Open Internet is sometimes referred to as “net neutrality.” Under this principle, consumers can make their own choices about what applications and services to use and are free to decide what lawful content they want to access,

⁵² BSG, “ISPs launch Open Internet Code of Practice”: <http://www.broadbanduk.org/2012/07/25/isps-launch-open-internet-code-of-practice/>, last accessed: 20th May 2016

create, or share with others. This openness promotes competition and enables investment and innovation.

*FCC*⁵³

”

While FCC’s position on NN (see above) actively targets the user’s perspective, it has also left many details open until recently. Building on US communication laws, Jordan & Schaffer [JS14] thoroughly reconstruct the rights for users and NSPs primarily from a device attachment perspective in the USA.

According to their position, the users are “entitled to”

1. “connect any legal device to a communications network” unless harmful for the network;
2. “run applications of their choice”;
3. a free choice of their NSP;
4. “transparency in terms of billing, traffic management, device restrictions and all other aspects of their communications (also see FCC’s Open Internet Order from 2010⁵⁴; services” (thus, strongly aligned to the position of Ofcom [OFC11]).

Correspondingly, NSPs are “entitled to”

1. “charge for communications services”;
2. “the use of reasonable network management”; (recommendation only)
3. “forbearance from regulations when they are not necessary to ensure user rights”.

While Jordan & Schaffer further argue that neither “fixed fees nor usage-based fees for communications services are controversial”, they have a different view on quality-differentiated pricing. While a QoS-differentiation for interconnection services such as peering should be unproblematic, towards the user the pricing should not be “unreasonably discriminatory”. To be precise, they see non-controversial discrimination only on layer 2 and 3 protocols [IEC 13236] (thus, network layer or below) comparable to up- or downstream bandwidth limitations. Based on this view, and according to the authors consistent with FCC regulations, they define “reasonable network management” as follows:

“ A network management practice is reasonable if and only if the user has control over the use of the practice or the practice controls Quality-of-Service on the basis of reasonable payment.

Joran & Schaffer, 2014 [JS14]

”

⁵⁴ FCC, “Open Internet Order”: https://apps.fcc.gov/edocs_public/attachmatch/FCC-10-201A1_Rcd.pdf, last accessed: 20th May 2016

However, FCC’s “no-blocking” and “no-unreasonable-discrimination” regulations have successfully been challenged in federal courts and, thus, had to be replaced by FCC recommendations.

In a prior publication [JG10], they have further detailed how reasonable common traffic management practices are—see Fig. B.1. Amongst other, they see degradation or enhancement of QoS unproblematic at or below the network layer. However, they consider a content- or application-based differentiation to be unreasonable.

Such an interpretation, thus, especially conflicts with the application-oriented differentiation of QoE, whenever the customers cannot control the applied practices. In our opinion, users further require an influence on the applied algorithms (automatisms), which allows the best redistribution of resources to their needs without blocking/degrading or preferring any specific service. For example, throttling external services, eg, Netflix, or prioritising the NSP’s or cable provider’s own services would be illegitimate against the interest of the customer. Apart from that, the QoE-differentiation is inherently application-sensitive and, thus, cannot solely be realised at or below the network layer, ie, layer 3.

From mid 2014 onwards FCC’s NN position became more turbulent. On the one hand, FCC’s position received a substantial shift of direction⁵⁵. Probably partially fed by prior court rulings and intensive market pressure, FCC proposed rules that still disallow the throttling of services of any kind. However, for the first time the ruling has actively allowed NSPs to charge content and application providers for faster or more reliable delivery. Relating this view to Ofcom’s position on discrimination, it can be concluded that while active sabotage of network services may still be disallowed, side payments would be enabled, ie, prioritisation of paid traffic, as an inverse action to active sabotage. Understandingly, this new position has received broad media attention such as the outrage by comedian John Oliver⁵⁶, intensive and highly emotional public feedback⁵⁷, (partial) disagreement by the current administration⁵⁸, and a series of protests by concerned application and service providers such as Netflix⁵⁹. Netflix, in particular, has regarded the proposal of the FCC more harmful than having no NN statement at all.

On the other hand, with *FCC 14-61* [FCC1461] the FCC has suggested banning any paid or unpaid discriminatory actions on the Internet. Only a few months later⁶⁰ the FCC has “passed the strongest Net Neutrality Rules in America’s History”⁶¹ in February

⁵⁵ FCC, “GN Docket No. 14-28”: https://apps.fcc.gov/edocs_public/attachmatch/DA-14-211A1.pdf, last accessed: 20th May 2016

⁵⁶ “Last Week Tonight with John Oliver: Net Neutrality” (HBO): <https://www.youtube.com/watch?v=fpbOEoRrHyU>, last accessed: 20th May 2016

⁵⁷ http://apps.fcc.gov/ecfs/comment_search/execute?proceeding=14-28, last accessed: 20th May 2016

⁵⁸ Bloomberg, “Obama Appears to Contradict Wheeler’s Paid-Prioritization Proposal”, <http://www.bna.com/obama-appears-contradict-n17179893364/>, last accessed: 20th May 2016

⁵⁹ <http://apps.fcc.gov/ecfs/document/view?id=7521491186>, last accessed: 20th May 2016

⁶⁰ <http://www.fcc.gov/document/fcc-adopts-strong-sustainable-rules-protect-open-internet>, last accessed: 20th May 2016

⁶¹ <http://gizmodo.com/fcc-passes-strongest-net-neutrality-rules-in->

2015—see *FCC 15-24* [FCC1524]. FCC is now actively regulating the Internet as a “public utility” where any form of paid priority handling of traffic is disallowed. With this, the FCC has disallowed any side payments or premium traffic models being operated on the Internet infrastructure. While end users may rejoice at FCC’s new rulings, legal pressures created by NSPs will again bear unpredictable risks for the continuity of FCC 14-61 or comparable rulings⁶².

Conclusion #4.19. *Paid QoE-differentiation appears to be unlikely in the USA but also in the EU. Depending on the convergence of the legal situation and the functioning of the market, niches or separate services may be offered. Single class QoE optimisation will not be targeted by current NN rulings, but may be of limited benefit to end users.*

Conclusion #4.20. *Globally, the transparency that allows customers to evaluate services properly, seems to be commonly understood as one of the key principles of NN and remains in full force, eg, in the UK and the USA.*

For the Canadian case, a Canadian Radio-Television and Telecommunications Commission (CRTC) rule from 2011 explicitly allows usage-based charging, which goes beyond right 1, which leaves many details open. The Canadian case is otherwise only regulated by general telecommunications laws.

4.5.4 Side Payments: Structural Market Power

This section investigates the economic appeal of content neutrality violations, and thus NN violations, from a market structure perspective. The results are later contrasted by a game-theoretic and price-centric perspective in Section 4.5.5. In particular, the structural analysis focuses on the introduction of charging regimes involving side payments.

4.5.4.1 Overview.

When considering a preferential treatment of one customer’s traffic or one service’s traffic, typically side payments come into play. Side payments refer to cross-payments between entities in order to assure the delivery of the desired service and quality. In other words, Internet access-specific payments are complemented by additional payments in order to compensate particular stakeholders for their efforts. Such side payments are typically considered to violate content or application neutrality, and are assumed to favour the receiver of the additional side payments, typically the NSP. To the best of our knowledge, the structural impact of transitioning to such a charging regime has not been studied in literature.

americas-h-1688204371, last accessed: 20th May 2016

⁶² Estes, A.C.: “Net Neutrality Wins: What Now?”, <http://gizmodo.com/net-neutrality-wins-what-now-1688183094>, last accessed: 20th May 2016

Typically CPs spend side payments to protect their business model from interference or even sabotage by NSPs. For example, Netflix has agreed on terms for side payments to Comcast⁶³ in order to provide the expected download speeds from their video streaming servers to their customers. The deal between Netflix and Comcast does not violate NN according to FCC statements—a view that is not shared by the **application neutrality** definition in [CKA10]. For completeness, works like [CKA10] follow a more generic viewpoint, which theoretically also considers side payments from NSPs to CPs, eg, the special case where an NSP aims at raising the appeal of its network. This view is confirmed by some recent concerns by cable companies⁶⁴ that see a rising risk of side payment from NSPs to CPs as well.

4.5.4.2 Quantification.

Market power shifts within a VN can be induced by the introduction of side payments. Such effects are studied with the help of the VNQ technique that has been introduced in Section 2.4. In particular, the quantification of Section 2.5.2 is rerun with a revamped VN that makes use of side payments and is otherwise kept identical to the reference case (also see [ZR15]). The revised VN drastically streamlines the business interactions, as the CP takes on a central role in all monetary business interactions. On the one hand, it collects the overall payments from the end users, but also directly orders the network transmission services from every required NSP. As other works have focused on modelling the revenue effects of side payments, eg, see [CKA10; Rei+13b] (see Section 4.5.5), this section will entirely focus on role alterations (structural shifts).

When using the input data from Section 2.5.2 and retaining the separate Internet access contract between end users and Access NSPs for basic Internet surfing, the revised dependencies depicted in Table 4.4 (with side payments) are yielded. The lowest dependencies (equals to the highest market power) exists for the CP and the Access NSP in the USA. Very high values are obtained for the end user, the Transit NSP and the Access NSP in the UK.

Table 4.4: VN dependencies in a VoD scenario (cf. Table 2.3).

	End user	CP	Access NSP (USA)	Transit NSP	Access NSP (UK)
w/o side payments	0.999	0.820	0.600	0.967	0.989
with side payments	0.998	0.755	0.620	0.991	0.988

⁶³ <http://time.com/80192/netflix-verizon-paid-peering-agreement/>, las accessed: 20th May 2016

⁶⁴ “Cable companies: We’re afraid Netflix will demand payment from ISPs”, <http://arstechnica.com/business/2014/07/cable-companies-were-afraid-netflix-will-demand-payment-from-isps/>, last accessed: 20th May 2016

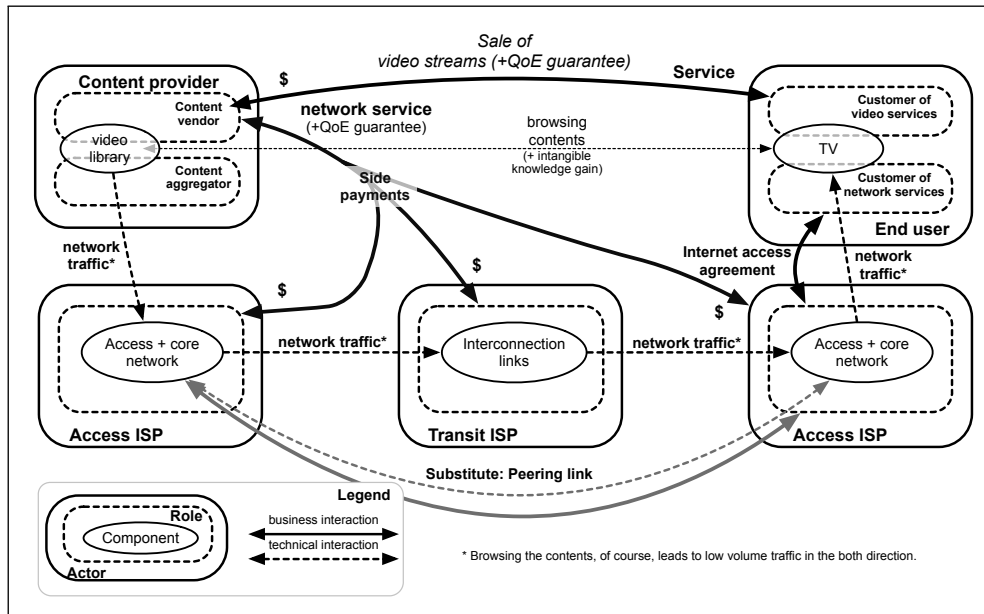


Figure 4.8: VoD VN with side payments

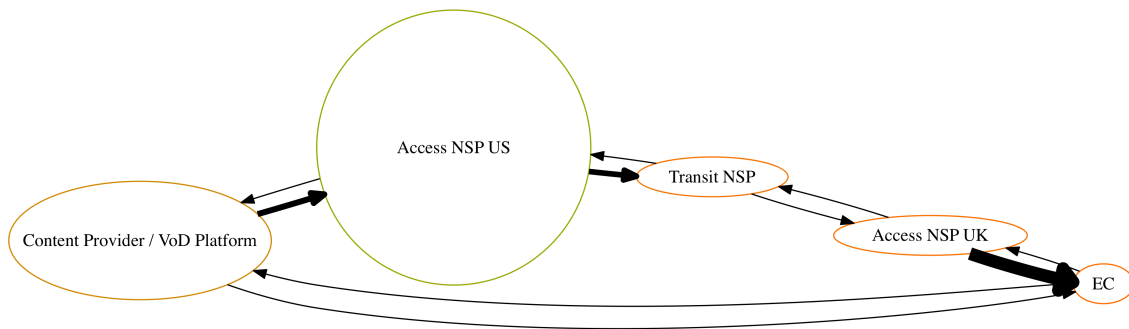


Figure 4.9: VoD VNQ w/o side payments results encoded as graph—colour: dependency; node size: entity size; edge width: utility per sale.

As the initial VoD case has only been altered with respect to the inclusion of side payments, the structural effect of side payments can be directly assessed at this point. The results for both variants are depicted in Table 4.4, and coded as graphs in Fig. 4.9 and Fig. 4.10⁶⁵ resp. Despite the common expectation that the introduction of side payments favours the NSP rather the CP, the strategic position of the CP is substantially improved (−17% for the CP; −1.4% for the Access NSP in the UK) by replacing the cascade of NSPs (ie, it takes the role of the central money distributor). This creates a structural market

⁶⁵ The graph does not directly reflect the relationship instance level. The node colour indicates the dependency result, whereas the node size shows the entity size. The edge width is the per sales utility (customer relationship; outgoing edge).

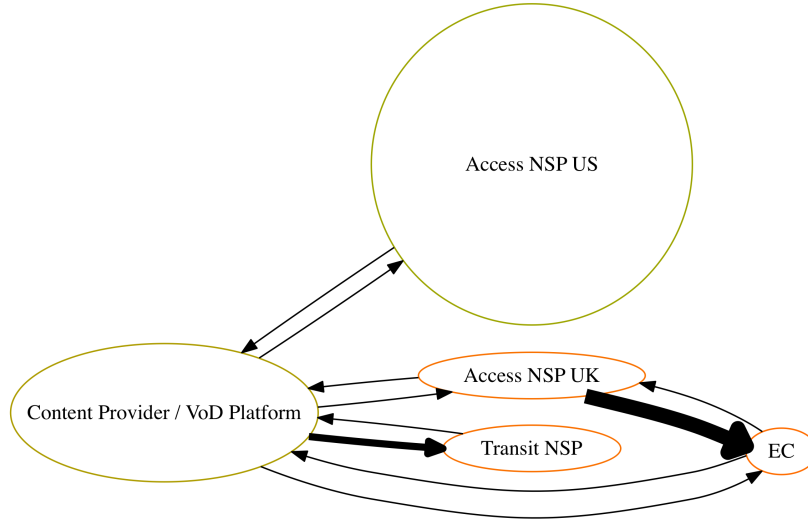


Figure 4.10: VoD VNQ results with side payments encoded as graph—colour: dependency; node size: entity size; edge width: utility per sale.

power paradox of side payments for Internet services (in analogy to the terminology used in Section 4.5.5 as coined in [ACK10]). This paradox effects is caused by the diversification of its business interactions, due to the fusion of the *transmission* and the *content* value streams (see Section 2.5.1) at the CP. The associated strategic gains are yielded at the expense of all other stakeholders, including the NSP in the USA. While in the revised VN, the Access NSP in the USA still holds the highest bargaining power, ie, it has the lowest dependency on the VN and its stakeholders, due to its enormous market volume and moderate competition levels, the CP has almost closed the gap. These entities will have the highest structural influence on future VN evolutions, eg, wrt the introduction of new revenue sharing modalities.

Conclusion #4.21. *From a structural VN point of view, a side payments paradox exists in analogy to [ACK10], which favours the sender of side payments, typically the CP, over the receiver, typically the NSP.*

4.5.5 Side Payments: Game Theory

♣⁶⁶ There are voices in industry and academy that claim that side payments may be a practicable tool for counteracting the claimed loss of revenue in the telecommunications industry. On that account, this section shifts the assessment of side payments to a game-theoretic and price-centric perspective in order to establish more clarity on the revenue potentials of introducing new charging regimes involving side payments. Such revenue po-

⁶⁶ This section primarily revisits the information presented in [Rei+13b], which is partly based on prior outcomes in [CKA10; ACK10].

tentials may further have to outweigh structural disadvantages for NSPs when introducing side payments, as illustrated in Section 4.5.4.

4.5.5.1 Effectivity.

In [CKA10] and [ACK10], the authors claim to have identified a paradox of side payments. According to their model, whenever an NSP receives additional payments from the CP or vice versa, their revenue does not increase, as intuitively expected, but under certain conditions even decreases. While this may pose severe impacts on telecommunications markets, some unclarity is created by works in literature such as [MSW09] where a highly context-specific argumentation seems to be preferred. The results are further challenged by the usage of fundamental assumptions, which might have a decisive character for the overall outcome.

4.5.5.1.1 Model. ⁶⁷ The model of [CKA10] and [ACK10] compares a neutral setting with a non-neutral counterpart. It is created in functional form as a one-shot game that considers various degrees of side payments in the non-neutral case. The model considers the price for the CP's and the NSP's service, as well as side payments p_s . Positive p_s are payments from the CP to the NSP, while negative values describe the opposite case. Both CPs and NSPs can freely choose their service price (ie, their strategy space) while p_s is set by a regulator. The demand is modelled relative to the price, and is, thus, affected by any fee charged to the customer. Thus, the best responses of players have to consider the demand effects of the overall pricing, their own rational pricing, and the rational pricing of the other player. Fundamentals equations of this model are subsequently briefly revisited.

The actual demand d is modelled around the maximum achievable demand d_{max} subject to the prices charged by both NSPs (p_i^N) and CPs (p_j^C) as follows

$$d(p_i^N, p_j^C) = d_{max} - w^N p_i^N - w^C p_j^C, \quad (4.24)$$

where w^N and w^C describe corresponding weights. We further assume that $w^N = w^C = w$ and define p_{max} to be

$$p_{max} := \frac{d_{max}}{w} \geq p_i^N + p_j^C. \quad (4.25)$$

The demand d is represented by a generic unit metric that is agnostic to product types or quality levels, and is thus directly comparable across player types. The “customer stickiness” σ , ie, a loyalty factor, for NSPs (and correspondingly for CPs with σ_i^C) is originally defined to be

$$\sigma_i^N := \frac{1/p_i^N}{\sum_{k=1}^n 1/p_k^N} \quad (4.26)$$

⁶⁷ Due to the functional form of the model in this section and its comprehensiveness, for further details we refer to the original sources and the specific instantiation in Section 4.5.5.1.4.

and subsequently integrated in the formulation of NSP N (n players) and CP C (m players) utilities

$$U_i^N := \sum_{j=1}^m \sigma_i^N \sigma_j^C d(p_i^N, p_j^C)(p_i^N + p_s) = \sigma_i^N d(p_i^N, p^C)(p_i^N + p_s), i = \{1 \dots n\} \quad (4.27)$$

and

$$U_j^C := \sum_{i=1}^n \sigma_i^N \sigma_j^C d(p_i^N, p_j^C)(p_j^C - p_s) = \sigma_j^C d(p^N, p_j^C)(p_j^C - p_s), j = \{1 \dots m\} \quad (4.28)$$

respectively.

While equations (4.27) and (4.28) appear to model individual players, harmonic means (p^N and p^C respectively) are used in the body of the original works to aggregate the players of a type. In other words, a single representative player is used per type, which is only affected by the σ function.

4.5.5.1.2 Results. Solving this model, the authors have found two internal Nash Equilibrium Points NEP_1 (side payments $p_s = 0$) and NEP_2 ($p_s \neq 0$) plus a series of boundary NEPs (NEP_B) with the following properties:

- Interior equilibria $NEP_{1/2}$ only exist under sufficiently small side payments p_s , ie, $|p_s| < 4.64\%$ of the maximum possible price p_{max} (cf. theorem in [ACK10]).
- The player receiving side payments at $NEP_{1/2}$ have a lower revenue than the sending player (based on (4.27) and (4.28) as revisited later on).
- Nevertheless, NEP_2 is an unstable saddle point equilibrium with higher NSP revenues than at NEP_B .
- Any boundary NEP NEP_B only exist whenever the price for one provider type is 0, ie, the NSP or CP cannot charge a price greater than 0, which may cause market exits in practice whenever side payments do not sufficiently compensate for the efforts of this player.

4.5.5.1.3 Critical Review. By reconstructing the theoretic results of [CKA10] and [ACK10] in a numeric experimental setting, this part elaborates on the causes for the claimed paradox. For this purpose, our results published in [Rei+13b] will be summarised hereinafter. This work begins with a critical inspection of assumptions being centrally used in the original model in order to especially highlight some existing inconsistencies.

Homogeneity: For simplifying the problem scope, all exchanged services are regarded to be identical at all times and for all users, eg, the network service QoS is held static for all customers, which does not reflect real world variations.

Competition: While competition among players of a particular type, ie, NSPs and CPs, is assumed, it is insufficiently integrated into the model. In the course of their mathematical argumentation, the authors rule out the case of market exit and concentrate on harmonic means of NSP and CP prices only. Harmonic means, however, only represent aggregate players, ie, a 1×1 game, which can neither sufficiently capture market entrance nor direct competition among players of a specific type. In particular, it is assumed that NSP and CP prices respectively are identical contradicting any competition among individual NSPs or CP (also see [Rei+13b]). This definition may be regarded as severe inconsistency in the model formulation.

Regulation: The amount of side payments needs to be exogenously set by a regulating body [ACK10; CKA10], as individual players could selfishly set side payments otherwise. We have further been able to illustrate that no naïve Nash Equilibrium (NE) can exist for the unregulated case, as both NSPs and CPs would set high side payments to their favour. This issue may only be resolved by identifying mixed NEs (without a regulatory body) or correlated equilibria (set by a regulatory body) resembling coordination games such as the classical “battle of the sexes”.

Market Power: Besides the regulation of side payments, market powers have not been considered in the model. In practice, however, the market power of players may drastically differ according to their role in the VN, eg, wrt customer ownership. Market powers may further be influenced by raised competition (due to more player on a market) or on the opposite by specialisation on niche products, eg, assured quality network services that cannot be provided by all NSPs.

Loyalty: In [ACK10] the customer loyalty or churn, ie, (un-)intentional of monetary differences, is captured by the “stickiness” factor σ . This factor non-historically describes the share of overall demand being assigned to specific NSPs or CPs respectively. As we will later describe, [ACK10] concentrates on aggregated players, and as such σ does not sufficiently reflect its important role for the overall model behaviour. Besides that, the given shape of the used σ function may be questioned, as there exists no sufficient price difference between two competitors to drive any player’s demand to 0. This behaviour may appear counterintuitive to experiences with real world marketplaces, and will be questioned in a separate analysis run later on.

Demand: Although the demand function given in equation (1) in [ACK10] and [CKA10] is formulated for an NSP i and a CP j , the interrelations between and the parallel existence of i and j are not captured in the entire model, eg, the requirement for a coexistence of both players. The demand function further appears to be inconsistent with the concept of σ_i^N and σ_j^C , which is rather used as indicators for market shares of different providers. It may further be claimed that the demand for CPs’ and NSPs’ services may hardly be described through a single demand metric, ie, “a symmetry of demand and revenue equations”. While the introduction of side payments p_s

(with $p_s > 0$) provides a valve resolving this symmetry, the means for creating this symmetry may not be “realistic in practice though” as confirmed by the view of the paper authors themselves. This concern may, for example, result from the assumed linear relationship between service types, which may not fit the different market stages of different kinds of services. A further (over-)simplification for algebraic solvability refers to the usage of an aggregate demand curve, which contributes to the inconsistent modelling of competition by this model.

Prices: While the non-zero price assumptions in [CKA10] may first appear comprehensible, we argue that in the wake of large enough side payments negative prices may represent the natural equilibrium point without any further regulation.

Despite the numerous assumptions used, which are necessary in order to achieve algebraic solvability, and a few identified inconsistencies, the discussed stream of works represents the best available approach in literature that addresses the economic feasibility of side payments, to the best of our knowledge. Thus, the revealed findings can be regarded to be highly relevant and deserve more attention in detail.

For this purpose, the model given in [CKA10] and [ACK10] will be replicated in a numeric rather than algebraic experimental setting. This setting allows not only to test 2×2 games but also to make root causes for the observed paradox explicit. As we will see later, our comprehensive analysis approach will reveal that paradoxical states indeed exist, however only under rather subtle conditions that are rather non-trivial to be specified.

4.5.5.1.4 Base Model Simulation. ♣⁶⁸ For solvability reasons, only representative players of each type were considered, ie, a 1×1 game. For the purpose of numeric simulations, the original equations in Section 4.5.5.1.1 need to be slightly revised (as presented in [Rei+13b]). The aggregate NSP utilities can, in particular, be rewritten to be

$$U^N := d(p^N, p^C)(p^N + p_s) \quad (4.29)$$

for the aggregated service demand—see [Rei+13b]. Correspondingly the aggregate demand for the CP can be expressed by

$$U^C := d(p^N, p^C)(p^C - p_s) \quad . \quad (4.30)$$

On the basis of these rewritten equations, pure Nash equilibria, ie, matching best responses, can be numerically computed for any possible price p combination (where p is normalised in $[0, 1]$)—discrete pricing steps have been used. Due to the exogenous definition of p_s this matching processes needs to be repeated for any realistic p_s level, ie, $p_s \in [0, 1]$. The effect of σ can be neglected in this 1×1 case due to the absence of competition.

⁶⁸ The initial work from [Rei+13b] is briefly revisited in this section.

In this game with only one aggregated NSP and one aggregated CP, the side payments paradox cannot be reproduced. Nonetheless, the curves largely correspond to the expected bidding behaviours of both players: the revenues are shared fraternally by both types of players, which implies that the receiving player compensates the side payments. However, when the price approaches zero, the rising side payments cannot be compensated anymore (see Figure 4.11a). At this point, the utility of the CP suddenly decreases, while the NSP first profits from higher side payments. At last, both utilities expeditiously converge to zero (see Figure 4.11b). Thus, the reception of side payments p_s is beneficial (or at least neutral) unless p_s becomes extraordinarily high.

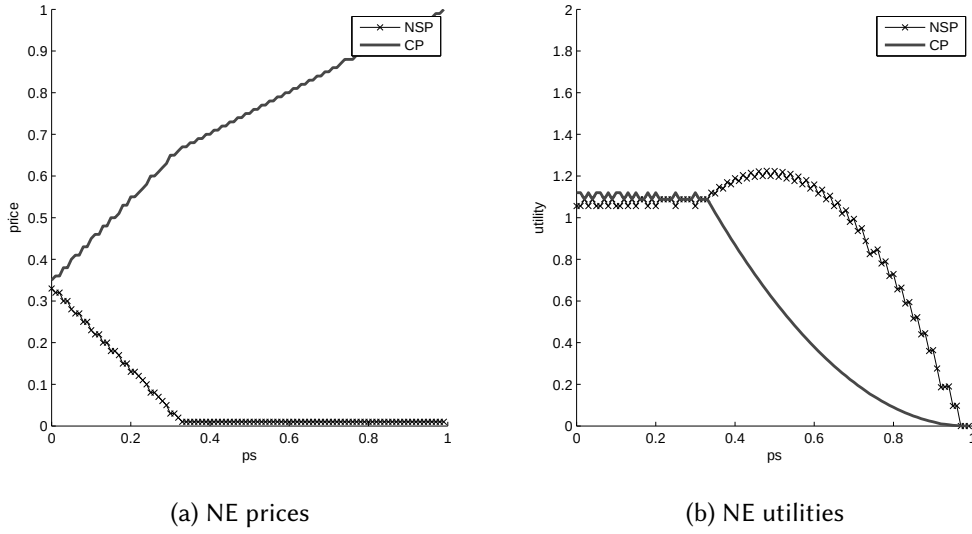


Figure 4.11: Base case: Equilibrium prices and utilities of the NSP and the CP for all side payments p_s [Rei+13b].

Despite the absence of revealing any kind of a side payments paradox, the characteristic of the results seems to be in line with those obtained in [CKA10]. To further elaborate on the side payments paradox, this game is extended to a duopoly in both the NSP's and the CP's market (ie, a 2 game). Thus, in this case, the loyalty functions σ^N and σ^C , ie, providing a customer distribution functionality among competing providers, become relevant.

W.l.o.g. we consider NSP 1 (and CP 1, respectively) to be a rational “representative” player charging p_1^N (and p_1^C) that faces a competing market follower NSP 2 (and CP 2) with p_2^N (and p_2^C). The price of the second player is set in response to the market leader's price by using a factor β , which we define by $\beta^N := p_2^N/p_1^N$ and $\beta^C := p_2^C/p_1^C$ and instantiate as follows (cf. [Rei+13b]):

- **“Low cost competitor** whose price is 25% below the price of the representative player (ie, $\beta = 0.75$)”

- “**Neutral competitor** who charges the same price as the representative player (ie, $\beta = 1.00$)”
- “**Premium price competitor** whose price is 25% above the price of the representative player (ie, $\beta = 1.25$)”

The cross-product of β^N and β^C values leads to 9 cases, which are depicted in Fig. 4.12). The case of $\beta^N = 1.00 / \beta^C = 1.00$ can be considered to be identical with the 1×1 game results (equal distribution of demand; σ is irrelevant; the sum of revenues are equal).

A side payments paradox occurs whenever $\beta^C = 0.75$ and exactly at $\beta^N = 1.25/\beta^C = 1$, ie, at the equilibrium the player receiving side payments has a lower revenue than the sender. On the other hand, Fig. 4.12 illustrates that the paradox vanishes at very high p_s . The magnitude of the side payment paradox increases with the growing discrepancy of the competitive situations between provider types especially when having premium players in the NSP and discounters CP market, ie, highest paradox is observed at $\beta^N = 1.25 / \beta^C = 0.75$.

For more extrapolations of the side payments paradox, we kindly refer to [Rei+13b] and our summary in Table 4.5; the four identified paradoxical cases are highlighted, while only a single case is considered neutral, ie, $\beta^N = \beta^C = 1$.

Table 4.5: Summary of results.

β^N / β^C	0.75	1.00	1.25
0.75	paradox (max. ≈ 0.2)	NSP gains	NSP gains heavily
1.00	paradox (max. ≈ 0.3)	neutral/base case	NSP gains heavily
1.25	paradox (max. ≈ 0.4)	paradox (max. ≈ 0.12)	NSP gains

As argued in [Rei+13b], the σ function causes any observed side payments paradox, for which subtle conditions have to hold to come into existence. Thus, aggregate 1×1 games cannot obtain any paradoxes in our approach.

4.5.5.1.5 Extended Model Simulation. Numerous works in literature (cf. the overview given in [HKH05]) have identified a positive relation between loyalty and CS, as well as an S-curve-shaped relationship between CS and WTP (insensitivity to small CS changes). While our simulation has not found any direct indications supporting the claims of an S-curve-shaped relationship, no direct empirical parameterisation for the the relationship between loyalty and WTP is available. Thus, only the assumptions used in the base model can be critically assessed in the light of these general indications.

In [CKA10], the factors σ_i^N and σ_j^C describe the distribution of the end customers among the various NSPs and CPs. They satisfy two conditions:

- C1. If all providers charge the same price, the distribution is assumed to be uniform;

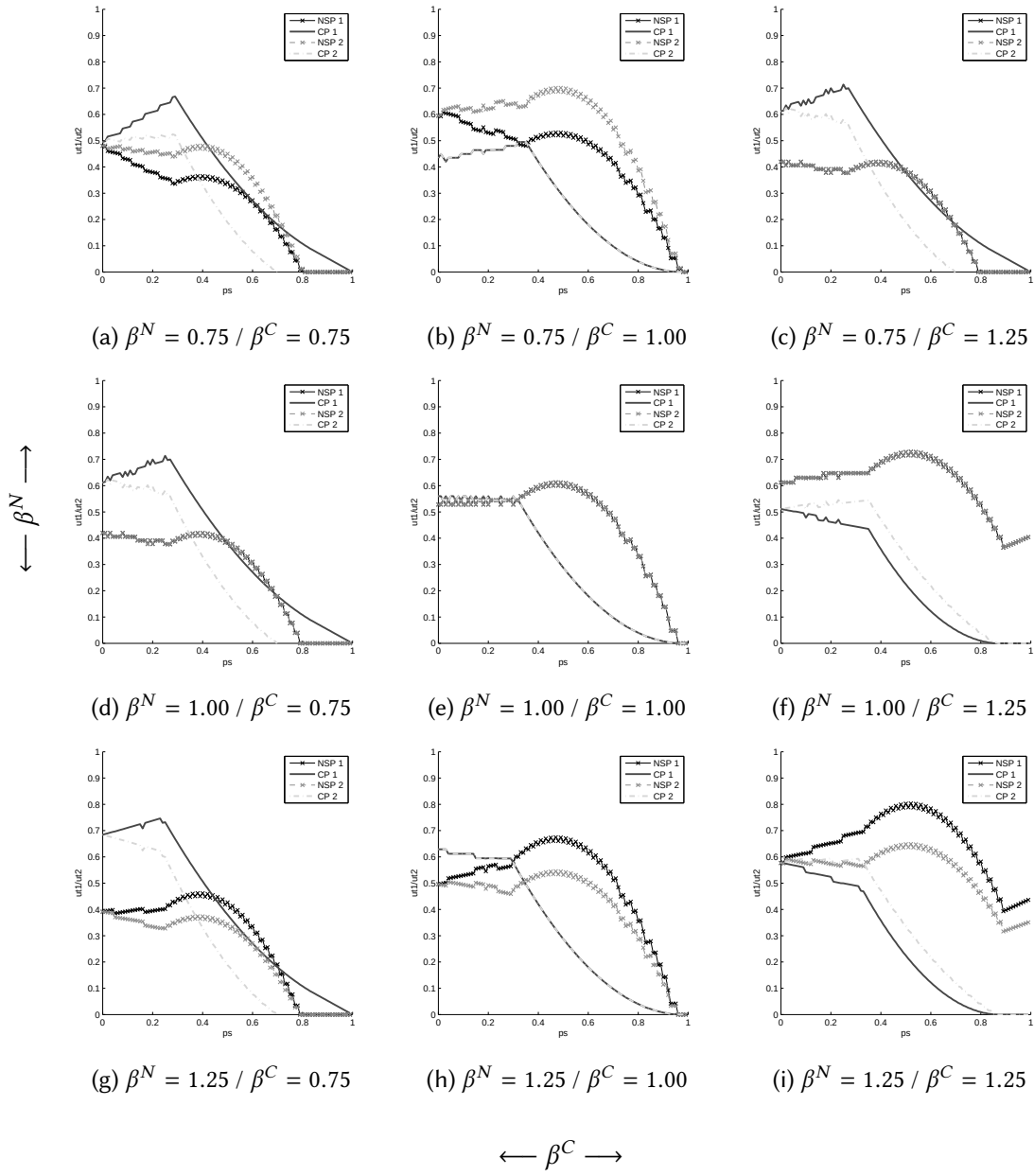


Figure 4.12: Utilities for nine competitive cases in the 2×2 game.

C2. A more expensive provider has less customers than a cheaper one.

However, the assumption-prone loyalty function σ in its current form quickly flattens with increasing price differences. In other words, it is almost impossible to shift the entire demand to a single provider if all others are very expensive. Thus, we will design a new and more realistic σ' that will satisfy a third condition:

C3. With a sufficiently high price difference, the entire demand shifts to cheaper competitors.

Thus, σ' is supposed to influence demands continuously around a neutral inflection point (insensitivity to small price changes). Such a σ' that meets all three conditions will, hereinafter, be formulated.

Let us assume there are two competing players (either NSPs with superscript N or CPs with C) in a market with $p_1, p_2 \geq 0$ with $p_2 \leq 2p_1$, resulting in a distribution of $\sigma_2 \geq \sigma_1$ with $\sigma_1 + \sigma_2 = 1$. Suppose p_2 is altered while p_1 remains static (or vice versa), then it can be assumed that the resulting change of σ_2 , ie, $\Delta\sigma'_2$, is directly proportional to the difference between p_1 and p_2 —thus, meeting the above requirements and the expectations given in [HKH05] the demand will shift between player 1 and 2. Unless $p_1 = p_2 = 0$ we can formulate $\Delta\sigma'_2$ by

$$\Delta\sigma'_2 = -\alpha \cdot |p_2 - p_1| \Delta p_2 \quad , \quad (4.31)$$

where $\alpha > 0$ as constant of proportionality. The directionality of σ' and, thus, d shifts are linked to the direction of price changes, ie, whether the price has been increased or decreased. This is reflected in the following differential equation for σ'_2 :

$$\frac{d\sigma'_2}{dp_2} = -\alpha \cdot |p_2 - p_1| \quad . \quad (4.32)$$

When minding condition C1. the two cases

$$p_2 \geq p_1 \Rightarrow \sigma'_2 = \alpha(p_1 p_2 - \frac{1}{2} p_2^2) + \frac{1}{2} - \frac{\alpha}{2} p_1^2 \quad \text{and} \quad (4.33)$$

$$p_2 \leq p_1 \Rightarrow \sigma'_2 = -\alpha(p_1 p_2 - \frac{1}{2} p_2^2) + \frac{1}{2} + \frac{\alpha}{2} p_1^2 \quad (4.34)$$

have to be distinguished.

As a result of conditions C1. and C3.⁶⁹, the demands are equal at the inflection point (ie, $p_1 = p_2$ yields $\sigma'_2 = 0.5$), while they entirely shift to the second provider when using the maximum price $p_{max} = p_1 = 1 \Rightarrow \sigma'_1 = 0$ and $\sigma'_2 = 1$. When considering the boundary condition $p_2 = 0 \Rightarrow \sigma'_2 = 1$ (ie, if provider 2 is offering the service for free, it attracts all customers), the constant of proportionality can be derived to be $\alpha = 1/p_1^2$. When defining $\beta^{N/C}$ to be

⁶⁹ Apparently condition C2. is satisfied too

$$\beta^{N/C} := \frac{p_2^{N/C}}{p_1^{N/C}} \quad (4.35)$$

σ'_1 ($\sigma'_1 = 1 - \sigma'_2$) and σ_2 take, hence, the following form for the case of NSPs:

$$\beta^N \geq 1 \Rightarrow \sigma_2'^N = \beta^N - \frac{1}{2}\beta^{N,2} \text{ and } \sigma_1'^N = 1 - (\beta^N - \frac{1}{2}\beta^{N,2}) \quad ; \quad (4.36)$$

$$\beta^N \leq 1 \Rightarrow \sigma_2'^N = 1 - (\beta^N - \frac{1}{2}\beta^{N,2}) \text{ and } \sigma_1'^N = \beta^N - \frac{1}{2}\beta^{N,2} \quad . \quad (4.37)$$

The case of CP works in analogy by replacing β^N with β^C .

For the case where $p_1 = p_2 = 0$, obviously $\sigma'_1 = \sigma'_2 = \frac{1}{2}$ can be defined (in response to C1.). In all cases, this new distribution function evidently satisfies C2. and can, thus, fully replace the initial σ formulation. By swapping p_1 with p_2 , the roles of providers can at any time be exchanged, ie, σ'_1 and σ'_2 are swapped alike.

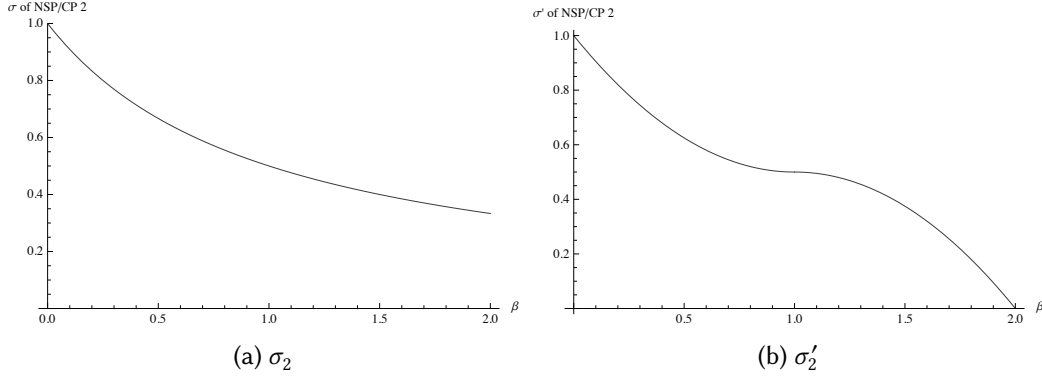


Figure 4.13: A comparison of σ_2 and the revised σ'_2 in β .

By replacing σ with σ' , the loyalty relationship between price and market share is altered (cf. Fig. 4.13b). The new loyalty relationship is stable around a neutral point when both prices are identical ($\beta = 1.00$). Unlike σ (cf. Fig. 4.13a) the new σ' non-linearly diverges with any alteration and converges to the minimum (at 0.0) or the maximum demand (at 1.0) with a large enough price difference, ie, $\beta \geq 2$ or $\beta \leq 0$ respectively.

In a subsequent simulation run, σ will be replaced by σ' in an otherwise identical model and test setup. While the new results (cf. Fig. 4.14) seem to resemble the results given in Fig. 4.12 at the first glance, some noteworthy differences exist: The new σ' eliminates the symmetrical emergence of the paradox of side payments for all players of the same type, eg, all CPs. In some cases, the “representative player” (NSP_1 or CP_1) can substantially profit from a paradox and any competitive advantage while the same does not hold for the second player. Apart from that, the extent of paradoxes is changed where especially the case of $\beta^N = 1.25/\beta^C = 1.00$ is affected (the paradox is rendered insignificant in this case).

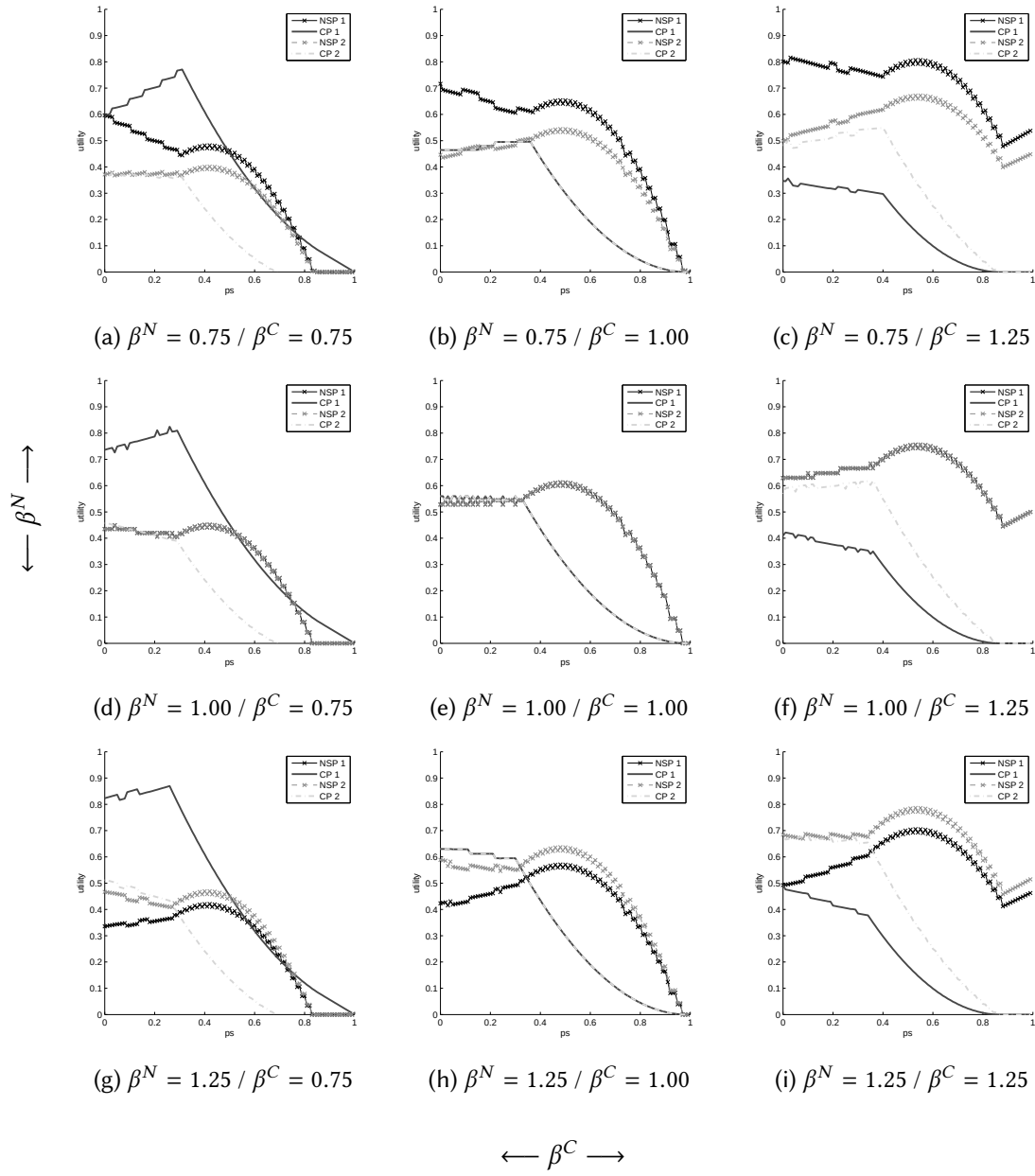


Figure 4.14: The nine cases for the alternative version of σ'^N and σ'^C .

In [Rei+13b] (see partitions below 0 in Fig. 3 and 4), we also conclude that moderate side payments may lead to paradoxical states. When side payments become large enough, they cannot be compensated anymore, which may limit the possibility of a paradox and may lower the utility levels of all players. Contrary to the theoretic results in [CKA10] (Theorem 1) and [ACK10], the side payment threshold of 4.64% of p_{max} has, however, not played any important role in our results. In the initial claim, an equilibrium has only been identified below the 4.64% threshold and has yielded a paradox. This threshold cannot be verified in our simulations neither for the initial σ , the revised σ' , whether zero prices are allowed nor when altering the competition scenarios. The differences may be explained by the specific assumptions used in the original works. From Fig. 4.12 and Fig. 4.14 we can also infer that the revision of loyalty factor σ' lowers the risk of paradoxes for all player when increasing the level of side payments.

Hence, from a price- and demand-centric perspective the economic gains of introducing side payments remain impalpable. Due to the compensation of side payments with access prices and the resulting low revenue prospects, it can, however, be concluded that large-scale deployments are unrealistic. In combination with the social hurdles associated with the introduction of side payments, the economic gains are too obscure without further evidence on economic gains. Strategies to avoid side payments have briefly been discussed in [FP712a].

Conclusion #4.22. *The economic feasibility of applying side payments for network services seems questionable from a game-theoretic point of view. While paradox effects cannot be confirmed in this context (where receivers of side payments would obtain less revenue and less profit in total), the model clearly points to the absence of relevant revenue or profit gain potentials when considering a large-scale market introduction of side payments. Niche cases, like applying side payments as surcharges for specific contents or players, may still be economically attractive, as the used model does not capture such cases.*

4.6 Concluding Remarks

In this chapter, a series of mechanisms have been assessed for the successful supply-side configuration of a market for quality-differentiated network services. A review of current NN regulations, fairness, and a test of associated market instruments does not render a very prosperous perspective for deploying side payments-based charging mechanisms (both from structural and revenue perspective). According to many definitions, side payments violate content or application neutrality, and in consequence NN. While integral factors characterising side payments are still unknown, this chapter has illustrated that even a paradox can occur under very specific circumstances, ie, receivers of side payments end up with lower revenues than without them. For these reasons, we currently recommend the elaboration of other instruments following more neutral market configurations for the sales of QoE first. Nonetheless, the partial deployment of side payments, eg, for a handful of services rather than the entire Internet, might conclude differently. Apart

from that, evidence has been found for the inapplicability of pure quantity-based pricing mechanisms that aim at fostering fairness among users. The concept of another kind of quality differentiation and quality guarantee seems to provide better working points for targeting fairness in communications networks.

For quality-differentiated single-class network services, this chapter has revealed that the profitability shifts towards lower quality levels whenever the demand is very high, and the capacity is small. This effect is caused by strong demand in lower quality classes and the good cost-price relationship in the customer segment. In our empirically-backed simulation, low and intermediate quality levels perform best in almost any case, which corresponds to the logic of the mass market of today's Internet. In a corresponding multi-class network, higher quality classes are economically more attractive. While in the single-class system, high-quality offers suffer from the low resource efficiency and moderate demand levels, the multi-class variant will complete the bundle of quality classes that maximise the revenue per available resource. This configuration can include "premium" offers that had to be ignored before. This chapter has further highlighted that higher prices yield higher or equal profits in monopolistic markets, which may provide a key information when approaching a new market.

This chapter has also investigated the economic appeal and nature of offloading cellular traffic to the shared Wi-Fi spectrum. Based on an applied game-theoretic assessment, it can be concluded that offloading traffic is a dominant strategy whenever the customer yields no disutility. Thus, for bulk traffic an aggressive offloading strategy may be applied, while providers may protect the quality in the cellular spectrum. Due to the absence of hierarchically organised coordination functions and the lack of exclusive spectrum access, the QoS of unlicensed spectrum solutions such as Wi-Fi is not protected by rationally acting NSPs. With this strategy, freed resources in the cellular spectrum can be resold or premium services can be provisioned in the cellular network, which might raise the revenue levels. This chapter has further illustrated that the joint operation of cellular networks and fixed line networks (with unlicensed spectrum Wi-Fi solutions) is strategically favourable for NSPs.

QoE & WTP TECHNOLOGIES

This part of the thesis transfers the outcomes of Chapter 3 (empirical parameterisation) and Chapter 4 (optimisation results) to a technical representation. Therefore, means for targeting the following primary objective are presented:

Objective: Presentation of mechanisms allowing the QoE-, WTP- and utility-aware aware revenue-, profit- and/or social welfare-optimal distribution of network resources along the entire end-to-end transmission path, which may potentially span multiple Domains, ie, inter-carrier solutions

Specific attention will be dedicated to the case of controlled degradations.

Intelligent controlled degradation is a mechanism of proactively coping with resource scarcity by adapting the service to keep the experience or utility optimal under the given resource constraints.

This objective can be subdivided into the following five contribution blocks for the controlled degradation case:

- Inter-carrier / Inter-Domain frameworks providing SLA-based end-to-end network quality assurance and means for selling network quality as global service (cf. Section 5.2)
- NSP and user utility functions when applying controlled degradations (cf. Section 5.3)
- Computation algorithms for the corresponding end-to-end transmission paths using the later on introduced technique of Media Degradation Path (MDP) and following defined objectives (cf. Section 5.3.3 based upon Section 5.1.8.1)

- MDP-based access control and resource assignment techniques (especially relevant for access networks) aiming to maximise the defined objective based on utility optimisations (cf. Section 5.3.4 based upon Section 5.1.8.2)
- Generic WTP-aware network schedulers (cf. Section 5.4)

After starting with a review of relevant background literature in Section 5.1, the solutions for the individual contribution blocks are introduced.

5.1 Background

This section reviews the potential of known technologies to be used for quality differentiation especially in the IC (ie, inter-Domain) context. IP networks will serve as fundamental basis, which on their own provide QoS information in their IP packet headers: In IP version 4 (IPv4) [RFC 791], an 8-bit *Type of Service* field for QoS classification plus a *Differentiated Service CodePoint (DSCP)* [RFC 2474] flag (the *Precedence* flag [RFC 791] was used before) for priority marking of traffic classes exists. In IP version 6 (IPv6) [RFC 2460], DSCP is retained, and additional QoS information is foreseen in the extension headers.

5.1.1 Integrated Services Architecture (Intserv)

Integrated Services Architecture (Intserv) [RFC 2210] is a technology to enable QoS-differentiation in the Domains of deployment. Intserv is based on per-session resource reservations, which are stored and handled by all involved routers. The reservations are done by using the Resource Reservation Protocol (RSVP) [RFC 2205], which is placed on network layer 4 of the ISO/OSI model (transport layer; cf. ISO/IEC 7498-1 [IEC7498]). The admission control handles the assessment whether a reservation request can be met based on the current load levels along the path and the guarantees that have previously been given. Intserv supports three modes: BE (as without Intserv or comparable technologies), *controlled load* and *guarantees*. The mode with controlled load refers to the guarantee that the majority of classified packets will not be dropped within the Intserv-Domain and the waiting time in queues will be very limited. The guarantees mode gives stricter assurances for the reservations, ie, typically guarantees on the promised minimum throughput and the maximum latency are given.

According to [Bos07], Intserv is functional in practice, but insufficiently scalable due to the session-aware states in Intserv routers. With the increase of sessions, these states will have to keep track of and maintain “millions of reservations”. In other words, the session-awareness seems to limit the practical applicability of Intserv.

5.1.2 Differentiated Services Architecture (Diffserv)

Due to the high complexity of Intserv, other mechanisms have been researched such as Differentiated Services Architecture (Diffserv) [RFC 2475]. Diffserv is intended to in-

crease the scalability by providing a service quality that is somehow better than BE in the Diffserv-enabled Domain. It replaces the strict resource reservations of Intserv for individual sessions by class-based discriminatory forwarding schemes (ie, priorities). Hence, no RSVP signalling [RFC 2205] is required anymore. In this model, packets are classified and marked at the network edge (representing a complex task) and forwarded (ie, Per Hop Behavior (PHB)), based on assigned classes in the core (limited effort). Two PHB schemes exist: Expedited Forwarding (EF) and Assured Forwarding (AF). EF is intended to enable low latency services. AF in contrast differentiates in 4 quality classes with three priority levels plus BE.

As a consequence to the transition from Intserv to Diffserv quality guarantees cannot be provided anymore. Therefore, whenever congestion occurs, an overload cannot necessarily be avoided. Especially low priority packets may be dropped [Bos07]. This shortfall has been targeted by Resource Management in DiffServ (RMD) that adds resource reservations to Diffserv without storing session information for interior routers overtaking the PHB functions (see [Bos07]). Thus, the scalability is supposed to be improved over the Intserv approach. Nevertheless, the complexity of performing the access control tasks in the edge of the network will rule out Internet-scale deployments of RMD despite those scalability improvements. Diffserv and RMD can still be used to provide network quality differentiation in individual Domains, which need to be connected by other technical means—eg, see Section 5.2.

Depending on the chosen “service disciplines” Diffserv may or may not be work-conserving [RFC 3290]. Contrary to stricter priorities, the usage of Weighted Fair Queuing (WFQ) for example would assure a work conservation.

5.1.3 Border Gateway Protocol (BGP)

Border Gateway Protocol (BGP) is a routing protocol that is required inherently for the proper functioning of today’s Internet. Positioned on network layer 3 it is typically referred to as Exterior Gateway Protocol (EGP). EGPs are used for exchanging connectivity information between AS and as such are the primary choice for the IC case. Despite its primary functioning as EGP, which is cardinal for today’s Internet, BGP may, of course, propagate routing information throughout the nodes of each AS. Apart from that, specifically designed Interior Gateway Protocols (IGPs) such as Open Shortest Path First (OSPF) exist that are typically only used within a single AS.

Apart from rate limitations, QoS and especially QoE differentiation is often seen to be linked to path diversity (ie, multi-path routing). A short path with low capacity may be suited for highly interactive traffic (due to a delay minimisation), while a longer path with more capacity may lead to higher end user QoE for download use cases. In other cases, the capacity of the shortest path may be insufficient to carry the entire traffic between two destination-pairs, ie, only “premium” traffic may be routed on the shortest path, while BE traffic may have to take a detour. Today, BGP does not support route diversity, as it only advertises a single neighbour for forwarding to particular ASes. This shortfall exacerbates

its usage as sole QoE and QoS differentiation technique.

BGP is a(n) (advanced) distance vector mechanisms, which only directly considers the distance to a destination for a definitive routing decision¹ and is designed to handle incomplete information. An extension to support such a functionality appears to the best of our knowledge unlikely², due to the scalability issues that might be inferred when further expanding forwarding tables to support path diversity. Thus, enabling the direct control over the provisioned QoS is difficult with BGP. The commercialisation is hampered further by the absence of attached market frameworks.

Within AS, alternative mechanisms using a link state approach exist. Link state routing requires a complete topology information and is, thus, hardly applicable to the IC context, especially when adding the complexity of path diversity.

5.1.4 IP exchange (IPX)

The GSM Association (GSMA) has standardised IP exchange (IPX) [IR.34] for exchanging the IP traffic of users across Domains in a defined manner. Originating from the idea of telephone networks IPX constructs a private network aside the classical Internet, which provides a technical and economic framework for exchanging data traffic across Domain borders. With this concept, multiple end-to-end QoS classes can be supported. For this purpose, IPX uses a cascade of payments, typically based on SLAs, to compensate for the efforts of external NSPs on the end-to-end transmission chain.

In the last years, IPX has successfully supported the convergence towards all-IP systems in telecommunications, which is further supported by approaches such as IMS. Today, many Internet Service Providers (ISPs) are also IPX providers and are directly be connected to IPX exchange points (comparable to public peering points).

IPX defines three kinds of IC models:

1. **IPX Transport** (bilateral): Bilateral transport layer connection (see ISO/IEC 7498-1 [IEC7498]) comparable to private peering on lower network layers;
2. **IPX Service Transit** (cascaded IC): Service-aware IC including cascaded charging functionalities;
3. **IPX Service Hub** (multilateral): Service-aware multilateral IC service in analogy with services of Transit NSPs.

The variants 2 and 3 may involve IPX proxies, which provide additional functionalities such as accounting, route identification or media transcoding. Today, especially the multilateral hub model seems to be of highest practical relevance (ie, a series of IPX exchange hubs exists, eg, in Singapore).

¹ Indirectly costs or different kinds of quality may be encoded in the distances being propagated to other nodes. BGP supports local optimisations such as local routing preferences.

² There have been some standardisation efforts for both the case within a single AS (IBGP) [Utt+15] and across ASes [Wal+15], but those have received a limited echo within academia and industry.

All in all, IPX mainly resembles the peering and peering point principle on the level of more sensitive traffic, eg, voice. IPX, however, alone appears to target more dynamic end-to-end quality differentiation cases insufficiently (including access network solutions).

5.1.5 IP Multimedia Subsystem (IMS)

IMS is an IP-based standard by the Third Generation Partnership Project (3GPP) group [TS 23V12], which has initially targeted cellular IP services. Especially widely known is IMS's Session Initiation Protocol (SIP) protocol, which is commonly used for VoIP calls. The entire IMS standard is, however, very comprehensible and complex, as it aims to both link classical telecommunications network to IP data networks and to provide added services for the IP-only case, eg, QoS policies and charging.

The (QoS) policy and charging mechanisms are defined in the Policy and Charging Control (PCC) architecture, which consist of the Policy and Charging Enforcement Function (PCEF) and the Proxy Call Session Control Function (P-CSCF) with the Policy and Charging Rules Function (PCRF). The PCC architecture allows IMS to charge for, control and assure the QoS for end user sessions. IMS has recently received a new impetus with the emergence of LTE networks. LTE is a low latency cellular all-IP network that requires voice calls to be redesigned as data traffic. For this purpose, IMS is intended to be used as a cardinal element of the LTE's voice call standard called VoLTE [IR.92]. In the context of this thesis, this may be useful for application-aware QoS differentiation and associated charging mechanisms.

The PCRF acts as Policy Decision Point (PDP), which can activate or deactivate pre-defined PCC rules for authenticated users³ and their services. The PCRF obtains the profiles of subscribers (the allowed usage perimeters and other policy-relevant subscriber information) from the Subscription Profile Repository (SPR) and the authentication information from the Home Subscriber Server (HSS). In contrast, the PCEF (generically denoted by Policy Enforcement Point (PEP)) is responsible for enforcing the PCC rules at the gateways, ie, the PCEF enforces QoS policies for each packet and operates associated charging functions.

In analogy with IMS, Telecoms & Internet converged Services & Protocols for Advanced Network (TISPAN) (a standard [ES 282] of the European Telecommunications Standards Institute (ETSI)) has emerged for the fixed line case, but has received only limited attention in industry and academy in the last years. The lack of interest may be explained by the integration of fixed line network support in IMS in 2008 [TS 23V7]⁴, which has rendered IMS more attractive. Despite these advancements IMS is still not able to assure QoS beyond Domain borders—ie, IMS does not support truly end-to-end IC QoS⁵

³ Authentication is done in the P-CSCF.

⁴ The standard has been updated and maintained in version 7 until 2009.

⁵ In [TS 23V12] a rough sketch of IC QoS differentiation ideas is given, but dedicate solution and their relationship to IMS QoS policing are still missing.

Despite, the high relevance of IMS, its high complexity and rather static configuration, has limited the IMS usage to sensitive operator services in access networks. IMS will likely play an important role in the technological convergence for cellular networks, eg, IP-only LTE networks realised with VoLTE and comparable measures. The application for inter-Domain usages linking several ASes, is not yet envisaged.

5.1.6 Connection-oriented Technologies

Today's Internet is dominated by connectionless⁶ IP traffic and associated technologies. Nevertheless, there exist other technologies transferring the principles of telephony networks to the case of IP data traffic, which are briefly touched in this section for completeness purposes⁷. Connection-oriented network technologies try to increase the degree of controllability of data traffic by enforcing the traffic to follow on predefined paths⁸.

Most notable technology in this area is Generalized Multi-Protocol Label Switching (GMPLS) [RFC 3945] as the functional successor of Multi-Protocol Label Switching (MPLS) [RFC 3031]. In GMPLS each packet follows a Label Switched Paths (LSPs) with strictly defined path and resource reservation (eg, using Resource Reservation Protocol – Traffic Engineering (RSVP-TE) [RFC 3209]), which is determined based on labels being attached to packets at the Label Edge Router (LER). Due to the reduced address size of GMPLS labels in comparison to IP addresses (especially for IPv6), GMPLS labels may be used alternatively to rewrite packet headers at edge routers while otherwise following the connectionless principle (as for example discussed in the later on described ETICS project).

GMPLS may also serve as a basis for realising network layer 2 or layer 3 Virtual Private Networks (VPNs). The layer 3 variant may according to the recommendations of [RFC 4364] be extended to provide IC VPNs. More details on VPNs using GMPLS are described in [FP712b].

One of the main downsides of GMPLS, when being used for a connection-oriented traffic forwarding scheme, is the complexity of calculating, maintaining and managing a wide range of precomputed paths⁹. Thus, true end-to-end network services with improved QoS may only be sold as a niche service. This view is further strengthened by the low efficiency of dedicating resources within physical pipes to predefined GMPLS paths. In other words, GMPLS does not support work conservation (in analogy with the terminology used for scheduling disciplines), which may not be aligned to the cost pressure in today's Internet.

⁶ Connectionless, in this context, refers to technologies on network layer 3 and below. On layer 4, connection-oriented transport solutions such as TCP are not affected by this definition.

⁷ Due to the increasing convergence to low cost all-IP networks, the principles of old telephony networks have increasingly become unpopular in the last years.

⁸ There may exist precomputed primary and backup paths potentially specific to traffic classes.

⁹ For interconnecting a set of n source and m destination end points (where each quality class counts as individual end point) a cross-product of precomputed paths is necessary.

5.1.7 Energy- and Cost-aware Routing

Cuzzocrea *et al.* [Cuz+12] have proposed a topology control scheme (moderating QoS satisfaction as well as overall network vitality) with a noteworthy potential for reducing the energy consumption and, hence, costs of networks, while still minding QoS requirements. This scheme is based on the betweenness centrality measure, as known from Social Network Analysis (SNA). By encoding energy levels as edge weights in the graph, their algorithm selects the energy-wise best logical neighbouring edges for each node in a graph.

In comparison to alternative algorithms such as the Gabriel Graph (GG) [GS69], the Relative Neighborhood Graph (RNG) [Tou80], and the Closeness Centrality (CC) [Fre79], Cuzzocrea *et al.* [Cuz+12] claim superiority “in terms of logical neighbors found, energy consumption, latency and hit-ratio”, supported by a series of simulations. Additionally, their algorithm may also be paired with effective load balancing concepts such as proposed in [PD10]. Such algorithms provide an orthogonal optimisation cost and energy-centric optimisation to the mainly revenue-driven considerations presented in this thesis.

5.1.8 QoE-aware Routing, Access Control and Resource Assignment

Aligned to SIP-based communications, Media Degradation Path (MDP) [Sko07] allows for the request-based optimal (intra-Domain) end-to-end network resource assignment. MDP matches the constraints of a single user, ie, an Extensible Markup Language (XML) file describing budget, capability, etc. to known service profiles in order to optimise the assignment and allocation of resources. The result of MDP is a data set consisting of an order list of configurations with decreasing utility for the users, ie, a list consisting of the optimal configuration and ordered alternatives for controlled degradation purposes.

Building on the work of [Sko07] (and follow-up works such as [ISM14; Dob+14]), the configuration of the controlled degradation process consists of two main phases:

1. The end-to-end flow path, which optimises the QoE for end users minding service profiles specified in the MDP configurations, needs to be computed [Dob+14].
2. Having calculated an MDP, an access control mechanism targets the utility optimal admission of service requests by utilising a controlled quality degradation technique whenever resource scarcity does not permit an unrestricted admission. Later alterations of service requirements (eg, requests for more bandwidth) are handled in a comparable manner by a resource assignment module that intervenes whenever the resource assignment of a session needs reconsideration [ISM14], ie, those sessions that requested more bandwidth could be degraded to less resource demanding configurations, if the request cannot be entirely satisfied.

For both 1 and 2 two specific kinds of simulation solutions are briefly reviewed hereinafter, ie, Q-POINT, and the ADmission control and resource Allocation for adaPtive mulTImedia SErvices (ADAPTISE) and LTE simulators.

5.1.8.1 QoE-driven Path Optimization Model for Multimedia Services (Q-POINT)

Q-POINT [Dob+14] represents an end-to-end flow path optimisation model centring its attention on the maximisation of QoE for the entire population of users that establish multimedia services over a Software Defined Networking (SDN) infrastructure. While MDP treats the network as a black box, Q-POINT tries to utilise the information on the network (devices, nodes, links, ...) in order to operationalise MDP configurations optimally, ie, Q-POINT takes the MDP configuration file as input for the creation of optimal end-to-end flow paths.

For this purpose the Q-POINT model utilises QoE curves for QoS metrics like packet loss and delay as input for a complex linear optimisation program (executed with *IBM CPLEX*¹⁰). It makes further use of technical constraints like the media codec and bitrate, the network topology, the available capacity on links, the queuing in the nodes and the capacity of a link in order to allow a realistic assessment of the network. The result of running the linear optimisation program is the Q-POINT solution—ie, a resulting equation allowing for the objective-optimal assignment of best end-to-end flow paths in a multimedia session. For simplicity reasons, the rare case of packet losses on the links, timely varying QoS demands (eg, variable bitrates leading to variable bandwidth demands), and inter-Domain issues have not been considered at this research stage. Apart from that, user's service preferences are currently not backed by empirical evidence.

For its proper functioning, Q-POINT is meant to be placed in a central entity, which might be interfacing to IMS/SIP systems in order to receive certain kind of service requests and transfer them to optimal configurations for SDNs. The interfaces between the SDN module and the network nodes themselves handle the technical realisation of the Q-POINT results, which is currently being prototypically implemented by the dedicated team at UNIZG-FER.

5.1.8.2 Admission control and resource Allocation for adaptive multimedia Services (ADAPTISE) and LTE Simulator

In [ISM13; ISM14], an access control and resource assignment-based approach is presented, which has been realised using the ADAPTISE simulator plus an LTE simulator [Pir+11]. In the first stage, ADAPTISE executes utility-optimisations based on various arrival rate and service duration patterns. The optimisation algorithm is run whenever a new resource request is added or a resource shortage occurs due to increased demands of existing sessions. In the second stage, the LTE simulator provides information on LTE-specific effects when provisioning the network resources in practice.

For the purpose of utility-optimality, this approach targets the usage of the entire capacity that is available. Thus, whenever a request can be satisfied in optimal service quality, the access will be granted without any alterations. However, in the case of resource

¹⁰ “BM ILOG CPLEX Optimization Studio”: <http://www-03.ibm.com/software/products/en/ibmilogcpleoptistud>, last accessed: 20th May 2016

scarcity, an MDP-based controlled degradation takes place whereby maximising the utility of the affected ongoing sessions.

Comparing this approach with a naïve Admission Control approach, which drops requests whenever they cannot be satisfied entirely, [ISM14] is able to assign more sessions and obtain a higher aggregate utility than with more simplistic approaches. The utility is assessed based on assumptions for utility curves and individual service preference weights by users, ie, their relative preference of one service over another one. The absence of complete empirical parameterisation may hamper the interpretability of the discussed utility gains.

5.2 Inter-Domain Network Quality

•¹¹ The FP7 ETICS project¹², funded by the European Commission (EC) between 2010 and 2012, has investigated technological and economic barriers when introducing and marketing IC QoS. The project consisting of a series of research institutions and major European telecommunications providers has proposed an architecture [FP712b] for realising guaranteed “premium” quality offers across Domain borders by using specifically designed SLAs. This concept not only considers the heterogeneity of involved ASes, but also provides means for controlling payment flows along the transmission chain. For the latter purpose, the project consortium has envisioned mechanisms such as Sending Party Pays (SPP) regimes [Bor+11] that are supposed to fairer compensate NSPs for the (additional) efforts. Despite some dedicated studies within ETICS, QoE is not directly supported in its architecture. ETICS has further aimed at creating a more dynamic IC resource trading sphere by replacing the rather static peering and transit agreements by better controllable and more direct solutions. Those new solutions may provide additional features like route control, path diversity, QoS differentiation, hierarchy-free multi-hop transit services, support for innovative charging concepts, and graceful quality and access control.

ETICS may thus act in different ways in order to fill the void in the IC QoS technology sphere (as summarised from [FP712b]):

- It can form an IC complement to intra-Domain solutions such as IMS, ie, ETICS may provide the ‘glue’ between various IMS islands by attaching IMS end user session to aggregate IC quality classes (see Fig. 5.1).
- Unified quality classes may be taken from IPX to parameterise the more flexible SLA-based IC QoS scheme of ETICS.
- It can interconnect multiple IPX backbone networks, which may each consist of several IPX NSPs (including quality guarantees aligned to IPX quality classes).

¹¹ *The authors would like to disclose their involvement with and funding by the FP7 ETICS project during the years 2010 and 2012.*

¹² <http://https://www.ict-etics.eu/>, last accessed: 20th May 2016

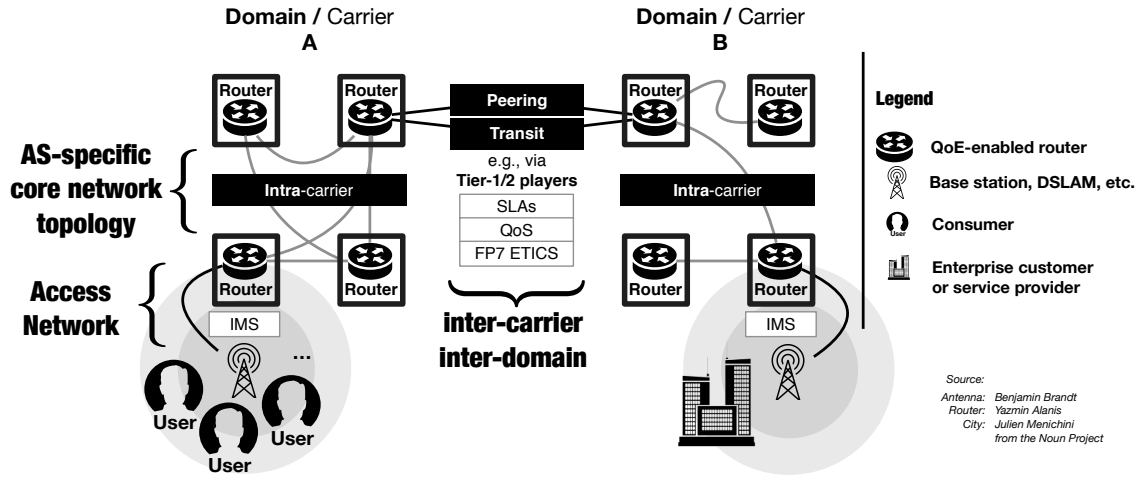


Figure 5.1: Formation of a global inter-carrier Internet network

- Smaller NSPs, which are not yet IPX providers themselves, may use ETICS for quality-assured IC services to connect to the IPX network of bigger providers.
- It may provide IC services for Domains that are currently separated due to their technological heterogeneity, eg, connectionless and connection-oriented Domains may be interconnected.
- ETICS can provide flexible IC services with QoS guarantees avoiding the classical IC hierarchy and peering limitations¹³.

IMS in this context mainly refers to the usage of the PCRF (PDP) that interacts with the policy profiles, ie, mainly the SPR. With this process the profile of the user SPR, eg, in the future reflecting quality class, WTP, QoE-sensitive to particular applications, can be applied to the last mile transmission where resources are typically very scarce. Beyond the last mile, QoS demands are very likely treated as aggregated request, as envisioned by ETICS. The IC case requires contractual agreements on top of heterogeneous technologies. Today's peering and transit solutions are incapable of mapping QoE demands of individual users and specific service usages to appropriate inter-carrier solutions. Hence, ETICS or comparable approaches are required in order to assure an end-to-end QoE- or even WTP-aware transmission.

In the context of this thesis, ETICS can, hence, provide a promising mechanism in the core network and IC parts of the end-to-end transmission in order to enable QoE- and WTP-aware network transmission. The utility figures as, for example, reconditioned in Section 5.3 and derived from the experiments in Chapter 3, can be integrated into the

¹³ Both private and public peering services are mostly designed to be data exchanges among equal partners without relevant monetary payments, which may exclude smaller players and may not be used for multi-hop IC.

resource provisioning close to the access network where resource scarcity is severest. As described in Section 2.2.1, within the access networks the transition points will translate individual user demands (involving QoS, QoE and potentially price considerations) to QoS or network demands, which are then attached to aggregate QoS classes. Aggregate QoS classes can efficiently be handled in the core and can easily be agreed upon in the IC segments, eg, by using SLAs as envisioned by ETICS. The agreement with the user is more complex due to the *experience good* nature of network quality and the complexity of coming to terms with consumers that are not legally trained. In this arena, the recently emerged concept of Experience Level Agreements (ELAs) has been proposed in [Var+15] in order to complement ETICS’s SLAs approach by a consumer equivalent. ELAs are, however, out of the scope of the present work.

5.3 Controlled Service Degradation: NSP and User Utilities

♣¹⁴ Many works like [Fu+13] treat QoE as equivalent to utility, although cardinal¹⁶ utility in the sense of Von Neumann’s definition [VM44] is bound to rationality and linearity. QoE, which is closely related to CS measures, however, is a non-linear measure of perception that quickly saturates. When switching from a user to a customer role, the transition from QoE to utility is required alike. Despite the obvious and notable influence of QoE on the user utility (and potentially NSP utility), QoE and utility are in particular disparate for a variety of reasons:

1. QoE non-linearly relates to QoS input, eg, often an exponential [FHT10] or logarithmic relationship [Rei+13a] (relating to the Fechner scale [Hei04]) has been found. In analogy with the discussions in Dagstuhl seminar 15022 [Zwi15], QoE could be measured using a “chocolate metric”. Both the joy of consuming more chocolate bars and iteratively improved QoS will drastically drop. While the first few chocolate bars may be pleasurable, the utility may converge towards zero after a few bars. Utility in contrast is a generic and rational concept, which linearly scales in (almost) generically representable metrics. A good example could be currencies¹⁷ like USD \$. The association of QoE to monetary spending behaviours is non-trivial and insufficiently systematised—eg, see the WTP measurements of the *2012 trial* [ZSR13], which clearly separates QoE curves from WTP curves.
2. The appreciation of high QoE (and indirectly QoS) need not trigger quality upgrades or monetary spendings in general. The monetary spendings, closest related

¹⁴ The results of this section have been reported in [Zwi+15] before. The *Mathematica*¹⁵ code used to obtain the results is publicly accessible at <https://github.com/pzwick1/PZ-Thesis/tree/master/Utility>.

¹⁶ Cardinal refers to the inclusion of strength differences in the assessment. By contrast, ordinal utilities are used for creating ordered ranks of utilities where absolute differences are ignored.

¹⁷ Of course when looking at the spending behaviour, the essential *need* for an extra dollar (ie, margin utility) will also decrease, but the monetary value will not be affected by the personal needs or spending behaviour. Hence, the metric is linear, the perception metric may still be non-linear.

to the NSP's utility, cannot be assumed to be identical to QoE representations on general basis. This effect is further underlined by observations of cognitive dissonance [Sac+12b] that have illustrated distortions of QoE curves induced by active purchasing situations, ie, prices.

3. Contrary to the generic utility concept, QoE ratings are subject to specific test scenarios, which are characterised by the used scenarios, test parameters and their ranges, tariffs, service preferences, context, etc. A QoE rating, eg, a value on ACR-5 scale [P.800], has only a local validity. This has partly been previously described in [Pin+12] where the comparison of MOS values is not recommended without the smoothing (data fitting) of the data and the usage of optimally the same group of subjects. This may enable the direct comparison of almost identical trials in terms of comparing result tendencies, but otherwise the results still remain of local nature (see Section 4.1.3 for a detailed discussion on QoE result's locality). This limitation is partly induced by training sessions (eg, as recommended in [P.910]) that expose users to the test conditions and ranges prior to the actual testing. In other words, the knowledge of test ranges biases the testing and limits it to local interpretations. This characteristic rules out meta-analyses and comparable techniques. On the other hand, different QoS input metrics and hardly comparable service conditions create additional challenges for an on par comparison of results.
4. The contextual service preference is typically not tested in QoE trials. For example, when users start to consume a VoD service, their utility of switching to an audio book will be strictly lower. The QoE in classical tests is, however, not affected. While integrated tests may provide indications on the forced service switching, the linkage and comparison of solitary QoE ratings is non-trivial.

Conclusion #5.1. *Utility and QoE are disparate concepts.*

Despite these disparities and the need for utility figures in order to parameterise technoeconomic optimisation models, the mapping of isolated QoE results to a utility-equivalent form has not yet emerged. Dedicated WTP or utility measurements are rare and limited to a few tested services (see the *2011 trial*, *2012 trial* and related work in Section 3). This section, hence, designs a technique to derive inter-service utilities mainly derived from the concatenation of QoE and WTP trials, and will later on illustrate this technique in a case study. This technique is tailored to the case of controlled degradation of services, as for example coined in [ISM14].

Conclusion #5.2. *Utility data are inherently required for further optimising the operations of communications networks.*

The focus on controlled degradation leads to the following assumptions that will be used subsequently:

1. In accordance with common QoE trials, a single **one-shot purchasing decision** is studied with constant scenario, location, time, budget, interests, WTP and user context (lighting, noise, etc.).
2. Users have an intention to use a network-based service, eg, video streaming or web surfing, and might become customers of this service.
3. Users have a clear relative preference between any two services. Thus, all users have a **strictly ordered preference list**¹⁸ for all considered services $s_{1...n} \in S : s_1 > s_2 > \dots s_{n-1} > s_n$ (where $s_1 > s_2$ represents a strict preference of s_1 over s_2 under the current conditions). For example, when users want to watch an episode of their favourite TV series at home at this moment (primary service), the utility for any different activity (eg, watching a sportscast) is lower.
4. Users are **willing to pay** more for **services they prefer** over any alternative (eg, TV series over sportscasts). In other words, when users are intending to watch a specific HD video stream, they have a lower consumption utility for any alternative activities. Whenever the QoE is acceptable, the following condition will hold:

$$WTP(s_k) > WTP(s_{k+1}) > WTP(s_{k+2}) \dots ,$$

where $k = 1$ is the primary service choice and $k + 1$ is the next best alternative. Otherwise, the situation may be reversed.

5. The controlled degradation is triggered whenever the QoE is critical, non-optimal and can be improved by switching to a less media-rich service, eg, switching from poor HD to acceptable SD streams.

5.3.1 Approximation Model

The subsequent 3+1-stage approximation model for NSP and user utilities will be carefully designed around known empirical anchor points. On the one hand, classical QoE data are used as central starting point, as comparable data including price cognitions are rare. In addition, available WTP data of services are transferred to the actual case on the other hand.

The first three stages are treating representative subjects, ie, subjects (users or customers respectively) with average usage behaviours as aggregates of real world individuals. The fourth stage provides a complementary individualisation to reflect the practical distribution of customers to customer segments.

¹⁸ This assumption represents a simplification of the reality for modelling purposes, which may not be generally applicable, especially for less concrete usage scenarios.

5.3.1.1 Stage 1 (S1): Subjective QoE Data Without Price

At first, solitary QoE assessments need to be concatenated, eg, linking the ratings of tests on SD and HD video streams (cf. Fig. 5.2). The concatenation requires a recalibration of the raw QoE data in order to incorporate the strictly ordered user preferences.

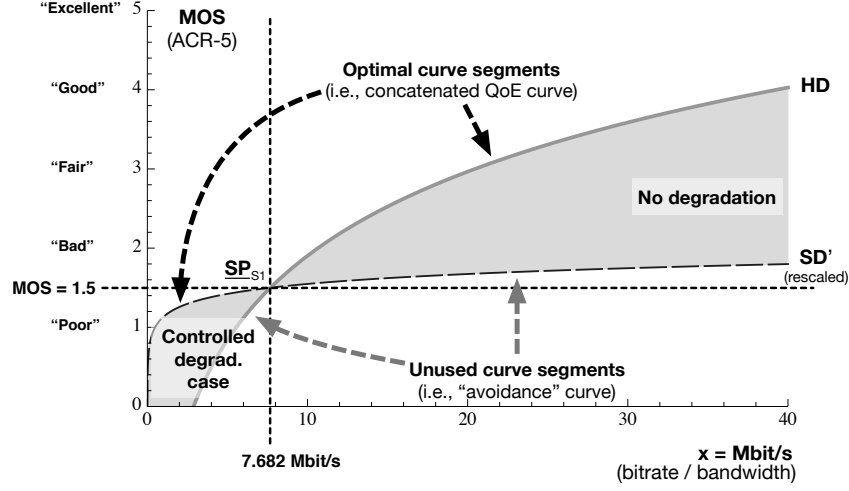


Figure 5.2: QoE for controlled VoD service degradations.

The process is as follows:

1. The isolated QoE curves for all considered services are transferred to a functional representation, eg, via the commonly used exponential [FHT10] or logarithmic curve [Rei+10] fittings.
2. A switching MOS is defined at the point where the user becomes indifferent between two services, the initially preferred service k and the next best alternative $k + 1$ (ie, adjacent services in the ordered preference list). The switching MOS has to be outside the range of acceptable QoE conditions for service k . Together with the QoS, eg, bandwidth, requirements of the service, a *Switching Point (SP)* for stage 1 SP_{S1} is formed where a service adaptation is triggered, ie, the system switches from service s_k to s_{k+1} . This process stands in analogy with Gomez *et al.*'s [Góm+13] QoE- and QoS-aware network management where adaptations actions are only triggered whenever the QoE drops below a defined threshold.
3. Whenever no empirical evidence exists that specifies the switching MOS at the SP, we recommend an estimation from the extent of the service (type) degradation and the associated disutility. For limited quality differences, for example when degrading from Full HD (1080p) to 720p video streams, we suggest to set the switching MOS to 3.0 (on ACR scale) for s_k , ie, “fair” quality, while a drop to Common Intermediate Format (CIF) is better reflected by a MOS of 1.5. Whenever not even the content

type can be retained, eg, instead of a video clip an audio stream is provided, the switching MOS should be below 1.5. With the existence of more comparable QoE data, a more fine-grained chain of service degradations can be constructed (based on a more extensive list of service (type) preferences and their parameterisations).

4. The MOS data of each service s_{k+1} (eg, SD video) are rescaled to peak at the switching MOS of the directly preferred alternative s_k (eg, HD)—see the derivation of the rescaling function in (5.1). With this process, the QoE for service s_k will strictly be higher in an acceptable QoS range. Outside this range, users may reconsider their preference. SP_{S1} in particular forms the new MOS maximum for service $k+1$ relative to k instead of the locally measured maximum MOS_{max} of 4.5, for example. As the considered QoS range (minimum, maximum) will not be altered, the new QoE curve will be less QoS-sensitive, ie, in relation to the QoS input the curve will receive a flatter shape.

For the rescaling of s_{k+1} 's QoE function relative to s_{k+1} 's, a weighting function m is used:

$$m(MOS_{k+1}(QoS)) := MOS_k(QoS) \quad (5.1)$$

where we further know that the MOS of service k at the switching point SP_{S1} ,

$$MOS_k(QoS) = MOS_{SP_{S1}} \quad , \quad (5.2)$$

and the next best alternative $k+1$'s maximum MOS,

$$MOS_{k+1}(QoS) = MOS_{max} \quad , \quad (5.3)$$

has to be equal when applying m , as in (5.1). Finally, we can formalise that the minimum MOS is not affected by the functioning of m ,

$$MOS_{k+1}(QoS) = m(MOS_{k+1}(QoS)) = MOS_{min} \quad . \quad (5.4)$$

This forms an equation system, from which m can be derived when applying empirical data input, as illustrated in Section 5.3.2. The parameterised m allows the formation of a concatenated experience curve across services.

Conclusion #5.3. *Separate QoE results, as local QoS-to-QoE mappings, need to be concatenated using context-specific adaptations in order to reflect the user's appreciation across service types and test scenarios.*

5.3.1.2 Stage 2 (S2): User Utility

♣¹⁹ In this phase, the attention shifts to the analysis of QoE ratings that involve price cognitions p , ie, QoE^p in order to derive user utilities. Price cognitions refer to the cognitive

¹⁹ This section aims at providing an economic complement to QoE-based metrics.

process that is triggered by the creation of a purchasing situation, where customers also have to set the price information in relationship to the quality offering. Due to some effects induced by the addition of price cognitions (eg, expectations or justification of purchases), the shapes of classical QoE trials such as given in [KRD12] can neither be directly compared to QoE^p nor retained in this stage. The shapes will be derived from an empirically tested case (w.l.o.g. we will assume it to be s_k). The curves for s_{k+1} 's can be approximated as follows:

1. SP_{S1} is transferred to its stage 2 representation SP_{S2} as follows: The point on the QoE^p curve for s_k , ie, QoE_k^p , where the QoS demands are identical to the value at SP_{S1} , eg, 7.682 Mbit/s in Fig. 5.2, forms the SP at stage 2, ie, SP_{S2} .
2. In analogy with S1, we rescale QoE_{k+1}^p such that its considered maximum intersects at SP_{S2} with the preferred service's curve QoE_k^p . As the empirical evidence for QoE with price cognitions is limited, the following approximation is recommended whenever insufficient data exist: In order to form QoE_{k+1}^p the known QoE_k^p curve can be rescaled between the considered minimum MOS_{min} ($QoE = 1.0$) and maximum point $(QoS(SP_{S2}), QoE(SP_{S2}))$. The service-specific minimum point is inferred from S1 for s_{k+1} and shifted according to the slight relative movement of s_k 's minimum point from S1 to S2. The entire rescaling process is again described by a weighting function $w()$.

In analogy with S1, the curves can be concatenated and normalised in $[0, 1]$ to form $QoE_{k \cup k+1}^p$, ie, the **user utility** aggregating price and quality considerations.

Conclusion #5.4. *User utility has to integrate both purchasing decisions (ie, monetary aspects) and QoE considerations.*

5.3.1.3 Stage 3 (S3): NSP Utility

The third stage concentrates on the formation of NSP utility figures. Central concepts for NSP utilities are profits and revenues. WTP and demand characterise the expected revenues. However, both WTP and demand data are sparsely available for communications services, especially when involving third-degree price discrimination or inter-service adaptations. The QoS-to-QoE relation is further distinctly different from QoS-to-WTP equivalents.

Conclusion #5.5. *As the NSP utility does only indirectly relate to QoE measurements, a separate quantification is required.*

Conclusion #5.6. *The NSP utility directly relates to WTP and demand figures, which are insufficiently available—an approximation from QoE is required.*

In order to match untested services, the tested WTP curves will be shifted by a factor derived from S_2 , ie, the **Average Service Preference Weighting (ASPW)**—ie, the average degree of preferring the primary service k over a given alternative $k + 1$. The ASPW has to satisfy the following conditions:

1. **Relative proportions:** The more often²⁰ an adaptation to s_{k+1} is necessary, the proportionately less valuable is s_k .
2. **Service-agnostic:** ASPW should be service agnostic. As long as services can be represented on comparable scales, a calculation should be possible.
3. **Indifferent at SPs:** As the user is indifferent between the quality of services s_k and s_{k+1} at SP_{S_3} , their WTP needs to be identical at this point.
4. **Sensitivity-aware:** When lowering the QoS sensitivity for a service, the WTP should be higher or kept more stable for low QoS inputs.

An ASPW factor satisfying these conditions can be formulated as follows:

Definition 6 (ASPW). *W.l.o.g consider the concatenation of service $k + 1$ to service k . Then ASPW is the ratio of the minimum QoE point P_{min}^{k+1} 's Euclidean distance $\mathcal{E}$ to SP_{S_2} and P_{max}^{k+1} , where point $P_{max}^{k \cup k+1}$ is where $QoE_{k \cup k+1}^p = MOS_{max}$ (MOS_{max} , for example, remains 4.5) and point P_{min}^{k+1} is where $QoE_{k+1}^p = MOS_{min}$, eg, where service $k + 1$'s MOS is 1.0:*

$$ASPW_{\{k,k+1\}} := \frac{\mathcal{E}(P_{min}^{k+1}, SP_{S_2})}{\mathcal{E}(P_{min}^{k+1}, P_{max}^{k \cup k+1})} . \quad (5.5)$$

The resulting factor ASPW ranges in $[0, 1]$, as the distance in the denominator will also be greater than for the numerator and for both Euclidean distances $\mathcal{E}(\dots) > 0$ holds.

The illustration in Fig. 5.3 gives an example of ASPW as used in Definition 6, based on data obtained from the *2012 trial* and normalised according to the process given in S_2 . The Euclidean distances $\mathcal{E}$ (blue arrow to SP_{S_2} ; red arrow to $P_{max}^{k \cup k+1}$) are necessary in order to accommodate for non-linear QoS to QoE^p in various shapes. The fraction is necessary for normalisation reasons. When shifting SP_{S_2} to the position of SP'_{S_2} , as depicted in Fig. 5.3, the ASPW value will be lowered.

With these characteristics, the ASPW is intended to be used as multiplier to shift the WTP for s_k to match s_{k+1} (ie, satisfying condition 3 and 4)—when lowering the ASPW value the WTP for s_{k+1} will drop. The NSP utility, as normalised revenue metric, is primarily influenced by the demand for a quality-price combination (eg, see WTP curves for *2011* and *2012 study* and the charged price. Due to the non-trivial estimation of the demand or WTP curves for s_{k+1} , the underlying (often linear) tariffs will be adjusted instead in order to match the known demand points, eg, from the WTP values of the

²⁰ In the case of dropping QoS, this refers to the requirement of an earlier adaptation.

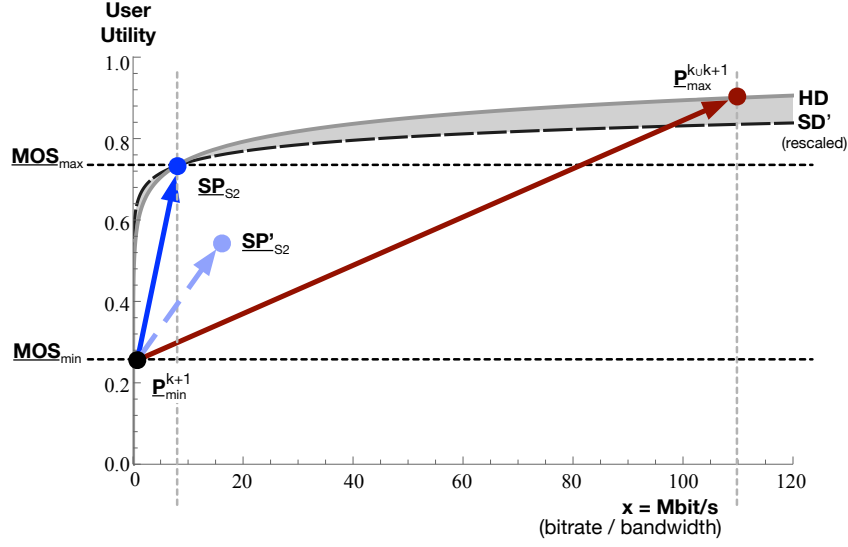


Figure 5.3: Illustration of the ASPW using normalised data from the 2012 trial.

2012 trial, across services s_k and s_{k+1} . Thus, while the demand is kept static the price is affected by ASPW in order to shift the revenues and, thus, the utilities. The full extent of ASPW will be effective at the maximum of the considered bitrate range b_{max} (eg, where $QoE_{b_{max}}^p = MOS_{max}$) the difference will be determined by the multiplier ASPW.

Thus, when shifting the SP_{S2} to a point with higher bandwidth demands, the service k becomes more sensitive, which would increase the value of the service $k+1$. The MOS value at SP_{S2} reflects the satisfaction with the current pricing, ie, the higher the value, the smaller the quality sensitivity of the demand. Hence, the condition 1 is satisfied. Apparently, this formulation of ASPW is service-agnostic (see condition 2). When normalising the prospective revenue in $[0, 1]$, the **NSP utility** results in a comparable form to the user utility.

5.3.1.4 Stage 4 (S4): Individualisation (Optional)

Finally, orthogonally to representative subjects, the ASPW (and other factors) may not entirely characterise the preference of each particular subject. Specific effects may be introduced when considering real world customer distributions, ie, customer segments. Contrary to the aggregate data of representative subjects, subjects may prefer “premium” qualities, for example, at almost all costs for one service, eg, audio, while they may have a low WTP for video qualities. The details of stage 4 go beyond the scope of the present work, but we recommend to conduct future work on the corresponding parameterisation.

5.3.2 Case Study

Comparable to the case study in Section 2.5, the utility approximation model of Section 5.3 will be applied to the controlled degradation of VoD streams from HD, s_k to SD (ie, CIF), s_{k+1} . Wherever publicly available and compatible, empirical data will be used. Otherwise, estimator models will be used as replacement. For QoE, MOS-based curves will be used.

5.3.2.1 Stage 1 (S1): Subjective QoE Data Without Price

The Video Quality Metric (VQM) [PW04] is an efficient QoE estimator[Vid03] for CIF videos²¹. Through a reverse engineering of the highly QoE-sensitive “soccer” data from [Fu+13], the following logarithmic curve (as recommended by [Rei+10]) can be derived:

$$MOS_{SD}(b) := 4.355 + 0.697 \cdot \log(b) \quad . \quad (5.6)$$

The curve $MOS_{SD}(b)$ has a coefficient of determination R^2 above 0.99²², which supports the used fitting approach.

Correspondingly, from the reconditioned empiric QoE data²³ for the “CrowdRun” sequence (quality-demanding; 1080p50) in [KRD12] a logarithmic MOS fit for HD²⁴ videos can be created with an R^2 of almost 0.99:

$$MOS_{HD}(b) := -1.624 + 1.532 \cdot \log(b) \quad , \quad (5.7)$$

where b represents the bitrate in Mbit/s. While both the S1 and S2 data can be retrieved from [ZSR13; SZR13a], the more extensive results from [KRD12] will be used for the QoE-only case in S1.

As a result of the significant experience difference between HD and CIF content, the SP_{S1} will be positioned at $MOS = 1.5$. This corresponds to an intermediate position between “bad” and “poor” experience of s_k . Using a specific form of the weighting function m , MOS_{SD} can be rescaled correspondingly:

$$MOS_{SD}(b)' := z \cdot (v + 4.355 + w \cdot 0.697 \cdot \log(b)) \quad , \quad (5.8)$$

where the parameters v (offset), w (steepness) and z (maximum rescaling) are instantiations of the $w()$ function for logarithmic curves. We obtain $v = -0.716$ and $w = 0.606$ around the $SP_{S1}(b, MOS_{SD}, MOS_{HD}) = (7.68, 4.5, 1.5)$. The parameter $z = \frac{1.5}{4.5} = \frac{1}{3}$ where 4.5 is SD’s original maximum and 1.5 is the new maximum²⁵. The resulting concatenated QoE curve is depicted in Fig. 5.2. At any available bandwidth the highest MOS is preferred by the user, ie, the concatenated result only consists of optimal curve segments from the initial SD and HD curves.

²¹ CIF resolution is only 352×288 pixels.

²² Almost 99% of the variance is explained by the model.

²³ The data have been rescaled to ACR-5.

²⁴ HD resolution is 1920×1080 pixels.

²⁵ Values above 1.5 result from the functional form and are practically irrelevant.

5.3.2.2 Stage 2 (S2): User Utility

Corresponding empirical material for the S2 and HD case is used from [ZSR13; SZR13a]. When logarithmically fitting the data MOS^p for an aggregate maximum price p_{max} around €3 (due to sample size reasons) the following functional form is yielded:

$$MOS_{HD}^p(b) := 3.014 + 0.316 \cdot \log(b) \quad , \quad (5.9)$$

where b denotes the video bitrate in Mbit/s and p the price cognitions. This shape will be reused for the untested SD service (s_{k+1}), which is, however, rescaled to match its specific S1 minimum, relative to HD's point, and maximum at the SP_{S2} (7.68 Mbit/s as in Stage 1; $MOS_{HD}^p(SP_{S2}) = 3.65927$). Thus, the following SD curve results:

$$MOS_{SD}^p(b) := z \cdot (v + 3.014 + w \cdot 0.316 \cdot \log(b)), \quad (5.10)$$

where $v = 0.999$, $w = 0.753$ and $z = 0.813$ (calculated as in S1). When normalising the concatenated curve $MOS_{SD \cup HD}^p(b)$ (best MOS for each bandwidth setting) to $[0, 1]$ with the considered minimum 0 at $MOS^p = 1.0$ and the considered maximum 1 at $MOS^p = 4.5$, the per service **user utility** $\mathcal{U}^u$,

$$\mathcal{U}_{SD \cup HD}^u(b) := \frac{MOS_{SD \cup HD}^p(b) - 1.0}{3.5} \quad , \quad (5.11)$$

and its functional equivalent,

$$\mathcal{U}_{SD \cup HD}^u(b) := \begin{cases} 0.286 \cdot (-1 + 0.813 \cdot (4.014 + 0.238 \cdot \log(b))) & \text{if } b \leq 7.682 \text{ Mbit/s} \\ 0.286 \cdot (2.014 + 0.316 \cdot \log(b)) & \text{otherwise} \end{cases} \quad , \quad (5.12)$$

is obtained.

The normalised $\mathcal{U}_{SD \cup HD}^u(b)$, as also depicted in Fig. 5.4, is substantially flatter, ie, less QoS-sensitive, than the classical QoE results from S1.

Conclusion #5.7. *The user utility $\mathcal{U}^u$ can slightly be increased when applying an intelligently controlled degradation. Benefits due to low prices for low qualities mitigate the potential for user utility gains.*

5.3.2.3 Stage 3 (S3): NSP Utility

The known demand d_{HD} (see the beta distributed results $B(x)$ in Section 3) and customer segments curves (see the normally distributed data $N(\mu, \sigma)$) for HD is reused for the SD case. The underlying price figures will be adapted to match the SD case by using the ASPW. As a result of the Euclidean distance $\mathcal{E}$ between both SP_{S2} and P_{max} ($b = 22.557 \text{ Mbit/s}$, $MOS = 4.5$), and P_{min} (0.484 Mbit/s , 1.0), $ASPW = 0.345$ can be retrieved. Satisfying the ASPW condition 3, the SD price p_{SD} can be derived around the given p_{HD} in order to keep $d_{SD} = d_{HD}$ and $WTP_{SD} = WTP_{HD}$:

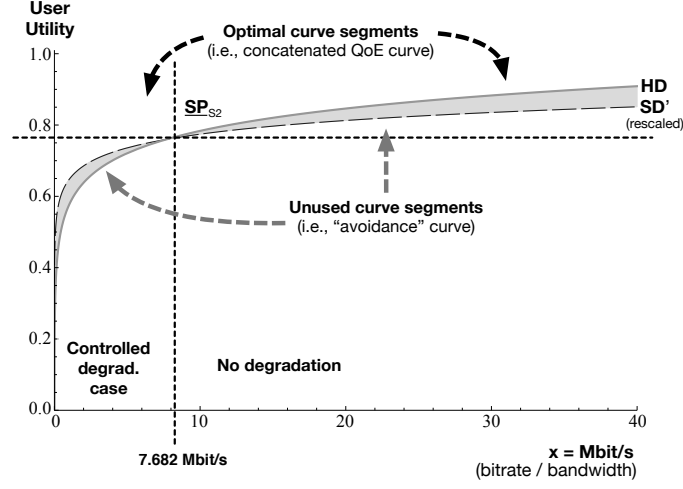


Figure 5.4: The inter-service user utilities for HD and CIF VoD streams.

$$p_{SD}(b) := 0.577 + 0.0153 \cdot b \quad , \quad (5.13)$$

$$p_{HD}(b) := 0.0919 \cdot (-0.128 + b) \quad . \quad (5.14)$$

The NSP revenue follows from both the price and demand figures (see *2012 trial* in Section 3.3.3 or [ZSR13]). Due to separate demand figures, a differentiation in the case where users can choose from a single offer (Fig. 5.5a) or multiple offers (Fig. 5.5b) is made. A linear price curve, subject to the chosen quality level, is used. The NSP can always profit from controlled degradations whenever the QoS is below the critical point SP_{S3} (which mirrors SP_{S2} point to revenue curves). Counterintuitively, high revenue levels are further obtained for very low qualities. These values result from the high demand for quality levels Q_x with $x < 10$ (cf. Fig. 5.5), which are shaped by the linearly increasing prices. This effect is even more apparent when looking at the SD content, where users cannot profit from high-quality classes and are, hence, not willing to purchase those for the given price.

The NSP utility follows from the normalisation of the CDF of the revenue curves in $[0, 1]$ (cf. Fig. 5.6 based on both single- and multiple-choice data; also see Fig. D.1 for an illustration of utility gains). Any QoS decline will lower the NSP utility (cf. Fig. 5.6). Below the inflection point SP_{S3} , controlled degradation successfully improves the NSP utility.

Conclusion #5.8. *The NSP utility $\mathcal{U}^n$ can be increased substantially when offering multiple alternatives to customers. Whenever focusing on a single offer, $\mathcal{U}^n$ resembles the limited gain potential of $\mathcal{U}^u$.*

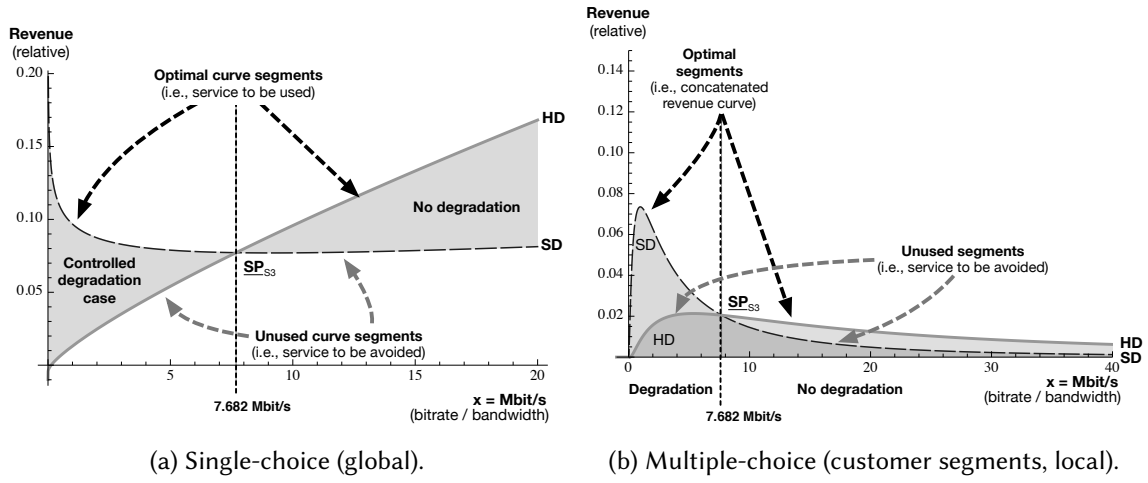


Figure 5.5: NSP Revenue figures for tariff B.

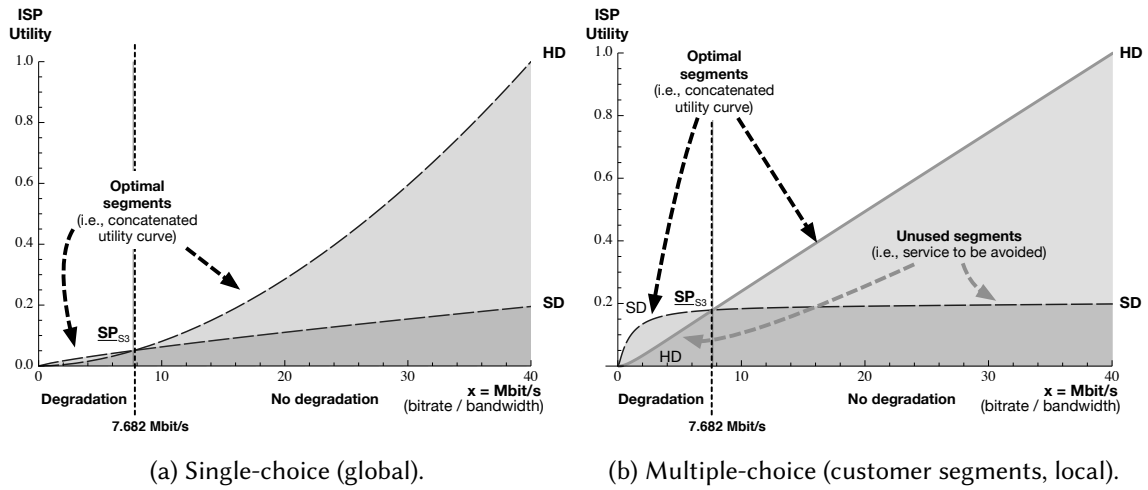


Figure 5.6: NSP utility figures $\mathcal{U}^n$ for tariff B in $[0, 1]$ based on demand CDFs and new price curves.

5.3.2.4 Practical Implementation & Operational Use

For a sustainable configuration in operational use, the position of the SP has to reflect both the user $\mathcal{U}^u$ (long term retention of customers) and the NSP utility $\mathcal{U}^n$ (short term profits) to some extent. Depending on the chosen objective, various flavours of social welfare optimisations may be applied:

$$c \cdot \mathcal{U}_{SD \cup HD}^u(b) + d \cdot \mathcal{U}_{SD \cup HD}^n(b) \quad (5.15)$$

with $c + d = 1$ where the state $c = 1.0$ is user utility, $d = 1.0$ is profit, and $c = d = 0.5$ is classical social welfare optimisation. Later on d will be replaced expressed by $d = (1 - c)$.

The proper parameterisation of SP_{S1} , which is propagated to $S2$ and $S3$, requires empirical evidence. This will remain for future work. Empirical validation of the entire approximation model is additionally required in order to clarify the model's accuracy and to scrutinise the optimisation potential.

For a technical usage of this model, the quantification of a studied use case is transferred to a code representation, eg, Java. Only limited modifications are necessary to the functional forms that are created in $S3$. Alternatively, utility arrays, as given in Section D.1, can be used as inputs for code-based computations whenever no concatenated curves $\mathcal{U}_{SD \cup HD}^u$ and $\mathcal{U}_{SD \cup HD}^n$ can be calculated.

5.3.3 Transmission Path

♣²⁶ Building upon the collected utility information, the network transmission path can be made utility-aware for each flow in order to run social welfare or profit optimisations. This section will design a mechanism that replaces Q-POINT's QoE-aware path optimisation, as described in [Dob+14] and revisited in Section 5.1.8.1, by a utility-aware model. The focus on controlled service (type) degradation is retained and the approximated data are taken from Section 5.3.

Q-POINT minimises the collective (considering a series of network flows and service types) QoE impairments due to packet loss for the entire flow path. Due to the focus on QoE, utility data have been left aside. On the other hand, the utility approximations given in Section 5.3 are subject to a single-link assumption and study utility variations induced by bitrate variations, which relate to network bandwidth demands. The present section, thus, has to overcome this mismatch of perspectives in order to systematically optimise social welfare for a set of end-to-end network transmission flows.

“ Interaction between metrics: Naturally, the various quality metrics are inter-dependent on each other. For example, streaming video at a higher bitrate would lead to better quality. However, ..., it would take longer for the video player buffer to sufficiently fill up in order to start playback leading to higher

²⁶ This section results from a joint work with UNIZG-FER.

join times. Similarly, streaming video at higher bitrates leads to higher rates of buffering ...

Balachandran et al., 2013 [Bal+13]

“ One important research question with regards to the perceived quality of the video flows, is to find whether the relevant parameters have any influence between them and with the MOS score. For example, if we could state that there is a direct correlation between certain codec-related parameters and the MOS under specific network conditions, then we could use this information for deciding how to code and transmit the video stream in the most efficient way. ”

Frank & Icera, 2006 [FI06]

When linking the data perspective of Q-POINT with utility approximated from the 2012 trial, the linkage of various QoS metrics affecting QoE, eg, delay, packet loss or bandwidth (bitrate), is crucial [Bal+13; FI06]. While there is a plurality of single impairment models that concentrate on “packet-level and bitstream” [Sch+13] effects such as [SFC10; Hoß+11; AL08], there is no agreement on the interdependence of these metrics. While a few models (eg, [LC12; Kil08; PMH06; Gee+10]) have described the conceptual link between various kinds of QoS metric, sometimes including both application and network QoS parameters, a quantitative relationship has only been openly discussed in a few works on packet-level:

1. Yamagishi & Hayashi [YH08]: The authors use a multi-metric model that links bitrate-induced video distortions (integrating objective quality estimations) to packet loss effects. As shown in

$$Vq := 1 + Ic \cdot \exp\left(-\frac{l}{v4}\right) \quad (5.16)$$

$$Ic := v_1 - \frac{v_1}{1 + \frac{b}{v_2} \cdot v_3} \quad , \quad (5.17)$$

where Ic describes a bitrate- and codec-specific video quality value, $v_1 \dots v_3$ are codec-specific parameters derived from subjective evaluations, $v4$ is the packet loss sensitivity of the used codec, the packet loss l^{27} only lowers the overall video quality Vq if $l > 0$ and b is the video bitrate (corresponding to the network bandwidth demand). The exponential function $\exp$ represents a weighting function for l for the aggregation process towards deriving Vq .

2. Siller & Woods [SW03]: The authors consider both the interdependence of application and network QoS metrics. Network QoS parameters, eg, jitter, delay and

²⁷ The model of [YH08] focus on bursts of packet losses in a given time span.

packet loss, are aggregated by using weights. While the numeric values (or corresponding equations) are not given in [SW03], a memory effect over measurement rounds is described in detail. Siller & Woods assume that the qualitative evaluation is relative to previous experiences, which is consistent with the market entrance effects on utility level described in Section 3.6.

Another interesting model, which describes the interdependence between QoS metrics, is given in [Kil08; Kil12]. However, quantitative details are not openly accessible.

On that account, the interdependence of metrics is best assumed to be moderated by weighting factors or weighting functions.

Conclusion #5.9. *Following QoE literature, the interdependence of QoS metrics jointly affecting the QoE outcome is most likely characterised by specifically parameterised weighting factors or functions.*

Due to linear nature of the Q-POINT optimisation model, the following objective function is derived in analogy with (5.15)²⁸:

$$OBJ_{max} := c \underbrace{\left(w \cdot \mathcal{U}^u(l) + (1 - w) \cdot \mathcal{U}^u(b) \right)}_{=weighted\ user\ utility, ie, \mathcal{U}^u(b,l)} + (1 - c) \underbrace{\left(z \cdot \mathcal{U}^n(l) + (1 - z) \cdot \mathcal{U}^n(b) \right)}_{=weighted\ NSP\ utility, ie, \mathcal{U}^n(b,l)} \quad (5.18)$$

where $\mathcal{U}^n(l)$ further refers to a normalised form of $MOS(l)$ (ie, normalised QoE function for packet loss impairments l , eg, see [Dob+14]), $\mathcal{U}^u(b)$ is the user utility subject to the video bitrate (see Section 5.3.2.2), $\mathcal{U}^n(l)$ is the unmeasured NSP utility for the NSP utility under packet loss impairments, and $\mathcal{U}^n(b)$ is the corresponding NSP utility subject to bitrates (see Section 5.3.2.3). All factors such as $\mathcal{U}^n(l)$ are normalised in $[0, 1]$ and in conformance to the results obtained in Section 5.3.

In addition, the following three **weighting factors** are used in (5.18):

- c:** The weighting factor c describes whether the user utility (c converges to 0), the NSP utility (c converges to 1) or social welfare (ie, $c = 0.5$) is optimised in accordance with the considerations in (5.15).
- w:** w and z , which are set to 0.5 in the case no further evidence is available, moderate the interdependence between QoS metrics. w is, hence, used for weighting the user's utility sensitivity of packet loss l relative to bandwidth b .
- z:** The factor z is the complement of w for the NSP utility.

As illustrated in Section 5.3 and described in [SS14], price cognitions primarily affect the end points of the QoE curves, ie, the QoE curves are flatter around the observed maximum and minimum values, which brings the curve closer to a linear representation. QoE

²⁸ While the focus is hereinafter set on HD video quality, the described process is transferable to a multi-service system.

curves involving price cognitions (ie, a purchasing decision) represent user utility, but as described in Section 5.3 disallow a direct characterisation of the NSP utility. Due to the linear programming approach of Q-POINT, which linearises QoE curves, the discrepancy between the measured user utility $\mathcal{U}^n(l)$ (followed from QoE under price cognitions) and classical QoE ratings $MOS(l)$ will fade. As a result, it can be assumed that the used and linearised QoE curves are an acceptable approximation for user utility $\mathcal{U}^u(l)'$, which introduces additional marginal error and will likely outperform assumption-prone alternatives²⁹. Considering the trade-off between calculability, ie, complexity reduction due to the linearisation induced by linear programming in Q-POINT, and precision considerations (affected by the linearisation), the Q-POINT approaches seems to be balanced for complex end-to-end quality-differentiated networks with dynamic provisioning (ie, highly complex and dynamic networks). The parameterisation of $\mathcal{U}^n(l)$ deserves even more attention, which is discussed in Section 5.3.3.1.

Conclusion #5.10. *The utility-aware network flow path computation requires a considerate handling of complexity increases due to precision advancements.*

5.3.3.1 Parameterisation and Proof

Equation (5.18) will be parameterised in this section in order to match the empirical data of the VoD case study from Section 5.3. The $\mathcal{U}^u(l)'$ data can be derived from (5.19) by normalising it in $[0, 1]$ due to the absence of QoE ratings under purchasing situations³⁰.

The corresponding $\mathcal{U}^n$ data required more attention. $\mathcal{U}^n$ will be synchronised around the maximum quality where the utility is the highest. This approach follows the key assumption that utility converges to a single utility maximum point whenever neither bitrate reductions nor packet losses occur, ie, optimal quality (context-specific!) is available. Bitrate reductions or packet losses will deteriorate the QoE, which will lower the WTP. In other words, the utility of 0% packet loss will be assumed to be identical to the highest bitrate considered in the *2012 trial* where the packet loss was 0%. To retain the measured utility as presented in Section 5.3, the parameterisations of the factors w and z can be derived³¹: $w = z = 0.5$. Due to a focus on social welfare optimisation c can further be fixed at 0.5, which equally weights user and NSP utilities. All these definitions are consistent with the model definition around (5.18).

The empirical WTP results have been measured for a range of 128 kBit/s (ζ^{min}_{HD}) to 32768 kBit/s (ζ^{max}_{HD}) where higher bitrate tendentially leads to better QoE ratings. The maximum of QoE of 4.5 has not been reached by the smoothed QoE curves involving price cognitions when using a realistic QoS range. Packet loss, in the form of mean percentages

²⁹ This design decision is bound to the linear programming solver and may not be suited to other kinds of optimisation.

³⁰ Due to the linearisation process required for Q-POINT (linear program) the introduced additional error is limited—see detailed argumentation above.

³¹ This follows the assumption that the user utility is also highest where the QoS is best. Due to pseudo irrationalities this effect may not always be observed.

of lost packets, will negatively affect the QoE as it grows. When reconditioning and rescaling the packet loss curve used in [Dob+14],

$$MOS(l) := 4.7 - 0.63 \cdot l \quad , \quad (5.19)$$

to take the input from a positive QoS input metric l comparable to bitrate (ie, bandwidth) in the range between QoE_l 's minimum QoE of 1.0 (equivalent to 4949.25 kBit/s of the corresponding WTP data) and the empirically measured QoE maximum (ie, 4.11803) at 32768 kBit/s (ie, ζ^{max} , which is shared by both services), a revised QoE curve results:

$$MOS^p(l) := 4.118 - 0.095 \cdot (32.758 - l) \quad . \quad (5.20)$$

The result has been obtained by applying a rescaling factor to the packet loss input l (as percentage) in order to make it compatible with $MOS^p(b)_{HD}$. This factor can be formulated to be

$$l := \left(32768 - pl \right) \cdot \frac{\zeta_l^{min}}{\zeta^{max}} \quad (5.21)$$

where ζ_{pl}^{min} is the corresponding minimum to ζ_{HD}^{min} at 4949.25kBit/s at which point the MOS is 1.0 (the considered minimum). The l function can be simplified to $6.61878 \cdot pl$ for the inspected case.

When using a linear price curve as in the *2012 trial* also for the case of packet loss impairments, the delta to $MOS(b)_{HD}^p$ can be characterised by two points. The first point is the common maximum, ie, a synchronisation point where demand and price are equal, the second point is formed at the mean QoE, ie, 2.56 of the considered range, ie, [1, 4.12], in analogy with the *SP* concept. Due to the different positioning of the MOS rating with 2.56 on the $MOS(b)_{HD}^p$ curve and its counterpart for packet loss variations, different price curves result:

$$p_{HD}(b) := 0.092 \cdot (-0.128 + b) \quad , \quad (5.22)$$

$$p_{HD}(l) := -2.979 + 0.183 \cdot l \quad , \quad (5.23)$$

where $p_{HD}(b)$ is a repetition of (5.13) and l is a positive packet loss metric. As in Section 5.3, the price curve is varied in order to keep the empirically measured demand curve static. $p_{HD}(l)$ is substantially steeper than $p_{HD}(b)$, which reflects the high demand sensitivity of packet loss impairments.

Conclusion #5.11. *The demand for network video quality is substantially more sensitive to packet loss increase than bitrate reductions (see price curves compensating the higher quality sensitivity as given in (5.22)).*

The suitability of this approach is hereinafter proven by applying a technique based on [Zwi+15]: Defining a flat theoretic reference curve θ with constant pricing that is capable of retrieving the full demand theoretically available, any deviation can be expressed by ASPW factors. ASPW represents the impairments introduced by reducing the bitrate, eg, for an HD video stream, or increasing the packet loss (which equals to a reduction of l). Thus, we can form

$$p(b)_{HD} \cdot ASPW_{HD} := \theta \quad \text{and} \quad (5.24)$$

$$p(l) \cdot ASPW_l := \theta \quad , \quad (5.25)$$

where $p(b)_{HD}$ is empirical measured, but $p(l)$ is unknown. However, from this we can infer

$$p(l) := p(b)_{HD} \cdot \frac{ASPW_{HD}}{ASPW_l} \quad . \quad (5.26)$$

Following the definitions given in Section 5.3.1.3, the $ASPW_l$ value can be defined by the ratio of the minimum point P_{min}^l 's Euclidean distance to the mean MOS point for $MOS(l)_I^P$ and the maximum point P_{max}^l (with P_{min}^l and P_{max}^l being placed where the MOS is 1.0 and 4.118 respectively). The factor $ASPW_{HD}$ can be formed analogously.

Proof. When defining

$$p_l \left(l(MOS = 2.56) \right)' = p_{HD} \left(b(MOS = 2.56) \right) \cdot \frac{ASPW_{HD}}{ASPW_l} \quad ,$$

around the mean MOS of 2.56, where $l(MOS = 2.56)$ and $symBW(MOS = 2.56)$ are the respective QoS demands at this point, and

$$p_l \left(l(MOS = 4.118) \right) = p_{HD} \left(l(MOS = 4.118) \right) \quad ,$$

around the maximum MOS at 4.118, we can derive that

$$p_l = p_l' = p_{HD}$$

has to hold. □

The function p_l' represents an unparameterised linear price curve for packet loss impairments (structurally comparable to p_l). This proves that a ASPW-based formation of $p_{HD}(b)$ is identical to the representation in (5.24). The rescaling factor of (5.26) can, thus, be used to reverse from p_l to p_b in (5.24). This further confirms the consistency of the model and its characterisation of QoS metric interdependences.

When using the beta distributed demand d for tariff B ($p_{max} = 3$) of the *2012 trial* (see (4.2)) a deeply nested hypergeometric utility function results (similar to the case of bitrate

variations), which cannot be handled by a linear optimiser. For this purpose, a linearisation has been applied, which leads to the following optimal curve:

$$\mathcal{U}_{lin}^n(l) := -0.070 \cdot (-4.949 + l) \quad , \quad (5.27)$$

which can be rescaled to $\mathcal{U}^n(l)'$,

$$\mathcal{U}_{lin}^n(l)' := 2.010 \cdot \left(-0.070(-4.949 + l) \right) \quad , \quad (5.28)$$

in order to satisfy the common maximum with $\mathcal{U}^n(b)$ at 32768 kBit/s.

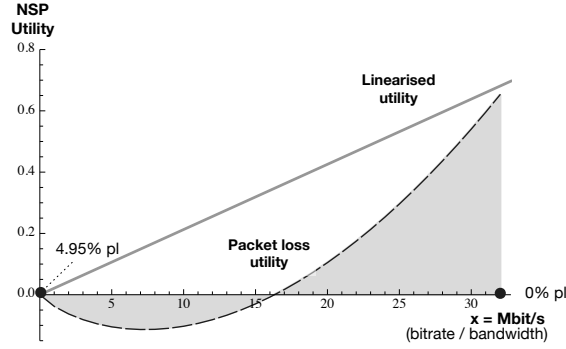


Figure 5.7: NSP utility under packet loss and corresponding linearisation.

The result is depicted in the perspective of a positive QoS metric (in analogy with bitrate or bandwidth demands) in Fig. 5.7.

Conclusion #5.12. *The utility converges to zero even for moderate packet loss levels. High packet losses may easily create a negative utility, which reflects the highly negative quality experience.*

Conclusion #5.13. *The NSP utility $\mathcal{U}^n$ is highly sensitive to packet losses and moderately sensitive to bitrate and associated bandwidth variations.*

5.3.3.2 Integration in Q-POINT

♣³² Utility data is important for the operation of technical optimisation. For this reason, the obtained utility data will be conceptually integrated in flow path optimisations. In particular, the $\mathcal{U}^u(l)$, $\mathcal{U}^n(l)'$, $\mathcal{U}^u(b)$ and $\mathcal{U}^n(b)$ (see result tables in Section D.1) can be integrated in Q-POINT (see Section 5.1.8.1 and [Dob+14]) as input information. While Q-POINT is capable of handling multiple traffic types, eg, video, audio and bulk traffic, the parameterisation derived from the *2012 trial* is only available for video. Due to scalability

³² The results, especially the numeric outcomes of the simulation runs, will be provided in an upcoming publication shortly. The material presented in this section arises from a collaboration with the Q-POINT authors.

reasons, the bandwidth request needs to be fixed upon starting the simulation. Thus, the expected purchases, which are derived from the winner determination output from Section 4.3.2, are taken as static input. The packet loss is, however, dynamically varied according to the network condition (ie, load levels).

Hence, Q-POINT has to read the b , $\mathcal{U}^u(b)$ and $\mathcal{U}^n(ul)$ values from the input. For the integration of the linearised $\mathcal{U}^u(l)$ and $\mathcal{U}^n(l)'$ data and in order to satisfy the revised objective on social welfare maximisation, the objective function has to be revised according to (5.18) (where the bandwidth/bitrate-related information does not affect the complexity of the simulation due to their static form).

5.3.4 Access Control & Resource Assignment

Partially proprietary resource assignment and access control mechanisms such as ADAPTISE [ISM14] can substitute the WTP multi-class access control sketched in Section 4.3. Contrary to the winner determination presented in this thesis, ADAPTISE focuses on a controlled degradation scenario. The quality may be degraded, ie, the initial request is not entirely satisfied, at the access control or the resource assignment may be dynamically altered whenever a reconsideration becomes necessary.

ADAPTISE's dynamic access control and resource assignment mechanism has been parameterised without empirical data. The empirical utility data given in Section 5.3 can easily be integrated by adapting the utility function in the ADAPTISE Java code. The multi-service type character of ADAPTISE can be reflected by using both HD and SD data. The mechanisms have to be aligned to the network flow path computation described in Section 5.3.3. Flow requests may, for example, be admitted when a suitable flow path can be found.

Due to proprietary character of parts of the ADAPTISE code, simulation results cannot be presented in this thesis. Nevertheless, a corresponding experiment is planned in collaboration with the ADAPTISE authors at UNIZG-FER.

5.4 Network Scheduler

Once the network sessions have been admitted, and network resources have been assigned, the practical enforcement is necessary. For the assurance of end-to-end quality levels, every network node, eg, a router, needs to enact dedicated quality policies. In the classical case, edge routers assign traffic to a certain QoS class, which is enacted by corresponding network scheduling policies. Corresponding to the findings of [WZR12] as revisited in Section 3.2, a QoE-aware scheduler requires a separation of charging quality classes, ie, charging for QoE classes, and provisioning classes, ie, QoS. When further integrating utility figures derived from WTP-findings, ie, creating a WTP-aware networks scheduler, a utility optimisation has to be executed before.

Hereinafter, a novel WTP- and QoE-aware network scheduler design (architecture, in-

interfaces and parameterisation) is introduced. Due to complexity reasons, a generic network scheduler design is targeted rather than a technology-specific variant, eg, a network scheduler specifically designed for LTE. Lower layer effects induced by the choice of a certain technology needs to be studied in future work.

The discussed approach will target assured service qualities, ie, quality guarantees, for premium traffic solutions, which are complemented by a defined (non-strict) share of BE traffic. This practice is in line with Ofcom’s interpretation of traffic management [OFC11], where a coexistence between BE Internet and other kinds of service is preferred, and the Universal Service Directive 2002/22/EC [Eur02] allowing the regulation of minimum quality standards for the Internet. While assured service qualities, “premium traffic” needs to be guaranteed, BE traffic may temporarily slip below a defined traffic share threshold in failures cases. The parameterisation and setup of the WTP-scheduler will deserve special attention in order to support a multi-class solution with strict quality guarantees (k -class solution where $k \geq 2$) in the premium segment and a single class solution in the BE segment. Quality guarantees, in accordance with Teitelbaum & Shalunov’s understanding or risk and QoS [TS03], provide an orthogonal temporal quality characteristic that can facilitate the QoE of end users.

5.4.1 Scheduling Techniques

Openly accessible emulators for known queuing disciplines especially exist for Linux, ie, *netem*³³ based traffic control and NISTNet³⁴ (not actively developed or maintained anymore), and BSD unix, ie, Alternate Queueing (ALTQ)³⁵ and *Dummynet* [CR10], provided via the ipfirewall (ipfw). Despite accuracy variances (see [NR09] for a detailed discussion) also depending on used kernel version, all solutions can be regarded to be stable, sufficiently tested, and accurate. Thus, the subsequent analysis will concentrate on assessing the functional requirements for creating the solution in Section 5.4.2 and how available tools match them.

Two main features are essential for the realisation of an efficient QoE scheduler: 1) a multi-class, “fair” and work-conserving queuing mechanism, and 2) cascaded queuing emulation. Multi-class refers to the usage of various traffic classes, eg, QoS classes, as used in Section 4.3.

5.4.1.1 Features

Work-conservation describes a characteristic of schedulers where the scheduler is never idle when it could serve waiting packets. Work-conservation is an essential prerequisite for an efficient and effective utilisation of available network capacity. Especially Weighted

³³ <http://www.linuxfoundation.org/collaborate/workgroups/networking/netem>, last accessed: 20th May 2016

³⁴ <http://www-x.antd.nist.gov/nistnet/>, last accessed: 20th May 2016

³⁵ <http://www.sonycl.co.jp/person/kjc/kjc/software.html>

Round Robin (WRR), WFQ [DKS90] and Deficit Round Robin (DRR) seem to provide “fair” concepts for redistributing resources from inactive queues among the available work-conserving mechanisms. Both WRR and WFQ use weights to control the volume of traffic being scheduled from individual queues. WRR, as among others utilised in *netem*’s Class Based Queuing (CBQ), uses the weights to schedule packets from active queues in a round robin approach. The resources from inactive queues are proportionately redistributed to all active queues, ie, satisfying the work-conservation property. Only as long as packet sizes are equal across queues, WRR thus provides a straightforward concept for “fairly” distributing bandwidth among traffic classes. Building on Generalized Process Sharing (GPS) [Kle76], WFQ advances over WRR by scheduling packets according their finishing times. WFQ thus considers the number of “rounds” (eg, a byte being scheduled) to finish packets in order fairly distribute bandwidth proportionately to each queue’s weight. Following this approach, WFQ still has to know the average packet size, which is addressed by the related DRR advancement. DRR [SV02] creates a deficit counter for each queue, which is increased by the assigned Quantum whenever a packet cannot be served in the current round. A packet can only be served if its size is smaller or equal than the assigned Quantum plus the deficit being carried from previous iterations. Thus, both WFQ and DRR provide suitable means for the fair work-conservation in networks (some “unfair” special cases may still be observed [Qua+00] though), while DRR may slightly be preferred due to its better adaption to varying packet sizes.

Contrary to WRR and WFQ, priority queuing (like a multitude of other class-based queuing disciplines) does not provide means for protecting minimum bandwidth requirements of each traffic class associated to individual queues, ie, the “fairness” condition among traffic classes is violated. Strict prioritisation may allow a prioritised queue to consume the entire bandwidth.

The Hierarchical Fair Service Curve (HFS)³⁶, implemented in Linux, optimises the scheduling for handling bandwidth- and delay-sensitive (real-time) traffic at the same time. The co-existence of traffic types is achieved by a higher tolerance to traffic bursts for real-time traffic classes in the short-term. On the downside, the underlying characteristic traffic curves are static, which disallows a direct optimisation using any collected empirical data. Furthermore, HFS seems to be rarely used and only weakly documented.

Conclusion #5.14. *Work-conservation, as required for an efficient network resource usage, is optimally provided by DRR and WFQ.*

The *Dummynet* and *ALTQ* implementations support WFQ, while *netem tc* comes with various realisations of CBQ plus HFS. These tools are the candidate platforms for implementing a WTP-aware scheduler. The *Dummynet* implementation provides cascaded WFQ-queues through defining rules reentering egress traffic back to the *ipfw* firewall, where the *Dummynet* mechanism is executed. *Dummynet* not only comes with an elegant rule definition syntax, but also has an active and supportive community. Analogously, *ALTQ*

³⁶ <http://linux-ip.net/articles/hfsc.en/>, last accessed: 20th May 2016

for Free BSD provides possibilities to redirect traffic outgoing from a certain queue (following a WFQ discipline) to another queue through the manual definition of rules.

Even more design control is provided by the popular *netem tc* implementation for Linux. *netem* is capable of hierarchically modelling traffic classes, mainly through the tool of its Hierarchical Token Bucket (HTB) queuing approach. HTB is broadly used in practice and has among others been utilised by BACCARDI IMS QoS Implementation (BIQINI) [Egg+11], a QoS enforcement tool for IMS. Although HTB is not necessarily work-conserving, it can hierarchically wrap the work-conserving CBQ discipline following a WRR approach. Moreover, with the DRR implementation also a more advanced work-conserving mechanisms is available, which is also fed by a tree-like traffic class setup and organised separately to the HTB structure. The traffic Quanta for DRR traffic classes have to be manually configured within *netem*, which may not represent a major drawback for our purposes. For the configuration of the scheduler *netem* has, hence, been used.

5.4.1.2 Capacity Dimensioning

Capacity dimension can be used as effective risk management tool, eg, in the sense of [TS03], in order to maximise the gain for the NSP, ie, high efficient network utilisation, without violating contractual agreements or unnecessarily affecting the long-term customer relationship. In the present context, BE Internet refers to a relaxed capacity dimensioning where a certain degree of dissatisfaction (localised, timely, etc.) is accepted. With quality guarantees, the certainty (low risk) of high quality is sold as a product. As a result, a considerate risk planning is required. Subsequently a few known works on capacity dimensioning with a focus on CO will be presented. This information is essential for properly parameterising the scheduler, hence defining the capacity that can be sold as quality-assured service and the rest that stays for BE use.

For a single link, the minimal CO capacity C_{min} for $M/G/\infty$ system (Poisson process) has been defined in [MMC05]:

$$C_{min} := \min_{C'} \{1 - \sum_{c(x) \leq C'} p(x) \leq p_v\} \quad (5.29)$$

$$c(x) := \sum_{i=0}^{k-1} c(r_i) \cdot x_i \quad (5.30)$$

$$p(x) := \prod_{i=0}^{k-1} \frac{\rho_i^{x_i}}{x_i!} e^{-p_i} \quad (5.31)$$

$$p_i := \frac{\lambda_i}{\mu_i} \quad , \quad (5.32)$$

where $p(x)$ is the violation probability (QoS and QoE is affected), p_v is the accepted violation probability, eg, 10^{-6} as in [MMC05], λ_i is arrival rate, $c(r_i)$ is request capacity

for type i , i is the indicator for considered traffic class, and x is the current state of the model. For details, we kindly refer to [MMC05].

According to Telkamp [Tel02] the extent of CO is a matter of the desired reliability. For maintaining 95%, 99% or 99.99% reliability, the CO extent may not need to be higher than 25%, 50% or 100% respectively. In an analogous manner, Juttner *et al.* [JSS03] recommend an effective CO by a factor of 2 – 3 (100 – 200%). The common ground can be found for a CO extent not below 100%.

Conclusion #5.15. *A CO extent of around 100% is required in order to assure a realistic reliability, ie, 99.99% [Tel02], for providing quality-assured network services.*

5.4.2 Solution Approach

The hereinafter presented solution approach will focus on two separate aspects: 1) the WTP-aware scheduler design and 2) the integration with known policy frameworks.

The solution approach will be *generically* applicable, as it will not be tailored to a particular technology such as LTE. The consideration of a given technology can lead to additional effects at low network layers, which are not assessed in this thesis. Contrary to Section 5.3.3 and Section 5.3.4, this section is not placed in the scene of a controlled degradation case. On that account, the data from Section 4.3 serve as main input for shaping and parameterising the mechanism.

5.4.2.1 Scheduler Parameterisation

A specific challenge lies in the mismatch between charged for and the provisioned quality classes, ie, a differentiation between QoE-centric charging and QoS as described in [WZR12]. For this reason, a truthful revelation³⁷ of the service type is necessary. Truthfulness can be assumed to exist in private networks as well as in public networks whenever a collaboration with trusted application and content providers exists. The traffic classification at the network edge will remain critical, as the network stability is bound to proper traffic management practices. The service type can be matched against a database with known QoS demands, eg, low latency for VoIP calls. Together with the billing information a proper classification of sessions to QoS classes is possible.

Whenever utility figures (eg, derived from WTP data) for the considered services are available, the scheduler can be entirely parameterised using this information. In a first step, a multi-class optimisation, as illustrated in Section 4.3, can be conducted for deriving the winning quality classes. In a second step, the result is transferred to DRR Quantum information to meet the demand for each quality class. This transition has to include the CO considerations of Section 5.4.1.2. In the present thesis, a CO factor of 100% will be used.

³⁷ Users or software tools may masquerade the service type in order to assign their traffic to better quality classes without additional payments.

Table 5.1: Multi-class example for the parameterisation of a network scheduler.

Tariff	Q_{q^*}	$\mathcal{R}^*$
Population: 1500k, $C = 1\text{Gbit/s}$ (25% BE, 75% premium, 100% CO)		
$p_{max} = 2$	$\{5, 7'\}$	222.95
$p_{max} = 3$	$\{6, 5', 3', (0')\}$	384.0
$p_{max} = 4$	$\{6, 5, 7', 4', (0')\}$	1459.79

For a 1Gbit/s link, where at least 25% remain for BE, and 1500 users for guaranteed premium services and applying the CO factor (100%), 37.5% remains as capacity for premium quality sales. When recreating the optimisation from Section 4.3 under these conditions, as shown in Table 5.1 for smoothed data for the extended model, the following quality classes and their demands can be derived³⁸ (see Table 5.2). The winning quality classes Q_{q^*} are ordered by their economic appeal (provisioned quality class I to V) according to the multi-class optimisation that has been presented in Section 4.3.

Table 5.2: Multi-class example for the parameterisation of a network scheduler.

	I	II	III	IV	V	
$p_{max} = 2$						
Q_{q^*} :	5	7'				BE
Sellable C (kBit/s):	104980	288152				262144
Required C (incl. CO, kBit/s):	209960	576305				262144
Quantum:	0.2	0.55				0.25
$p_{max} = 3$						
Q_{q^*} :	6	5'	3'	(0')		BE
Sellable C (kBit/s):	270336	122356	362	(128)		262144
Required C (incl. CO, kBit/s):	540672	244712	724	(256)		262144
Quantum:	0.516	0.233	0.00069	(0)		0.2502
$p_{max} = 4$						
Q_{q^*} :	6	5	7'	4'	(0')	BE
Sellable C (kBit/s):	43008	247608	101360	1024	(128)	262144
Required C (incl. CO, kBit/s):	86016	495216	202720	2048	(256)	262144
Quantum:	0.082	0.4723	0.19333	0.00195	(0)	0.25024

³⁸ The partial provisioning of quality class 0 indicates that a share of the capacity is returned to BE. Hence, a four class premium quality setup suffices.

The efficiency, as the ratio of sold to sellable traffic³⁹, of utilising the capacity dedicated to quality-guaranteed services is as follows: 99.98%, 99.67% and 99.95 for p_{max} is 2, 3 and 4 respectively. In other words, almost no idle resources remain when applying the multi-class optimisation. As witnessed in Section 4.3, the revenue also rises with p_{max} .

The scheduler uses the derived Quantums (cf. Table 5.2) as input parameters. The corresponding network traffic has to be marked accordingly in order to receive the guaranteed quality level. A guaranteed bandwidth assignment is required in order to protect individual usage sessions from each other. IMS or a similar kind of technology can be used for reserving the required bandwidth.

Conclusion #5.16. *BE requires assigned minimum resource guarantees, eg, a DRR Quantums of 25%, for avoiding a complete erosion of the BE Internet product segment at exceptional load levels. Any residual capacity, which cannot be sold effectively for a premium price, are returned to the BE segment.*

5.4.2.2 Generic Scheduler Design

Due to its broad adoption, its integration in the Linux kernel and its solid DRR implementation, we prefer the cascaded DRR-implementation provided by *netem* over its *ALTQ* and *Dummynet* alternatives. Whenever a direct pairing with the system firewall or the manipulation of both ingress and egress traffic is required, *Dummynet* is a promising alternative, due to its elegant syntax and direct integration in the lightweight Free BSD Operating System (OS).

netem mainly differentiates in *rules* and *filters*. Rules construct *qdiscs* (ie, classless pipes) and in the classful case also *classes* (ie, queues). Both may follow a certain kind of queuing discipline such as HTB or DRR and might specify specific shaping parameters. Classes provide more fine-grained control over different kinds of traffic (ie, traffic classes) by for instance supporting different weighting across sibling-classes. So-called filters filter the traffic and subsequently attach it to the correct class. In contrast to *Dummynet*, *ALTQ* and comparable tools, *netem* is only capable of shaping egress traffic, thus the filters trigger the traffic control after the *forwarding* has been determined, but before the traffic eventually leaves the node via the physical interfaces.

As depicted in Fig. 5.8, the proposed WTP scheduler⁴⁰, which receives dynamic input from resource optimisers, starts by setting up a bandwidth limitation through a single-class HTB setup. This design allows for testing multiple bandwidth scenarios on the same physical infrastructure. After that, two DRR phases follow where the first one concentrates on guaranteeing the required minimum quality standards for BE Internet traffic and any prioritised “premium” traffic for the failure case. The second DRR phase focuses on allowing a quality differentiation within the premium segment. For this purpose, an arbitrary number of quality classes (varying in their assigned bandwidth Quantum) can

³⁹ Q_0 capacity is automatically returned to BE, as no separate business case exists.

⁴⁰ The QoE scheduler has been prototypically implemented in Python, being executable on common Linux platforms.

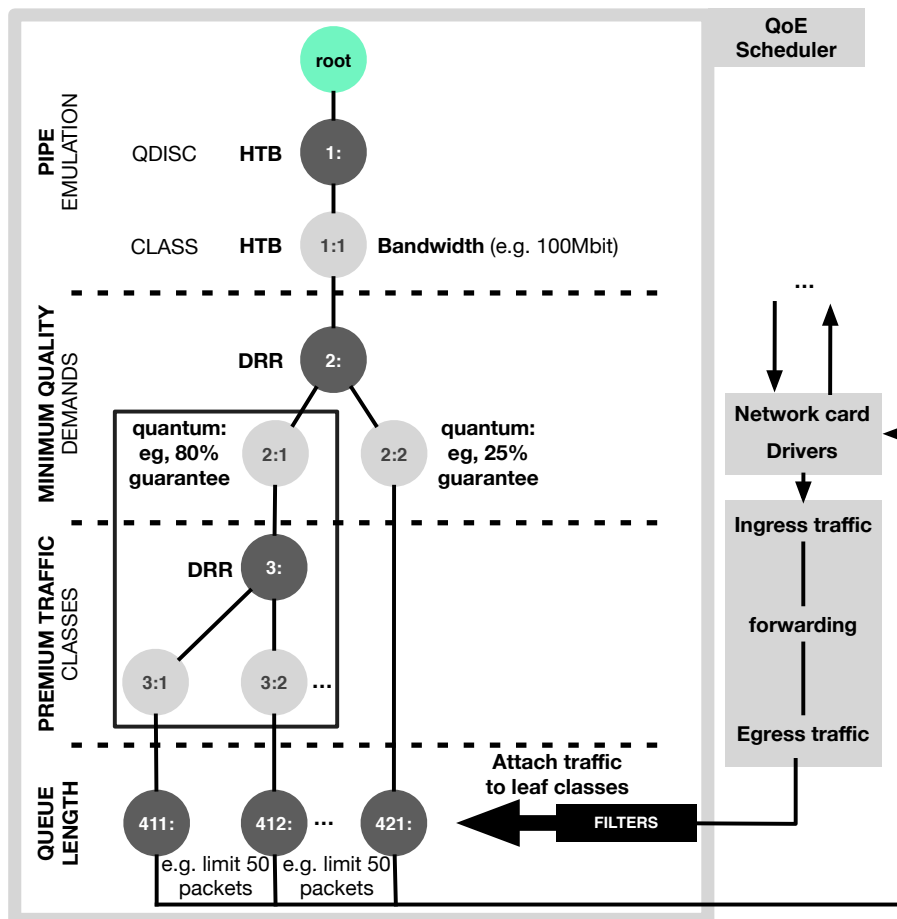


Figure 5.8: QoE scheduler traffic control approach

be created. The assigned Quantum for both kinds of traffic (and all nested quality classes) are derived from techno-economic multi-class optimisations and network capacity planning considerations. The latter incorporates the required capacity redundancy as preparation for failure cases (see Table 5.2). This step is assisted by an admission control process, which has to assure that all premium traffic can properly be handled in typical failure scenarios (see Section 5.3.4).

Conclusion #5.17. *Quality-assured network services need to be separated from basic connectivity demands in order to on the one hand protect BE Internet and on the other hand provide quality gains for “premium” customers.*

In a final step, the queue lengths are limited within *netem*. In the standard case, packets would be queued up to an infinite length, which may especially for real-time traffic be a serious quality concern. Thus, the present QoE scheduler provides means for meaningfully configuring these lengths for any kind of traffic. By default, the same queue length is used for each quality class, which have to be revised in a multi-service environment, eg, differentiating in bandwidth-intensive and delay-sensitive traffic. For a VoD use case, the current setup seems to suffice.

5.4.2.3 Link to Policy Frameworks

This section concentrates on the integration of QoE or WTP scheduling mechanisms with policy frameworks like IMS. Policy frameworks are required for associating particular user sessions to traffic classes. The presented view has been strongly influenced by the design of the open-source tool BIQINI [Egg+11]. Tools like BIQINI come into play for integrating a scheduler design, as sketched in Section 5.4.2.2, into policy mechanisms such as IMS. However, BIQINI does not satisfy a fair work conservation property and does not include WTP and QoE considerations. BIQINI has not been maintained for some years, and comparable toolsets are typically proprietary, and, thus not available for research purposes. The solution of this thesis, hence, concentrates on the scheduler design with *netem* and conceptually targets communication interfaces to IMS or other policy frameworks, as introduced hereinafter.

On the application- or service-specific layer, the billing for specific QoE classes is detected, and service-specific requirements are transmitted to a dedicated catalogue (see left communication interface in Fig. 5.9). The QoE billing feeds the PDP, which issues policies to be enacted by all PEPs, in our case any network node on the network layer. For the communication between the PDP and PEPs the Diameter protocol [RFC 6733] (right communication interface) is used in the IMS case. In our case, the PEP receives instructions on assigning specific service usages to a QoE class. The separate communications on service requirements allows the revised PEP to convert QoE demands to QoS parameters—following the approach sketched in [WZR12].

The presented design focuses on a limited number of QoS classes in order to keep the network complexity reasonable. A multi-class optimiser, eg, the WTP-aware component

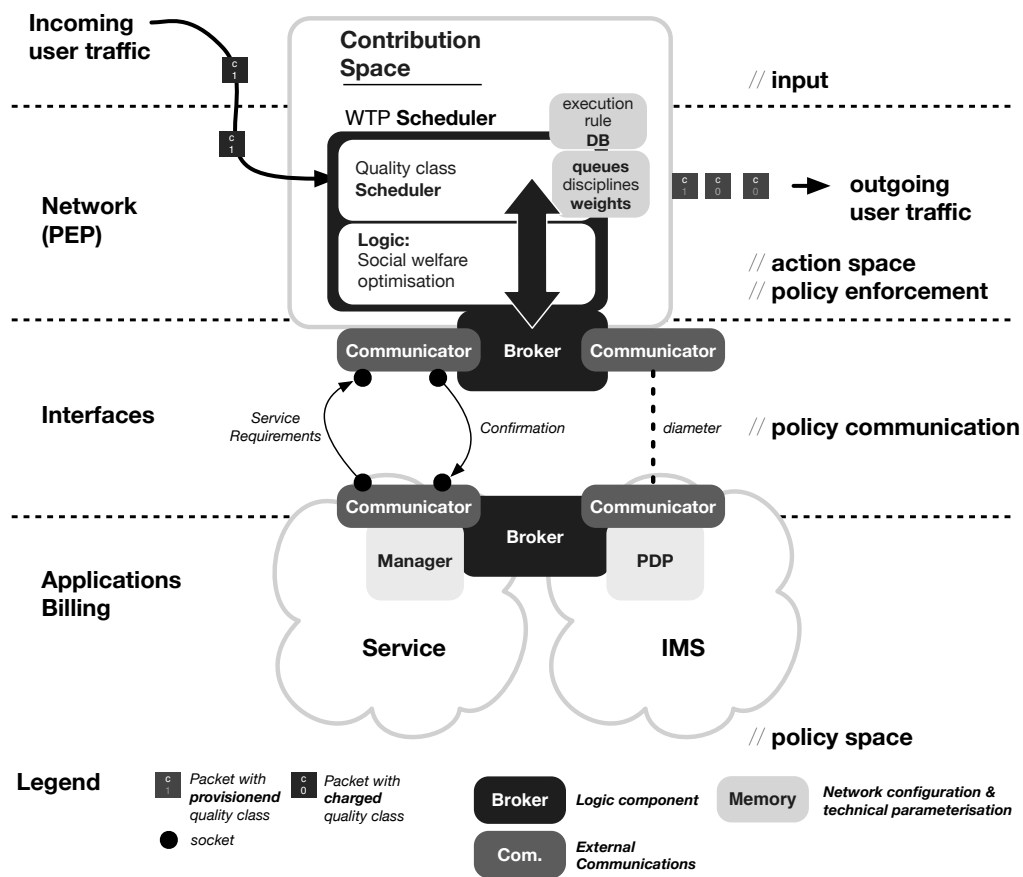


Figure 5.9: WTP-aware network scheduling with link to IMS

illustrated in Section 4.3, provides the essential input for the selection of winning quality classes relative to QoS needs, and assignment efficiency and effectivity. The classical PDP only receives winning quality classes and takes the responsibility of managing the enactment for specific users and usages.

The actual scheduling is done in an entirely revised PEP component, which contains the WTP network scheduler. The incoming user traffic is flagged according to the received QoS requirements and assigned to the appropriate queue. The charged for QoE class is not of relevance at this stage.

The optimality of both the experience and the profit need to be maintained throughout the network. Thus, the WTP-aware scheduler may need to run on every node in the network. Moreover, a linkage to QoE or WTP-aware access and routing mechanisms may be necessary and needs further exploration.

The service demand of an application is specified using a broker component. Typically, the NSP controls this broker. Untrustworthy information will likely automatically be ignored (especially on the receiving side) to maintain the financial and operational sanity.

5.5 Concluding Remarks

For the technical operations of networks and the optimal usage of available network capacity, utility information is crucial. Such information provides the basis for running efficient resource assignment and admission strategies, selecting the optimal flow path through a network and choosing promising customer segments.

The focus of this chapter was on the collection and application of the required utility data from the measurements in the *2012 trial* (cf. Chapter 3 and 4) in order to run social welfare optimisations both close to technology and users. Due to the limited available WTP data, approximation strategies are essential for meaningful simulations and optimisations. Classical QoE results (ie, user perspective of network quality) do not directly relate to purchases. Both user and NSP utilities, however, are subject to purchasing situations, which creates a substantial perspectival change. Thus, utilities and QoE are disparate concepts with an indirect mutual relationship.

This chapter illustrated that user utilities can be derived context-specifically from QoE ratings during a purchasing situation (involving price cognitions). WTP measurements can characterise NSP utilities. Both kinds of utility were approximated for the case of controlled degradations for video services, where the service preference of customers can be captured. This restricted case creates a sufficiently-specific customer context in order to expressively characterise inter-service utilities. The outcomes enable the modelling and simulation of the entire chain from network flow computation, access control and resource assignment, to lower network layers involving specific network schedulers can be modelled and simulated, as applied in this chapter. Due to the complexity of iteratively running end-to-end optimisations, simplifications such as linearisations are inherently required. A corresponding utility analysis around the *2012 study* results have led to the following not-

able insights: The obtained utility curves provide a useful source for optimising various kinds of traffic management tasks, which may affect the end-to-end resource assignment, traffic admission, go-to-market strategies, and traffic routing of future communications networks. When offering a choice of several quality-price bundles to customers, ie, a multi-choice scenario, extra NSP utility gains can be exploited over the case without choice. This supports the claim in [Sac+12b] according to which economically attractive niches could exist in the network quality market whenever customers can actively choose. Our results further demonstrate a high packet loss-sensitivity of both user and NSP utilities. Bitrate and the associated bandwidth demand reductions still moderately affect utility levels, but less sensitively than packet losses.

The conceptual integration with IMS illustrates the practical relevance of the collected utility information and the strategies applied upon it. IMS is used in modern telecommunication networks, such as LTE-based networks, for providing quality assurance for service types with specific needs—eg, the IMS-based VoLTE approach (see [IR.92]) in LTE supports the required low latency for audio services. While obtained data is conceptually compatible with the needs of modern usages of IMS technologies, a detailed analysis or implementation goes beyond the scope of this thesis.

Ongoing work has been further targeting the assessment of various network topologies to study controlled degradations using empirically-backed utility data and both Q-POINT and ADAPTISE simulators.

EVALUATION

♣¹ The findings of this thesis was derived primarily from empirical observations originating from dedicated laboratory trials. Empirical data was used to parameterise techno-economic optimisers as well as technical solutions. For the cross-validation of the associated outcomes, a new laboratory trial was conducted in Vienna, Austria in 2015 (later on referred to as *2015 trial*) and was rerun in Oulu, Finland (also in 2015). The trial was conducted at a new research lab at the University of Vienna (instead of reusing FTW's i:lab) and was operated by different personnel than prior trials. The first results of the Vienna and subsequently a brief summary of the Oulu results are presented hereinafter and related to the previous outcomes for validation purposes.

6.1 Demography

Twenty-two (22) test subjects have completed all stages of the experiment—while partly employees and students of the University of Vienna, a bigger share of entirely external subjects was tested. The number of test subjects exceeds the requirements of ITU-T's recommendation P9.10 [P.910] and BT.500 [BT.500-13]. Each test took ≈ 1.5 hours with three almost identical stages, thus a relatively large data pool, ie, 66 purchases and QoE ratings, was obtained for cross-validating especially the *2012 trial*, when considering the partial within-subjects test design of the later on described experiment.

Nine (9) out of 22 subjects ($\approx 41\%$) were female and 19 had graduated from a university (almost all subjects hold a master's degree or equivalent). The subjects belong to the following age groups: two (2) subjects are between 10 and 19 years old, ie, [10, 19], 11 subjects are in [20, 29], 6 subjects are in [30, 39], one (1) subject is in [40, 49] and 2 subjects are 50 years or older. Their experiences with VoD services are limited, ie, 11 subjects seldom purchase contents, one (1) subject does so weekly. Seven (7) subjects have one or

¹ The design of the laboratory study that is reported in this section has resulted from a research collaboration between VTT in Oulu and the Cooperative Systems (COSY) group of the University of Vienna. All trials have been conducted separately but in coordination. The results are currently in preparation to be published in direct content- and presentation-wise alignment to this section.

more subscriptions with a VoD service, eg, Netflix. Thus, mainly highly-educated young subjects with adequate gender balance were tested.

6.2 Technical Setup

The technical setup of the *2015 trial* follows the *2012 trial* structurally (see Fig. 3.4) with updated encoding techniques. The new test design uses Dynamic Adaptive Streaming over HTTP (DASH) [ISO/IEC 23009-1] video streams using one second segments and the High Efficiency Video Coding (HEVC) [ISO/IEC 23008-2; Sul+12] (also referred to as *h.265* [ITU-T H.265]).

In the 2015 experiment, the maximum bitrate was reduced by approximately 30% to 16384 kBit/s due to the codec advancements. The video quality was reduced from 17 to 8 bitrates, ie, $\{Q_0, \dots, Q_7\}$. The virtual quality classes, which test first-degree price discrimination, were not retested. The following bitrates were created from modern Blu-ray contents:

Table 6.1: 2015 trial: Quality levels Q_0 to Q_{17} in kBit/s.

Q_0	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7
128	256	512	1024	2048	4096	8192	16384

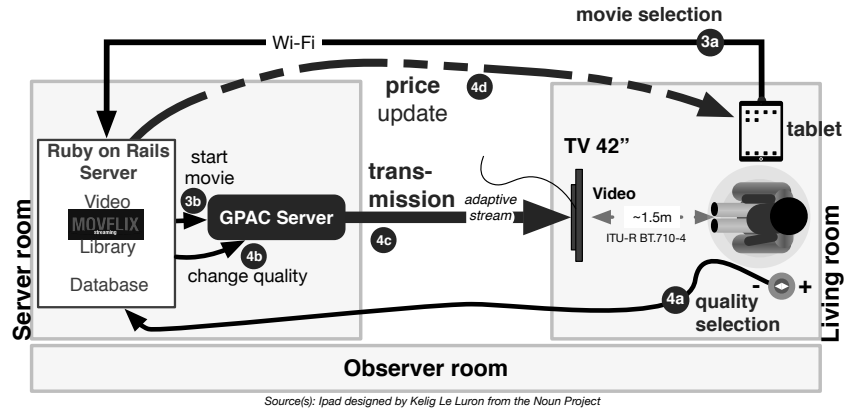


Figure 6.1: Technical setup (*2015 Trial*)

The test setup of the *2015 trial* (Fig. 3.4) used a more modern toolset (see Fig. 3.4):

- HyperText Transfer Protocol (HTTP) Live Streaming with VLC was replaced by DASH, which has been popular in research recently (eg, see [Sto11; El +13; OS12;

MLT12]), using GPAC². GPAC is capable of dynamically responding to DASH quality changes and provides more control over video quality switching. Some adaptations to the GPAC source code reduced the buffering and the associated switching time between quality classes. A Ruby script further moderated the quality switching according to the inputs of the jogwheel.

- For laboratory testing reasons, the content was streamed locally to reduce switching time between quality classes and to strengthen the controllability of the campaign
- A new paperless test suite was created with Ruby on Rails (RoR), which has been refactored, extended, documented and will be made openly available soon. The suite guided the test users through the entire trial. The subjects interacted with the system using a web interface on an iPad. The web interface showed a video quality rating (using ACR-based MOS scales and binary acceptance questions) and a content likability assessment page after each test round. The suite also provided the content selection itself.
- Subjects changed the quality (purchase video quality upgrades) by using a jogwheel identical to the one that had been used in the *2012 trial*. While the iPad displayed the balance and the price for the current quality, the quality class indicators were again intentionally hidden. The latter rendered the illusion of ‘endless’ quality levels and was used to direct attention towards perceiving the video quality and price values.
- Instead of a 40” flatscreen TV, a modern 42” Full HD display was used. The viewing distance, normed by [BT.710-4], was set to slightly above 1.5 meters (three times the height of the TV’s viewing area).
- An extensive set of video content, stretching several genres, was provided via a modern video library (with some similarity to video libraries of market leaders such as Netflix³)—cf. Fig. 6.2. The subjects were able to select the video material to their liking. Video materials were provided in English and German.

The new test design uses three tariffs, ie, A , B and C , with the respective maximum prices p_{max} of €2, €3 and €4. The prices follow a linear curve from €0 in each tariff for Q_0 to p_{max} at Q_7 :

² GPAC Multimedia Open Source Project: <http://gpac.wp.mines-telecom.fr/>, last accessed: 20th May 2016

³ www.netflix.com, last accessed: 20th May 2016

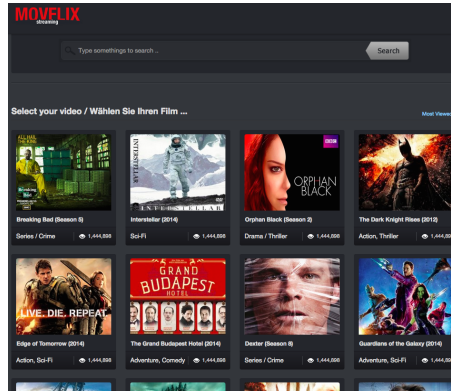


Figure 6.2: Movie library (2015 Trial)

$$\begin{aligned}
 A &:= [0, 2] \quad \dots \quad \{p_0 = 0, p_1 = 0.286, \dots, p_6 = 1.714, p_7 = 2\} \quad , \\
 B &:= [0, 3] \quad \dots \quad \{p_0 = 0, p_1 = 0.429, \dots, p_6 = 2.571, p_7 = 3\} \quad , \\
 C &:= [0, 4] \quad \dots \quad \{p_0 = 0, p_1 = 0.571, \dots, p_6 = 3.429, p_7 = 4\} \quad ,
 \end{aligned}$$

The number of test groups is reduced from three in the design of the *2012 trial* to two in the *2015 trial*—see Table 6.2. In the revised setup, increasing prices (ie, *Group I*) are compared to a control group with static pricing (ie, *Control*) in the first two measurements. In the third measurement, ie, t_3 , the control group is subdivided into two further groups in order to collect more data for tariffs *A* and *C*.

Table 6.2: User Groups.

	Description	Tariff			Sample
		t_1	t_2	t_3	
Group I	<i>Increasing prices</i>	A	B	C	11 Users
Control	<i>Constant prices</i>	B	B	A	6 Users
		B	B	C	5 Users

6.3 Test Routine

At the beginning of the trial, the users were briefly introduced to the lab environment. Each subject received a €10 deposit for video quality purchases, as in the *2012 trial*. Any money left on the deposit was paid out in cash after the trial, which was made clear

at the beginning. Test subjects were encouraged to try the system in order to make an educated decision and to perceive the tested experience product, ie, the network video quality. However, the test users were not encouraged to spend any money—they could freely decide on how to use their own money. Identically to the *2012 trial*, this, hence, represents an “own money experiment”.

The subjects then took the iPad and selected a language for the trial (German, English). Thereafter, the users signed a data policy declaration and received instructions regarding the used rating scales. In the next step, some demographic data was collected, and the measurement rounds began.

The trial consisted of three measurement rounds, where each has the following sequence:

1. The iPad shows some instructions.
2. The movie library opens up, and the subjects browse the available material.
3. Upon selection, a video (see (3a) in Fig. 3.4) with ≈ 20 minutes duration opens up on the TV (fullscreen) in the poorest video quality—(3b).
4. The start of the video initiates the QSP of 180 (three mins) seconds. Within the QSP, the user can select the video quality with the jogwheel (4a), which affects the video stream—see (4b) and (4c). The information for the QSP—remaining time, balance and price for the current quality selection—is updated on the iPad, see (4d).
5. When the QSP terminates, the NIP is initiated. The users can now focus entirely on the video now (no distraction is introduced).
6. Upon finishing the NIP (after ≈ 17 minutes), video quality ratings and video quality acceptance are assessed in a questionnaire on the iPad. Additionally, users are asked whether they have liked the content in order to allow likability assessments in the future.
7. The users are pointed back to the instruction page described in step 1.

After the completion of the three measurement rounds, a post-measurement questionnaire was triggered, which asked “Did it feel like spending your own money?” in order to understand the viability of the results better. A cognitive dissonance questionnaire [PS15], which was used in the *2012 trial*, was handed out to the subjects after the second measurement round.

6.4 Results

The pre-tests, including a series of full-length pilot tests with members of our research group, have concluded with the following outcomes:

- The *h.265* codec provides surprisingly good quality for low and medium bitrates, ie, Q_0 to Q_4 .
- Each quality class increase up to Q_4 provides a substantial subjective quality gain. From that point forward, the QoE gains are only perceived by a subset of the participants.
- While the highest quality, ie, Q_7 , provides a quality level slightly below Blu-ray, the difference was not recognised by most pre-test subjects.

6.4.1 WTP

In the main test, the high noise⁴ of the data (according to our observation and the results presented below) in measurement 1, ie, t_1 , suggests that the QSP was too short and the video material too heterogeneous for the chosen QSP duration. For this reason, the QSP was extended to 4 minutes during the conducting of the trial. Apart from this, a few videos with outlier results have been pulled from our marketplace during the experiment. Nevertheless, the noise in the t_1 data is still too high to present a clear analysis of market entrance effects with the obtained data. However, the careful trade-off management between price and quality benefits by subjects suggests a proper functioning of the *2015 trial* for providing more coarse-grained parameterisations. Subjects further liked the video content in 91% of all cases and rated the “own money” feeling with 2.9 (ACR-5) on average (median: 3.0; “Fairly”) in a post-measurement questionnaire. Some details of the *2015 trial* results are presented hereinafter.

The overall spending data (in € and relative to the respective p_{max} in %) are in line with the *2012 trial* (with slightly higher spending in 2015):

Table 6.3: Aggregate & per measurement t spending behaviour (*2015 trial*).

	Measurement			
	All	t_1	t_2	t_3
Median	€1.29 (43%)	€1.21 (52%)	€1.49 (57%)	€1.71 (43%)
Std. deviation	€0.87 26% ⁵	€0.73 25%	€0.76 25%	€1.07 27%

The histogram of the normalised spending (cf. Fig. 6.3a) and selected quality levels (cf. Fig. 6.3b) reveals a widespread interest in all quality levels and the existence of various customer segments. The majority selected intermediate qualities between Q_2 and

⁴ Noise in this context refers to the accumulation of outliers in this measurement compared to other measurements. The variance is in line with other measurements.

⁵ Percental std. deviation on the basis of normalised price data in $[0, 1]$.

Q_4 (where the maximum is at Q_2). The demand peaks at a normalised spending of 40% of p_{max} (ie, 0.4). Only one user did not spend any money at all during the trial, in analogy with the results in 2012, while 4 subjects have at some point spent the maximum. No subject consistently took the best quality, but several took only the two highest bitrates (Q_6 and Q_7). In the third iteration and for tariff C , the maximum p_{max} was never spent.

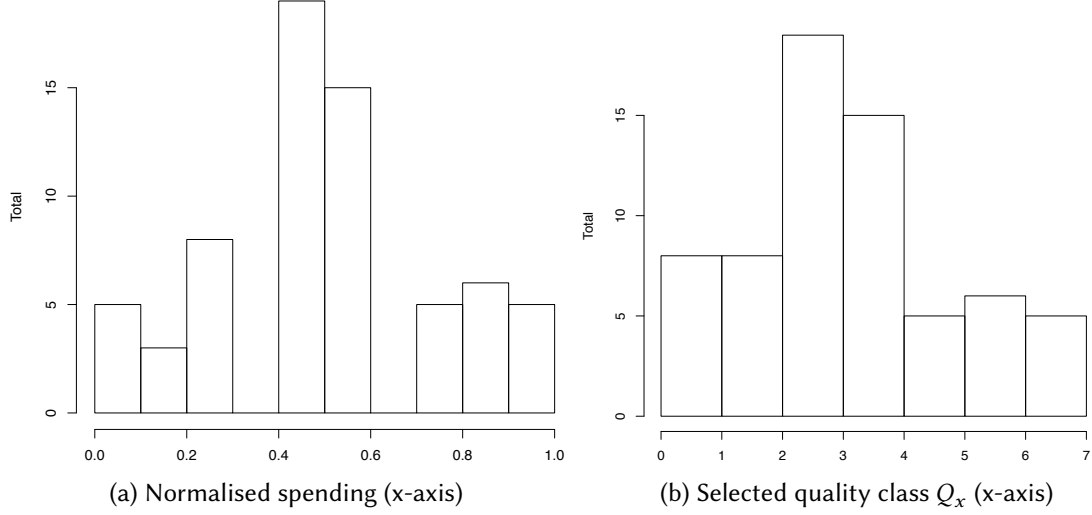


Figure 6.3: Histograms of purchases during the 2015 trial.

Table 6.4: Spending per Tariff (2015 trial).

	Tariff		
	A: $p_{max} = 2$	B: $p_{max} = 3$	C: $p_{max} = 4$
Median	€0.86 (43%)	€1.71 (57%)	€1.71 (43%)
Std. deviation	€0.43 22%	€0.80 27%	€1.16 29%

Some inconsistencies are revealed when reclustering the data per tariff A , B and C —cf. Table 6.4. The spending under tariff A was relatively low (43%) compared to the results obtained with B and C , especially in terms of the absolute expenditure of €0.86. As depicted in Table 6.3, the spending was low in t_1 despite the lower price on average (approximately two thirds had tariff A and the rest B), while t_2 (consistently using B) has obtained a higher spending on a normalised scale. The expenditure seems to be highly affected by the shy user interactions in t_1 . This confirms our observation of high data noise in t_1 and our recommendation to extend the QSP or schedule some training sessions. Training sessions should, however, avoid introducing additional biases, eg, by creating

reference points that could function as anchors for the main experiment, which creates a veritable design challenge.

Studying the normalised spending behaviour in t_1 and t_2 , a Pearson correlation of ≈ 0.42 (low correlation) results for both the *control group* and *group I*. The corresponding Spearman correlation⁶ for *group I* is 0.27 (no correlation) and 0.41 for the *control group* (low correlation). The normalised median of group I that moves from 0.43 (Q_3) to 0.71 (Q_5) and finally back to 0.43⁷ further reflects this correlation pattern. In the control group, the median is more stable with 0.57 (Q_4), 0.43 and 0.43, respectively. Between t_2 and t_3 , the corresponding Pearson correlation is 0.30 (Spearman: 0.22) in the *group I* and 0.42 (Spearman: 0.41) in the *control group*. The on average highest price in t_3 has lowered the relative spending and has led to a comparable absolute median spending of 1.71 in t_3 . Similar to the *2012 trial* the price increase has triggered a direct reaction—subjects chose lower quality levels—, but not at the a priori anticipated measurements t_1 and t_2 .

This effect is especially illustrative when comparing mean and median values for the subgroups of the control group over time⁸: In the subgroup with $B \rightarrow A$, the normalised spending remained constant, which relates to a lowered absolute spending. Subjects did not actively reconsider their t_2 choice in t_3 . In the subgroup with $B \rightarrow C$, the normalised spending has drastically decreased, which refers to an active decision. In absolute numbers, the mean spending under increasing prices in the control group's t_2 and t_3 dropped from 1.53 (median: 1.714) to 1.14 (1.1714). On the other hand, the absolute spending decreased proportionately to the price cut.

Table 6.5: Spending as percent of p_{max} in t_2 and t_3 per control subgroup.

	Measurement	
	t_2	t_3
	B: $p_{max} = 3$	A: $p_{max} = 2$
Mean	43%	45%
Median	43%	43%
	B: $p_{max} = 3$	C: $p_{max} = 4$
Mean	51%	29%
Median	57%	43%

When applying an ANOVA RM to the absolute and normalised spending data, no significant ($\alpha = 0.05$) time, group and group-time effects can be observed. Nevertheless this analysis illustrates that especially on normalised scale group effects are likely more

⁶ For a non-linear relationship of a data set that is not normally distributed, as likely in the presented case, the Spearman correlation may provide better coherence indications.

⁷ Semantically, the movement of the median may appear to be contradictory, as a price increase also raises the willingness to buy higher quality levels. This effect will be debated in detail hereinafter.

⁸ As shown in Table 3.6, one half of the control group was assigned to tariff A in t_3 and the other half had tariff B.

relevant than time effects. This stands in contrast to the comparison of t_2 and t_3 and is supported by the trends in Table 6.5. Despite the somehow promising results in t_2 and t_3 , our experiment was designed to illustrate group effects in a clean comparison between the *control group* and *group I* in t_2 (all subjects are on tariff B in t_2 , but have experienced other reference points in t_1). Hence, the sample sizes were too low for these subgroups to significantly verify or disprove market entrance pricing effects. Due to the high experimental noise in t_1 , this comparison cannot be repeated from the *2012 trial*. After the learning phase in t_1 , subjects have more consistently interacted with our system in subsequent measurements.

Conclusion #6.1. *In noisy test environments, eg, when subjects are unfamiliar with the purchasing system or the video content is too dissimilar for the chosen evaluation period (as is probably in our case), market entrance pricing effects can neither be confirmed nor disproved.*

Conclusion #6.2. *In the 2015 trial, subjects compensated average price increases with corresponding quality reductions.*

Conclusion #6.3. *In the 2015 trial, subjects did not actively responded to price cuts, which lowered the absolute spending proportionally to the applied price cuts.*

6.4.2 QoE

The QoE ratings—as measured in MOS on ACR-5 scale and illustrated as a box plot in Fig. 6.4 (mean values with errors)—are relatively insensitive between Q_2 and Q_6 , when considering the non-linear QoS increase from Q_0 to Q_7 . Q_0 represents an outlier due to the high dissatisfaction of users with the lowest quality level. The result for the best quality class Q_7 is noisy, due to the high price charged and the attraction of quality seekers that have partially indicated an interest in even better quality. Except for Q_0 , the logarithmic curve fit, as depicted in Fig. 6.4, provides a good approximation across all tariffs, which partly confirms the interpretation in [SS14] based on the *2012 trial* data. S-curve-shaped relationship of CS and WTP—where linearly increasing prices and, thus, WTP are captured on the x -axis—that has been anticipated by [HKH05], is not explicit in Fig. 6.4. However, a flattening of MOS ratings, as a particular CS representation, between Q_2 and Q_6 can be observed, which may have been induced by pricing. This effect is consistent with the results of our previous trials.

Clustering the data per tariff A , B , and C (cf. Fig. 6.5), the increased QoS-sensitivity with growing p_{max} stands out. With tariff A , the ratings for most quality classes range close to 4 (“Good”), which is notably higher than expected from classical video QoE trials. The high satisfaction of customers with their purchase probably causes this effect. In other words, the low price and active decision on a particular quality level and content seem to blandish the assessment of the actual quality experience. Only Q_1 (lowest quality class

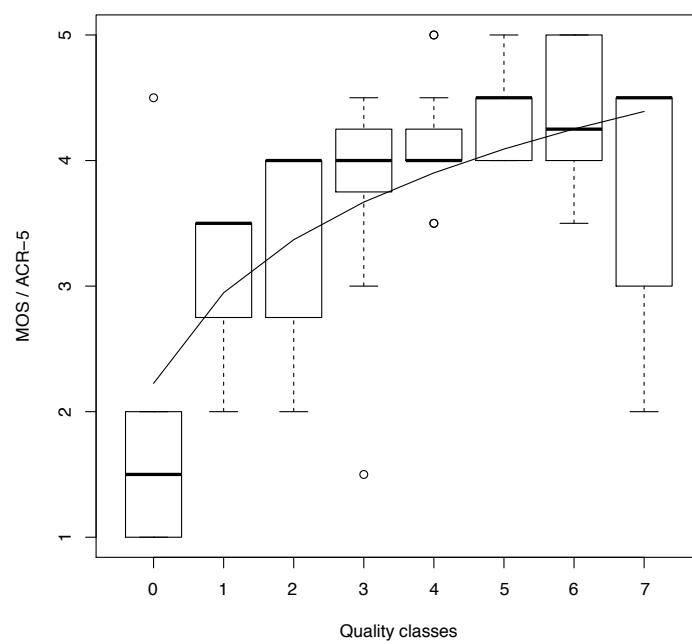


Figure 6.4: Box plot of MOS ACR-5 ratings across all tariffs with logarithmic fit (*2015 trial*).

chosen) and Q_7 represent outliers. With higher p_{max} the QoE ratings become more QoS-sensitive. Although superficially the results for tariff B and C may almost resemble classical QoE ratings, the MOS close to 4 between Q_1 and Q_4 still reflects a whitewashing effect in order to justify the purchase. The whitewashing distortion may be further alleviated when the price is raised above the tested p_{max} of €4.

Despite the likely relationship between quality blandishing and cognitive dissonance effects, as observed in previous trials, the used questionnaire has largely been assessed as too “emotional” by subjects, and has, hence, been inconclusive.

Conclusion #6.4. *Based on the 2015 trial data, it can be confirmed that prices and purchasing situations affect quality perception.*

Conclusion #6.5. *MOS ACR-5 ratings are less sensitive to QoS alterations, when subjects pay for quality upgrades. Both the 2012 trial and 2015 trial confirm this claim.*

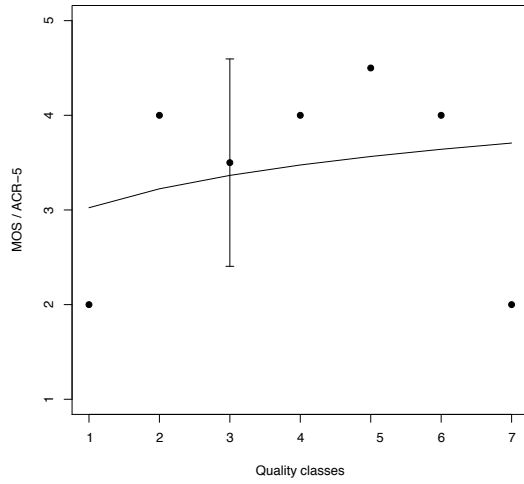
Conclusion #6.6. *While QoE ratings obtained in a purchasing situation, represented as MOS on ACR-5 scale, can be approximated using logarithmic curves, the quality of the fit is lower at QoS and price extrema than expected for QoE ratings without a purchase.*

As illustrated in Fig. 6.6, and in the details attached in Fig. C.2, acceptance has been high for wide ranges of the quality spectrum. Only the lowest quality levels Q_0 and Q_1 have not found acceptance, which seems to be caused by the particularly low bitrate of 128 kBit/s and 256 kBit/s respectively. In analogy with what was observed in the *2012 trial*, the price has sugarcoated quality acceptance especially in the lower to medium quality segments. The disagreements in Q_7 may relate to the wish of some subjects to purchase even better quality levels, as indicated in qualitative post-session surveys, or by the dissatisfaction with the price. Q_7 was not purchased with the most expensive tariff C ; hence the pricing could be around the absolute maximum the customers are willing to pay for network video quality.

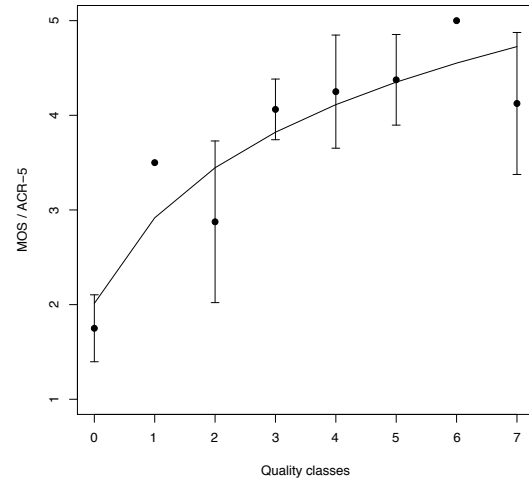
6.5 Recommendations

Upcoming trials, including the partnering trial in Oulu described in Section 6.6, should focally scrutinise group effects. Such group effects are relevant for understanding the market entrance parameterisation for novel experience goods. Based on our experiences with the *2012* and especially the *2015 trial*, we recommend the following alterations:

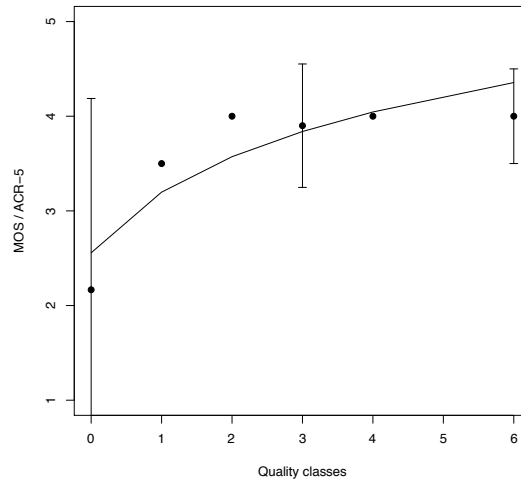
- Improvement of price visibility, eg, red for the price on the price indicator monitor (in our case, the iPad)
- Improvement of switching times between quality classes (in our case 2 – 3 seconds) in order to encourage experimentation with quality classes and to use the QSP optimally



(a) Tariff A, $p_{max} = 2$



(b) Tariff B, $p_{max} = 3$



(c) Tariff C, $p_{max} = 4$

Figure 6.5: MOS ACR-5 ratings per tariff with logarithmic fit (2015 trial).

- Extension of the QSP duration to 5 minutes or design of a non-parametric test session⁹
- Improved homogeneity and consistency of video materials, especially in the QSP

⁹ We define a non-parametric test session to be a dedicated round within the experiment where users can try the video adaptation and purchasing system without showing any prices nor revealing the available number of quality classes.

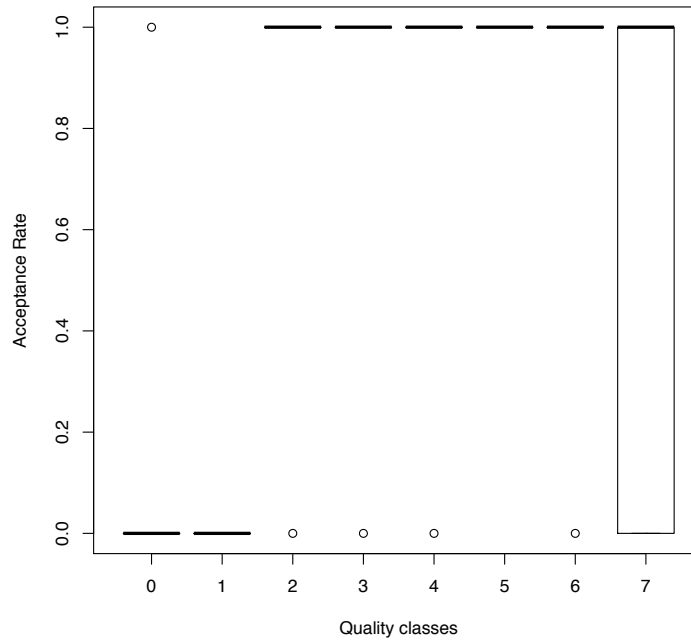


Figure 6.6: Acceptance rate across all tariffs (*2015 trial*).

6.6 International Cross-Validation: Oulu Trial

♣¹⁰ Following an almost identical test design and using identical tools, nineteen (19) subjects were tested in *Oulu trial* in 2015—primarily non-involved employees of VTT and students of a nearby university. In order to respond to the recommendations from the *2015 trial*, a few smaller modifications were applied:

- Clearer visualisation of the fee for the chosen quality level
- Removal of videos that have produced outlier results in the *2015 trial*
- Due to the unexpected high efficiency of the *h.265* codec, the maximum bitrate was reduced to 15000 kbit/s using a flatter increase from the minimum of 128 kbit/s (the 8 quality classes, and identical prices and tariffing structure have been retained from the *2015 trial*)
- Extension of the QSP from 3 – 4 to 5 minutes and reduction of the DASH segment size to one second in order to allow for a more dynamic and extended probing of the product offering

¹⁰ The results from the *Oulu trial* are preliminary, and despite the direct collaboration, the materials reported in this section are considered to be external works taken from VTT Oulu. For more details we kindly refer to our upcoming publication, which is produced in collaboration with VTT.

The retesting obtained a higher WTP than in the case of the Vienna campaign — the median rose from €1.29 to €1.71. While this effect could be subject to regional or sample size effects, the average increase of the WTP (22%) almost perfectly equals the QoS input difference between the trials (19% increase of the average bitrate). From this we can draw to interesting conclusions:

Conclusion #6.7. *The VoD and likely the network quality markets are of supraregional or even global character.*

Conclusion #6.8. *The offering of a larger set of “inadequate” quality levels—products that are unlikely purchased due to their unappealing quality—might entice customers to spend more money.*

Another difference was observed when considering the market entrance effects due to sequential tariff changes, which were acknowledged in the *2012 trial*, but also weakly in the *2015 trial*. In the *Oulu trial*, subjects have reacted rather passively to any kind of price changes, ie, once a selection was made it was likely retained over time. A possible explanation could be the even more prominent placement of the price information (a known inherent bias of all discussed WTP trials), which may have fostered the behaviour to look for the product with the identical price over all iterations. Cultural or sample size-related differences may also be at cause when considering campaigns with limited, despite satisfying all relevant recommendations, sample sizes.

Conclusion #6.9. *For future WTP trials, the inherent price bias of WTP trials deserves careful handling and moderation. Strategies that allow subjects to both keep an eye at the price levels and to perceive the network video quality might be beneficial (as in the 2012 trial).*

Based on an extensive literature review, Payne *et al.* [PBJ93] argued already in 1993 that decision-making is influenced by the range of attributes and their presentation. In our case, the importance of price and quality might be affected by the given number of quality classes (“choice”) and how prominently the price is presented to the customers. Comparing the Oulu with the Vienna trial, such effects might be plausible in the WTP context. Hence, more information is required from future work on how choice and presentation configurations may affect the purchasing behaviour of subjects in WTP trials.

Conclusion #6.10. *Market entrance pricing effects for network quality markets are likely, but deserve further testing in large-scale campaigns or with the help of meta analysis (if applicable). We further recommend a bigger number of quality classes as in the 2012 trial (or comparable strategies) in order to diffuse the range of available product offers and to centre the attention on the perceived product valuations. A low number of quality classes supports strategies to avoid active redecisions, which may lower the outcomes of campaigns.*

SUMMARY & CONCLUSIONS

7.1 Summary

The present thesis has addressed the telecommunications industry in its entirety. Despite its technology-centricity, telecommunications is a multi-disciplinary field today, for which analyses are required that use highly tailored methodologies from various Domains. This thesis has targeted one of the historic and most challenging problems of the industry: the sales of quality-differentiated network services. Attempts to establish network QoS have a dubious reputation today, which probably relates to the debatable success of associated solutions and research results. Nevertheless, in the last few years the success of QoE research has on the one hand given new life to the idea of quality-differentiated networks, but on the other hand has also brought back memories of the unsuccessful marketing of network qualities. Both QoE and the emerging NN discussions have further emphasised the need to approach this special breed of problem with multidisciplinary toolsets but have insufficiently aligned their perspective to market demands. This thesis takes up the balancing act between disciplines by integrating the viewpoints of users, markets, technology and society in order to design the more prosperous and utility-oriented telecommunications industry of tomorrow. For this purpose, dedicated but interlinked methodologies have been used that allow the assessment of specific sub-problems but also shed light on the fate of the entire quality-differentiated network market.

Network quality, like other experience goods, is not only hard to convey to customers but is also energised by its end-to-end character that integrates CP services with the services of various NSPs. Thus, a convincing storyline from promising market foundations via the needs of users to technical components is necessary, which has provided the common theme for this thesis. Originating from fundamental market designs, represented as VNs, empirical trials, economic models and optimisers have provided the required data in order to parameterise individual technical components. By especially concentrating on a VoD scenario, a cross-section through this enormous problem scope leads the way to further assessments. Most prominent results are reiterated in the subsequent section, while other interesting conclusions have been seamlessly integrated with every chapter.

7.2 Interpretation

On a high-level, the results of this thesis point to a clear potential to optimise the efficiency of telecommunications networks when transitioning from a BE Internet to a new service or market design (in response to the research question in Section 1.4). The introduction of a multi-class service design enables price and quality discrimination, which can substantially lower unused capacity while at the same time focusing on customer segments that, for example, yield the highest revenues. Who will profit from these efficiency gains is, however, subject to the chosen objective function and the market power of stakeholders. The market powers will vary from market to market where especially NSPs have the potential to absorb efficiency gains in markets with low competition (cf. Section 2.5). The associated revenue gains can be substantial (compare Table 4.1 with Table 4.2) when optimising for the NSP. Section 3.2 has further found efficiency gains when replacing QoS- by QoE-aware service provisioning. The extensive discussion in Chapter 3, however, also points to the need of replacing QoE by more market-oriented empirical data related to QoE, ie, WTP for network services. WTP data around a third-degree price discrimination, as in the *2011*, *2012* and *2015 trials*, are key parameters for a successful end-to-end multi-class network service provisioning that is utility- and quality-aware. The assessment in Chapter 5 further points to the multi-faceted technical realisation of such a network service design.

The detailed interpretation of the result of this thesis is oriented around the following *aspects: discrimination, perception, market niches, strategy, social compatibility, complexity* and *tools & data*. These *aspects* are fed by research results from the entire thesis and its heterogeneous methods.

Discrimination. From the current point of view, content and application discrimination, which involves so-called side payments, seem promising from neither a structural nor game-theoretic model assessment. Such kinds of discrimination may still be a lucrative market niche which can hardly be scaled to mainstream markets—due also to NN concerns and current regulations.

The present thesis has been able to confirm the functioning and practicability of customer-centric first- and third-degree price discrimination for network qualities (and other experience goods). First-degree price discrimination has been observed in the premium quality segment where customers have been seeking optimal quality. For some of these customers, the price was an authoritative quality indicator. Third-degree price discrimination has been illustrated especially in the *2012* and *2015 trials* where for different quality-price combinations diverging demand levels have been observed. Depending on the given choice, during the purchasing situation a beta distributed demand under third-degree price discrimination or a normally distributed association to customer segments appears plausible.

Perception. QoE and utility, whether NSP or user utility, are separate concepts with an indirect relationship. In Section 5.3, an extensive approximation process for the controlled degradation of video services has characterised this indirect relationship. Such services have the potential to increase social welfare through relating user utilities to efficient resource usage. The user utility for a given QoS level depends highly on the used service and the service preference of the user in question.

When contrasting QoE ratings during a purchasing situation with classical results, as given in Section 3.4 and 6.4.2, it becomes apparent that customers blandish their active choice for a product. In particular, the QoS-sensitivity of QoE ratings is lower for poor quality levels during purchasing situations. The purchase may also have led to a more critical handling of the highest quality segment where both the wish for even better qualities is aggregated and dissatisfaction with the pricing (despite the quality) may have been expressed. Thus, in analogy with the interaction of QoE and prices presented in [Rei+13a] (cf. Section 2.2), it can be concluded that the quality assessment is affected by pricing.

Despite the often witnessed systematic relationship between QoS and QoE, such as described in [FHT10] and [Rei+10], the evaluation of price information is highly relative and, hence, subject to any kind of anchor (reference point). In a study on market entrance pricing, the relativity of human valuation of goods has been illustrated, as historic price anchors have affected subsequent decisions.

Market niches. Active decisions can also be used to moderate the experience of users when confronted with network conditions below the targeted standard quality or expectations of the customer. Due to customers attempt to comply with their own active purchasing decision (ie, whitewashing effects), niche markets for very low quality and very low price offers may exist in equal measure to upmarket niches (see first-degree and third-degree price discrimination). Upmarket segments will attract customers that are willing to pay whatever it takes in order to get potentially a marginally better quality. While in today's market strategies the mainstream is focal, such niches may provide attractive additional sources of revenue.

Strategy. User, NSP and/or social welfare gains can be generated when migrating to discriminatory utility-aware network resource provisioning, especially when the appropriate prices have been set right at the market entrance.

While QoS and cost optimisation are focal in the backbone network (comprising core network and network IC parts of the end-to-end transmission), QoE and pricing become dominant wherever customer contact can be built up—mainly in access networks and also when CPs bundle the sales of their own service with network connectivity. Whenever a QoE perspective is applicable, the transition point where QoE demands are translated to QoS requests forms a critical role in the VN. The configuration of such transition points is, hence, critical for power distribution within an industry.

The positioning of a firm or role in the VN, especially when quantitatively understood, is critical. It provides valuable evidence on the future evolvement of the industry and the prospects of this firm or role. A case study has further illustrated the high economic pressure of Transit NSPs in today's telecommunications industry. A change of Transit NSPs business models is likely in the near future, and generally structurally advised.

IC agreements are essential in today's globalised networks. This is the same for quality-assured network services where guarantees have to be maintained end-to-end and across Domain borders. Available IC approaches, with ETICS leading the way, are SLA-based and typically highly complex. The integration of access network aspects further creates a parameterisation problem, as user-centric figures such as QoE ratings are essentially required for creating the expected service experience. The agreements between NSPs are QoS-, technology-, resource- and cost-centric. The transition to a human perspective requires a tool for conveying and selling network qualities to customers—see [Var+15].

Network offloading (eg, to Wi-Fi using the unlicensed spectrum) can be a dominant strategy, especially when the QoE of customers is not affected (cf. [Zwi+13]). This typically refers to bulk traffic with low QoS-sensitivity. The possession of both a fixed line backbone network (plus Wi-Fi) and an exclusive cellular spectrum can be a strategic advantage for NSPs, which can be used for generally lowering the Wi-Fi quality in order to raise the relative quality gain for customers in their exclusive cellular spectrum. In other words, the exclusive spectrum can become a critical asset for reliable or higher quality services, while unlicensed spectrum solutions will be the carrying technology for services with lower quality demands.

Social Compatibility. A fundamental QoS- and QoE-pricing model (cf. Section 3.2; [WZR12]) has illustrated the economic value of utilising QoE-aware pricing over classical alternatives. Due to the exploitation of human capability to perceive quality differences, originating from QoE as quality sensing metric, moral questions may be raised. On the other hand, Wahlmueller *et al.* [WZR12] have illustrated the increased inclusion of customers with very low purchasing power and/or WTP. Thus, it is fair to argue that a properly and theoretically optimal discriminatory QoE-aware system could be beneficial for society and the network infrastructure at the same time. Efficiency gains due to a more tailored service provisioning could provide a productivity advancement that positively affects the entire industry. However, the possibility of using greedy objective functions may substantially alleviate positive effects. Greedy strategies may profit from the difficulty of valuating services, especially when cost levels are often covert. Anchoring effects, as witnessed in Section 3.6, may be used for increasing social welfare or for exploiting consumers. According to our trials, historic pricing may be such an anchor that could mislead consumers to retain their WTP even in the case of falling provisioning costs. Further, QoE- or utility-aware pricing might not comply with common NN viewpoints, regulations or practices (also see comment on side payments above). Of course, regulatory intervention or high competition can also limit the feasibility of greedy market strategies. Thus, the role of

regulatory authorities in quality-aware networks can be regarded to be of even higher importance than for today's BE Internet.

Complexity. Proper quality differentiation is highly complex and social hurdles are high. Additional design-time complexity, ie, a parameterisation complexity, is generated when a QoE- and/or utility-awareness is introduced—cf. Chapter 5. This change requires empirical data backing, which only exists practically for selected services and scenarios. An extensive data approximation as illustrated in Section 5.3 for a controlled degradation case still requires substantial data input. Runtime complexity is generated in the dynamic winner determination phase of utility-aware networks where the assignment of resources to the most rewarding customers or customer segments is critical and challenging. The end-to-end optimisation approach, which requires a consideration of the entire flow path, further raises the complexity, which has been addressed in this thesis by model linearisation techniques. A purely technical complexity lies in the provisioning with today's QoS toolsets such as IMS, as used for VoLTE in LTE networks, where the tool complexity rubs of on the end-to-end resource assignment optimisation for quality-differentiated networks.

Tools & Data. In the course of conducting the experiments of this thesis, tools have been created and research data have been collected. Due to Intellectual property (IPR) restrictions, data can only partially be released publicly. The following items have been made available in our GitHub repository at <https://github.com/pzwickl/PZ-Thesis/>:

- The VNQ toolset, as presented in Section 2.4 and which is not limited to the telecommunications context;
- The exported data from the *2015 trial* and the associated analysis script for R^1 ;
- The Mathematica² script for calculating the utility figures presented in this thesis.

The laboratory tools that have been used to conduct the *2015 trial* are currently revised and extended. The polished RoR project is planned to be released as Open Source in the upcoming months.

7.3 Outlook & Future Work

The characterisation of the overall design of the telecommunications industry can profit from an extended VN quantification, as described in Section 2.4, for more cases. With this technique, creative new business ideas may be tested and compared to each other. Such

¹ The R Project for Statistical Computing: <https://www.r-project.org/>, last accessed: 20th May 2016

² Wolfram Mathematica: <http://www.wolfram.com/mathematica/>, last accessed: 20th May 2016

an assessment requires additional market data sources which have not been accessible for this thesis.

Economic niches deserve further exploration. In particular, their detailed realisation within NN bounds³ remains for future work.

The thesis intensively relies on empirical parameterisation. Additional confirmation of the used data is planned. The most concrete plan is the detailed analysis of both *2015 trials* (Vienna and Oulu). Untested services remain for future work, where especially WTP, demand and QoE figures involving pricing are recommended for testing.

The communication of network quality as an experience good is critical, but insufficiently targeted to date—see [Var+15]. From the perspective of this thesis, strategies to convey QoE to end customers is critical in order to link the QoS perspective of interconnected core networks to a QoE viewpoint of consumers.

The access and resource assignment as well as flow path computation will be followed up by technical simulations using the outcomes of this thesis. An extended Q-POINT simulation for the utility-aware flow path computation is scheduled for 2016.

³ During the thesis, the NN regulation situation repeatedly changed drastically. When the regulations appear to be sufficiently settled, fine-grained response strategies can be worked out.

VALUE NETWORK QUANTIFICATION (VNQ)

A.1 Inequality Measures

(In-)equality measures are required for two main VNQ purposes: first, bargaining powers (cf. $\delta_{s/c}$) require an understanding of the value distribution among business interactions (relationships) of an entity. The more distributed the value generation is, the higher is the bargaining power of this entity. Second, from a more fine-grained perspective also internal competition for each relationship needs to be characterised. With an increase of competition the internal pressure on relationship instances (firms competing for the exchange of resources) is raised, eg, an Transit NSP facing a series of similar priced offers may render price increases more unlikely. The present section, thus, reviews promising inequality measures for VNQ from the ground up by especially concentrating on the comparison with the used Gini coefficient—see [ZRG11; ZR12; ZR13].

Due to the seemingly arbitrariness of (in-)equality measures—eg, see [Pal10]—under superficial inspection; the yielded results may drastically vary from one metric to another—and the absence of clear recommendations in literature, a selection of the most promising quantification approach is particularly difficult. For this reason, subsequently an inspection of most common inequality measures from economics and information theory are reviewed, such as the Gini coefficient [RS04] and several entropy-based [Sha48] alternatives. This assessment is further complemented by voting power metrics. The results have been applied to Section 2.4. Without a systematic characterisation of their interplay, we further do not recommend to intermingle various metrics that measure inequalities, such as the Gini coefficient or the Theil Index.

From [Pal10] a few desired a-priori characteristics of metrics (some do not apply to the context, and have thus been disregarded) can be derived:

- **Axiom of Anonymity:** Ordering of values, eg, incomes, does not influence the

outcome of the metric. This metric is satisfied by all subsequently discussed mechanisms.

- **Decomposability:** Possibility to decompose the assessment in within and between group results. Decomposability leads to a better analysis of inequality sources and subproblems.
- **Bounds:** The results should range in clearly defined bounds. The existence of clear bounds allows an easier reasoning and relative comparison, which appears to be beneficial in the context of VNQ.

Further context-specific characteristics, such as the “entity size” as an orthogonal factor, also contribute to the tailoring of an adequate inequality measurement approach.

A.1.1 Concepts

While in [ZRG11; ZR12; ZR13] the Gini coefficient was used as measure for distribution of values among relationship but also among competing alternatives, good alternatives seem to be available. One of the very promising replacements is the Theil index—as derived from Shannon’s [Sha48] information entropy—which has gained in popularity for income and wealth distribution. Due to the important role of entropy and information gain in physics, electrical engineering and computer science, especially with respect to machine learning, the Theil index will be used as starting point of the analysis to identify the best suiting inequality measure for VNs.

A.1.1.1 Features & Calculation

Like the Gini coefficient the Theil index (esp. $T_{\alpha=0/1}$) can be used as Generalized entropy index, but has some noteworthy advantages:

- Several variants of the Theil index exist. Most notable are T_L and $T_{L'}$ (see definition below) that support the overweighting of the worst or the best alternatives in the case of *negated* and *unnegated* utilities. This modification renders the definition of a spreading factor obsolete.
- As illustrated in [WL03], the Theil index allows for decomposability in a way allowing *within* and *between* subject measurements. This could be used for assessing particular segments, for example.
- Weighted Theil indices may further be used in order to isolate contributing factors (from subgroups) while respecting the size of a subgroup (relating to decomposability). This can be used in this work to integrate the entity size $S_{in/out}$ directly into the measure for $\delta_{s/c}$, which eliminates the required weighting factors.

In order to computational achieve such a transition, the following series of definitions is required in order to fit the Theil index variations, starting with Theil-T T_T , to our particular usage. Theil-T T_T can be defined by

$$T_T := \frac{1}{n_r} \sum_{i=1}^{n_r} \frac{x_i}{x} \cdot \ln \frac{x_i}{x} \quad , \quad (\text{A.1})$$

where x is the arithmetic mean of all relationship utilities x_i with n_r being the number of available relationships $r \in R_{in/out}^e$. In order to overweight the best alternatives out of negative utilities¹(eg, in order to avoid spreading factors for the utility calculation), the known Theil-L T_L (using Mean Log Deviation (MLD)) is slightly modified,

$$T_L := \frac{1}{n_r} \sum_{i=1}^{n_r} \ln \frac{-x}{-x_i} \quad , \quad (\text{A.2})$$

by changing the arithmetic sign of x , which strengthens the influence of any resulting low positive utilities.

Contrary for positive utility values, ie, profit maximisation for alternatives, the modified $T_{L'}$,

$$T_{L'} := \frac{1}{n_r} \sum_{i=1}^{n_r} \ln \frac{y}{y_i} \quad , \quad (\text{A.3})$$

is defined where $y_i = x_i^2$ and y in correspondence to x represents the arithmetic mean of the y_i values. Thus, highest utilities are overweighted, as low utility alternatives are of lower interest.

At this point all Theil indices range from 0 (equality of relationships; maximum disorder) to $\ln(n_r)$ (highest heterogeneity), which can be normalised in $[0, 1]$ by a multiplication with $\frac{1}{\ln(n_r)}$. This represents a sufficient factor for the calculation relationship instance utilities, while in the case of bargaining powers (being at the same computational level as the entity size force) $\delta_{s/c}$ a weighting according to entity sizes may allow the elimination of an explicit factor and its weighting factor. Thus, by utilising Weighted Theil indices in order to reflect heterogeneity in the VNs, Equations (A.2) and (A.3) can further be reformulated to form the Weighted Theil WT index,

$$WT^{VN} := T \cdot S_{\{s/c\}}^e \quad , \quad (\text{A.4})$$

where $S_{in/out}^e$ represents the relative share of incoming (supplier) or outgoing (customer) relationship utilities for entity e (being scaled in $[0, 1]$) where $\sum_{e \in E} S_{in/out}^e$ is 1 (for all entities $e \in E$). Consequently, an entity heavily doing business with a limited number of other entities, may even be more dependent than considerably smaller entities. In practice, WT may neither be exactly 0 nor 1 for any of the entities, thus a final normalisation for

¹ Emphasis on cost-minimal solutions

comparability of dependency indicators for entity e is necessary. The resulting alternative form $\delta'_{s/c}$ for $\delta_{s/c}$ is formulated as follows:

$$\delta'_{s/c} := \frac{WT^e}{\max_{k \in E} WT^k} \quad . \quad (\text{A.5})$$

A.1.1.2 Drawbacks

While the modified Weighted Theil index streamlines the computation by removing two previously required weighting factors and by replacing a spreading factor, we have to acknowledge newly introduced representation problems: whereas the Gini coefficient ‘runs’ against a natural bound, ie, it ranges in the interval $[0, 1]$, this does not apply to the Theil index (whether T_T , T_L or combinations of above; also see [Pal10]). This may require, hence, some efforts in rescaling for easier comparison.

Despite Theil Index’s capability—originating from the power of information entropy [Sha48]—of measuring disorder or uncertainty (a qualitative description of distribution), the orthogonal factor of *choice*—what we may define as outcome impact of the number of alternatives n_r —is disregarded. In other words, the quantity of offered alternatives does not necessarily affect the inequality metric. The Gini coefficient considers both quantitative and qualitative aspects at the same time. Applying the Theil Index to a large enough set (of values, alternatives, etc.), optimally converging to infinity, the order may approximate the summation of all authoritative characteristics (down to infinitely small errors). Thus, for the measurement of income inequality of million inhabitants or even continents the absence of *choice* may be regarded to be irrelevant and may, hence, possess limited authority. For the quantification of VNs (or business interaction in general) the probability of quantitatively assessing a very small set of alternatives may be high though. In turn, the applicability of the Theil Index, despite its expression power across disciplines and the yielded calculatory simplifications, may without modification be inapplicable to the context of VNQ. Turning back to Gini coefficient (a comparable measure) *choice* is sufficiently captured.

Referring back to the mentioned measurement differences across metrics, the intermingling of ‘nice-to-have’ features seems to be unfavourable. Nevertheless, the subsequent subsection will aim transferring desired characteristics to the Gini coefficient.

A.1.1.3 Voting Power

There also exists a series of voting power indices such as Shapley–Shubik or the Banzhaf power index [GKT02], which allow the assessment of the influence of one vote on the outcome (coming from a scenario of shareholder voting where one shareholder may hold substantially more votes than others), ie, whether a particular vote would be decisive or not. While this concept may illustratively express the power of one voter (thus, in our context the power of a relationship instance) in the overall panel (set of relationships and

their utilities), it does not give a direct indication on the distribution of powers. In turn, another inequality measure is needed to assess the distribution of voting powers among the set of players, eg, shareholders.

A.1.1.4 Interpretation

Systematically reviewing other alternatives presented in [Pal10], it becomes apparent that decomposable metrics cannot capture *choice*. However, with substantial effort the Gini coefficient, capturing *choice*, can be made decomposable: according to [STY86] and adequate tooling², the Gini coefficient can be made decomposable in the form presented by Lerman and Yitzhaki [LY85], when a statistically validated linkage of subgroups can be established—see the adaptations presented in Section A.1.2.1.

To the best of our knowledge, neither advanced nor more sophisticated decomposition techniques are known for alternative metric such as the Krugman index [Pal10] or the Index of Inequality in Productive Structure [Pal10], but may be constructed in the future (especially due to the simplicity of the quantification). The well-known Atkinson index [Atk70], however, is decomposable, but as such does not support the representation of *choice*. However, it would again provide the possibility of continuously overweighting lower or higher values in the set, ie, providing a replacement for a spanning factor capturing the utility distribution of alternatives. Please take further note that per subgroup assessment, eg, through IBM’s *IntelligentMiner* serving the tree-based data splitting³, is sufficiently targeted in literature.

A.1.2 Lessons Learned

The present section aims at transferring advantageous features of alternative inequality measures required for VNQ to a revised Gini coefficient calculation.

A.1.2.1 Decomposition

In correspondence to the notation of Lerman and Yitzhaki [LY85] and the Gini extension described in [STY86],

$$G := \sum_{e \in E} \underbrace{\frac{\text{cov}(y^e, F)}{\text{cov}(y^e, F^e)}}_{RC^e} \cdot \underbrace{\frac{2 \cdot \text{cov}(y^e, F^e)}{m^e}}_{G^e} \cdot \underbrace{\frac{m^e}{m}}_{S^e} \quad (\text{A.6})$$

—in continuous form for a uniform value distribution F in $[0, 1]$ —, we will extend VNQ’s discrete Gini calculation (using the notion in [Bre+84]) such that it satisfies decomposability. Decomposability is required as technique for capturing *entity sizes* in analogy with

² López-Feldman, A., “Decomposition of the Gini coefficient using Stata”, http://www.stata.com/meeting/mexico09/mex09sug_alf.pdf, last accessed: 20th May 2016

³ Han, J. and Kamber M., “Classification and Prediction”, <http://www.itu.dk/people/pagh/ADBT06/classification.pdf>, last accessed: 20th May 2016

WT. G is the aggregated Gini coefficient of several subgroups and G^e represents the Gini coefficient for entity $e \in E$ —as previously introduced in (2.6). The subsequent elaboration constructs a relationship between G^e and the value G its enclosing group. The results directly affect the calculation of $\delta_{s/c}$, mainly through the weighting based on the factor S^e .

$$S_{\{s/c\}}^e := \frac{\sum_{r \in R_{\{in/out\}}^e} \mathcal{U}(r)}{\sum_{k \in R_{\{in/out\}}} U(k)} \quad (\text{A.7})$$

S^e is the share m^e (see numerator in (A.7)) of entity e 's utility from the utility m of all entities (comparable to WT; see divisor in (A.7))⁴ and RC^e is a new factor for the “Gini correlation” [LY85]. The latter is required in order to circumvent the non-decomposability of the classical Gini coefficient.

Rewriting RC^e , a so-called correlation factor, from (A.6) to fit the calculation of the used Gini coefficient from information theory [Bre+84] we obtain

$$RC^e := \frac{\frac{G}{2} \cdot m}{\frac{G^e}{2} \cdot m^e} \quad , \quad (\text{A.8})$$

which can be simplified again to G when inserted in (A.6), ie, proofing the equality⁵. In contrast to the notation of RC being given in [LY85] ranging from $[-1, 1]$, here it conveniently scales in $\mathbb{R}^+$.

Proof.

$$\begin{aligned} G &= RC^e \cdot G^e \cdot S^e \\ G &= \frac{\frac{G}{2} \cdot m}{\frac{G^e}{2} \cdot m^e} \cdot G^e \cdot \frac{m^e}{m} && \left| \text{(A.7) and (A.8)} \right. \\ G &= \frac{G}{G^e} \cdot G^e && \left| \text{cancellation} \right. \\ G &= G \end{aligned}$$

□

This relationship proves that from the coherent product $G^e \cdot S^e$ the value of G (representing the Gini value for the entire VN) is deducible with the help of RC^e without knowing

⁴ Due to a uniform and continuous income distribution in $[0, 1]$ the definition in [LY85] is alternating.

⁵ Similar, to (A.6) only k 's proportional share of G can be calculated by replacing G with $G/|E|$ where $|E|$ is the number of entities in the VN

any other $RC^{e'}$ or $G^{e'}$ for any $e, e' \in E$ and $e \neq e'$. Transitively, we can infer G^e for any entity e from G ,

$$G^e := G/S^e/RC^e, \quad e \in E \quad , \quad (\text{A.9})$$

by considering how much an entity e contributes to the overall inequality in the VN where inequality refers to a higher dependency or a lower market power. A high contribution of e to the inequality solely due to its high S^e , further relatives any high G^e on absolute terms. Thus, transferring these considerations to the context of measuring the dependency of e on a VN, we end up with $(1 - S^e)$. This will hence feed the computation of $\delta_{s/c}$ as derived from (4) in [ZRG11]⁶ (scaling the results in accordance with VNQ standards), allowing the integration of the *entity size* without requiring any weighting factors.

A.1.2.2 Overweighting

Following the Gini extension in [Yit83], we will further construct a comparable measure to the Atkinson index that allows a controlled overweighting of low or high values in the set of alternatives. Transferring the ideas of [Yit83] to this context, we introduce the inequality preference parameter v in the computation of $Gini(r)$ (cf. (2.6)):

$$Gini(r)^v := \sum_{j \in \{1, \dots, n_r\}} [p(j)]^v \quad , \quad (\text{A.10})$$

where $v = 2$ by default, which is kept static and unaltered for $\delta_{s/c}$. The computation of $\iota_{s/c}$, however, flexibly accommodates for the moderation of inequality preferences. When lowering v in analogy with the position in [LY85], the inequality of low values (utilities, incomes, etc.) is overweighted until constant inequality is reached at $v = 1$ —ie, an absolute indifference to inequalities. Thus, the result converges towards the “Rawlsian criterion”⁷ when v approximates 1. Contrary, raising v overvalues indifferences of high values until total indifference is reached again at $\inf$. The realistic scope of action may, thus, lie in $[1, 3]$ as symmetry from indifference to strong overweighting of inequalities centred around the default Gini calculation. For the quantification of $\delta^{in/out}$ we, hence, recommend the usage of 3 in order to overweight the inequality in the top range as a replacement for the spreading factor $y(j)$. In other words, if we consider two extreme outliers in the top range the classical Gini coefficient with $v = 2$ detects a high inequality level, while at $v = 3$ the inequality may appear much lower whenever the outliers are similar to each other, ie, the best offer is being matched by an attractive competing alternative.

⁶ The original paper uses the notion $\delta_{1/2}$ instead.

⁷ The Rawlsian criterion originates from a work on justice theory in 1971 [Raw09] where Rawls recommends to “maximize the prospect of the least advantaged”. In our case, the differences at the low utility end have to be overweighted to converge to a Rawlsian criterion.

NET NEUTRALITY (NN)

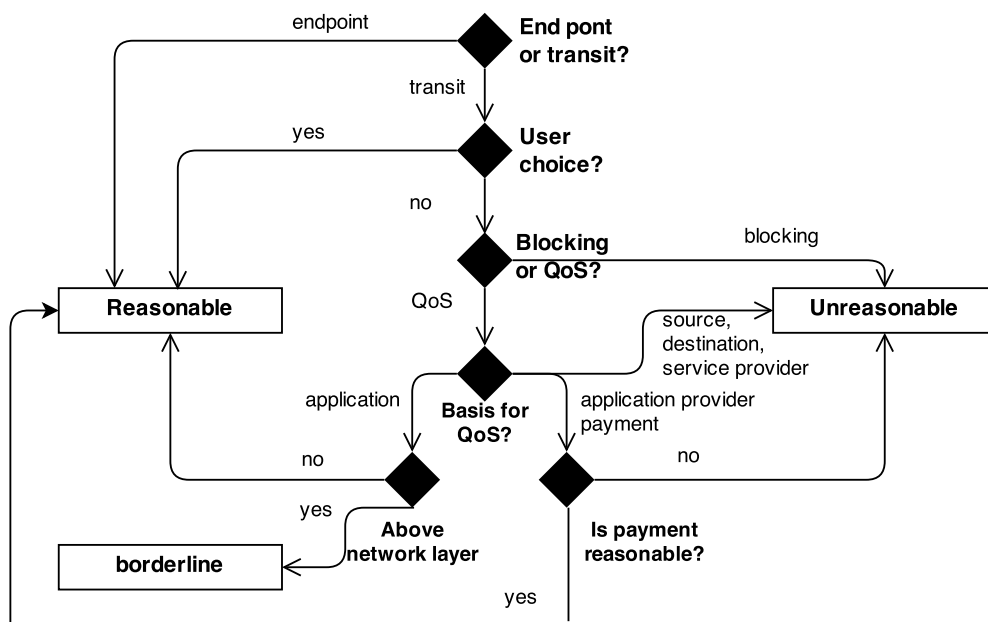


Figure B.1: Classification framework for traffic management techniques [JG10] (redrawn).

QoE AND WTP

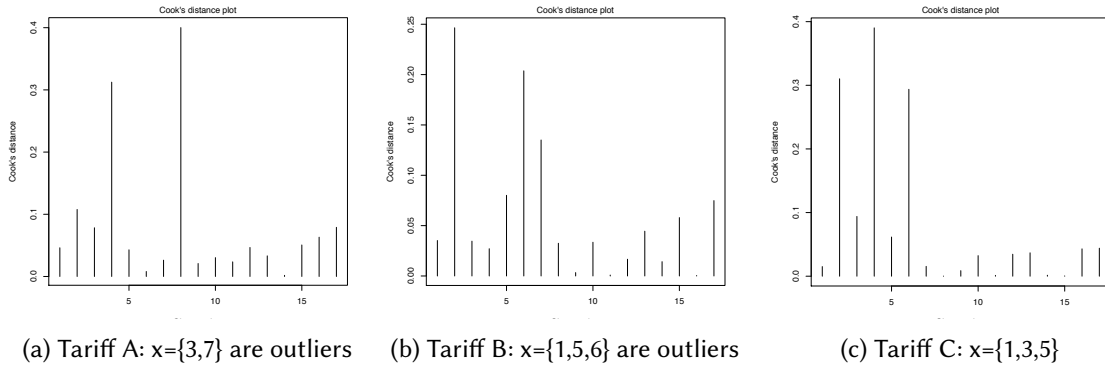


Figure C.1: Cook's distance diagram for quality the observation number obs , which is equivalent to $q - 1$, ie, $obs \in [1, 17]$ for $q \in [0, 156]$ (2015 trial).

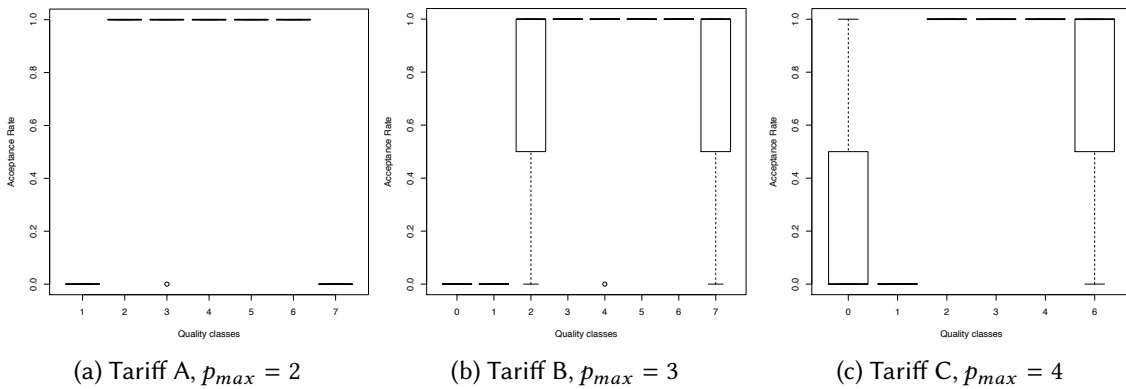


Figure C.2: Acceptance box plot for tariffs A, B and C (2015 trial).

UTILITY

D.1 Numeric Results

The user utilities for the controlled degradation VoD case of Section 5.3, (both SD and HD) can be represented in array form for the considered bitrate range of $[0, 40]$ Mbit/s as follows¹:

```
//User Utility: SD
```

```
[-∞, 0.781, 0.813, 0.832, 0.845, 0.855, 0.864, 0.871, 0.877,
  0.883, 0.887, 0.892, 0.896, 0.900, 0.903, 0.906, 0.909, 0.912,
  0.915, 0.917, 0.920, 0.922, 0.924, 0.926, 0.928, 0.930, 0.932,
  0.933, 0.935, 0.937, 0.938, 0.940, 0.941, 0.943, 0.944, 0.945,
  0.947, 0.948, 0.949, 0.950, 0.952]
```

```
//User Utility: HD
```

```
[-∞, 0.576, 0.638, 0.675, 0.701, 0.721, 0.737, 0.751, 0.763,
  0.774, 0.784, 0.792, 0.8, 0.807, 0.814, 0.82, 0.826, 0.832,
  0.837, 0.842, 0.846, 0.851, 0.855, 0.859, 0.863, 0.866, 0.87,
  0.873, 0.877, 0.88, 0.883, 0.886, 0.889, 0.892, 0.894, 0.897,
  0.899, 0.902, 0.904, 0.907, 0.909]
```

With a bitrate of 0 Mbit/s, no video can apparently be shown, which leads to a non-positive result. The corresponding NSP utilities can be represented as follows:

```
//NSP Utility: SD
```

```
[0., 0.009, 0.017, 0.0235, 0.0300, 0.0356, 0.0413, 0.047, 0.052,
  0.058, 0.063, 0.068, 0.073, 0.078, 0.082, 0.087, 0.092, 0.097,
  0.101, 0.106, 0.110, 0.115, 0.119, 0.124, 0.128, 0.133, 0.137,
  0.141, 0.146, 0.150, 0.154, 0.158, 0.163, 0.167, 0.171, 0.175,
  0.179, 0.183, 0.187, 0.191, 0.195]
```

¹ Due to the empirical answers provided by subjects in the 2012 *trial* the maximum of 1 in the $[0, 1]$ is not reached.

```
//NSP Utility HD
```

```
[0., 0.001, 0.004, 0.009, 0.015, 0.023, 0.032, 0.042, 0.0538,  
    0.067, 0.081, 0.096, 0.113, 0.131, 0.149, 0.169, 0.190, 0.212,  
    0.236, 0.260, 0.285, 0.312, 0.339, 0.367, 0.397, 0.427, 0.459,  
    0.491, 0.525, 0.559, 0.594, 0.631, 0.668, 0.706, 0.745, 0.786,  
    0.827, 0.869, 0.911, 0.955, 1.0]
```

For the packet loss values in percent,

```
//Packet Loss Range
```

```
[4.949, 4.798, 4.647, 4.496, 4.345, 4.194, 4.043, 3.892, 3.741,  
    3.589, 3.438, 3.287, 3.136, 2.985, 2.834, 2.683, 2.532,  
    2.381, 2.230, 2.079, 1.928, 1.776, 1.625, 1.474, 1.323,  
    1.172, 1.021, 0.870, 0.719, 0.568, 0.4167, 0.266, 0.115] ,
```

where 4.949 (poorest quality) and 0.115 (best) can be compared to 0 and 32 Mbit/s respectively from the HD NSP utility data, the following utilities are retrieved:

```
//NSP Utility Packet loss (HD-based)
```

```
[0., -0.038, -0.063, -0.081, -0.095, -0.105, -0.111, -0.114,  
    -0.113, -0.109, -0.103, -0.094, -0.082, -0.067, -0.05,  
    -0.03, -0.008, 0.017, 0.044, 0.073, 0.105, 0.139, 0.175,  
    0.213, 0.253, 0.296, 0.341, 0.387, 0.436, 0.487, 0.54,  
    0.595, 0.652]
```

The corresponding user utility can be followed from (5.19), which needs to be normalised to $[0, 1]$.

D.2 Figures

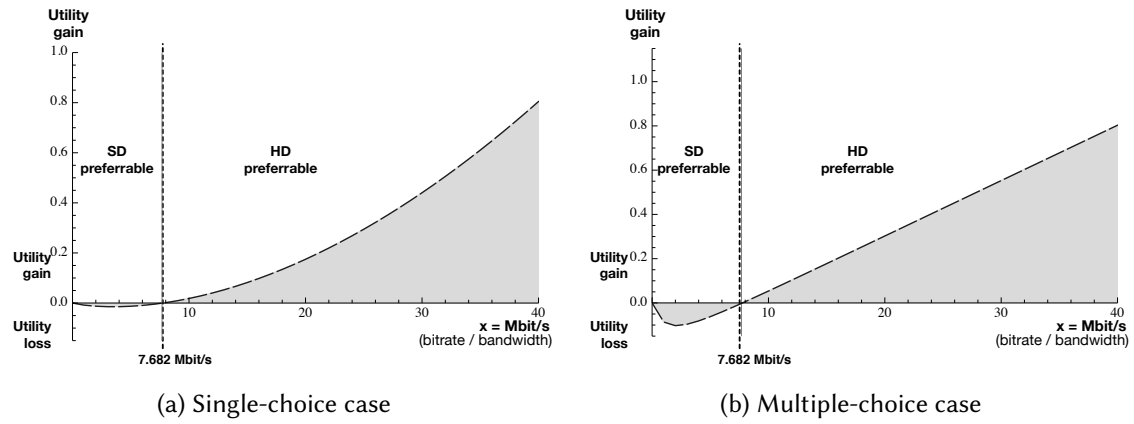


Figure D.1: Utility gain of remaining in HD over switching to SD video streams (see Section 5.3.2.3).

List of Conclusions

The following list reiterates the conclusions from the main sections of this monograph. An interpretation of outcomes is given in Section 7.2.

Conclusion #2.1. *It is economically preferable to transit from a single-class to a multi-class Internet service provisioning when providing deliberately selected quality classes.*

Conclusion #2.2. *The VoD scenario aggregates at least two separate value flows in one VN, ie, the network transmission and the actual content sales. While both value streams have to coexist, their business models are only loosely coupled. Whenever a platform controls both value flows, a central bottleneck role with high bargaining powers results.*

Conclusion #2.3. *The VoD platform, as probable bottleneck role, represents an intermediary two-sided market for the service sales. While two-sided markets can be attractive, the market entrance phase is more challenging.*

Conclusion #2.4. *If the NSP is also operating the VoD platform, strategic advantages can only be utilised, if target customers also include the Internet service customers from other NSPs, or if no cross-NSP competitor exists.*

Conclusion #2.5. *In a limited VoD scenario, the Transit NSP has the most notable dependency on the VN. If configured as a separate entity, this position may need to be reconsidered in the future. The consumer business of small NSP can also be regarded to be structurally challenging in competitive markets.*

Conclusion #2.6. *Entity sizes strongly affect the market powers in a VN.*

Conclusion #3.1. *The inclusion of QoE in the charging process for Internet qualities can yield revenue and social welfare gains, and can be beneficial for customers when the scheme is properly configured.*

Conclusion #3.2. *The 2011 trial demonstrates the existence of a substantial WTP or RTP for purchasing network video quality upgrades.*

Conclusion #3.3. *Due to the existence of several customer groups in the 2011 trial, it can be concluded that a homogenous population of customers seems unrealistic in the market for network video quality.*

Conclusion #3.4. *QoE ratings, both acceptance and MOS ratings (on ACR-5 scale), are non-linearly affected by pricing. A WTP approximation from QoE without thorough characterisation of the relationship between QoE and WTP appears to be notoriously difficult.*

Conclusion #3.5. *Active user decisions may trigger cognitive dissonance effects that may be exploited by tailor-made business models for niches, eg, low-quality and low-price niches.*

Conclusion #3.6. *Classical QoE data cannot be assessed within a WTP trial, as the price cognitions will affect QoE figures.*

Conclusion #3.7. *The unclear absolute WTP bounds from the 2011 trial anticipate the testing of **higher priced tariffs**. This step should be paired with a test of **price variations** in order get clarity over purchasing motives (ie, satisfaction with quality level vs. monetary considerations).*

Conclusion #3.8. *A substantial maximum WTP exists for network quality enhancements.*

Conclusion #3.9. *The WTP monotonically increases in the charged price.*

Conclusion #3.10. *While the first interaction phase during the purchase of a network service is dominated by quality cognitions (ie, identification of the desired quality level), the few final seconds are dominated by price decisions (ie, reductions to a point where the price-quality balance is acceptable).*

Conclusion #3.11. *Tariff levels affect the interaction behaviour of customers with a market system.*

Conclusion #3.12. *A reasonable likelihood exists that historic prices for network services may affect future purchases with unrelated price figures.*

Conclusion #3.13. *Monopolistic price increases in large steps are preferable over small modifications for NSPs².*

Conclusion #3.14. *From a per-user perspective, NSPs are **not** recommended to offer unsustainable low teaser prices at market entrances that cannot be retained over time.*

Conclusion #3.15. *In monopolistic network markets, stable pricing³ always provides the best revenue figures. However, if necessary, price reductions from high initial prices are preferred over price increases from low teaser prices.*

Conclusion #3.16. *In a competitive setting, larger but sustainable price cuts may be more useful to attract new customers, which at the same time reduce the marginal cost, ie, positive scaling effects.*

² It may be tested whether small price steps could recreate this effect.

³ Prices that do not require modifications after their initial announcement are referred to as “stable” in this context.

Conclusion #3.17. *Immature markets, ie, market entrance with or without high competition, may be best served by progressively penetrating the market, eg, through aggressive pricing, in order to decrease unit costs.*

Conclusion #3.18. *Teaser price strategies for the market entrances phases, eg, “filling the pipe” strategies, may trade short-term interest for the long-term sustainability of the market. Such pricing strategies should, hence, be only used carefully and optimally as measure of last resort, eg, when marketing strategies insufficiently meet the demands of an immature market.*

Conclusion #3.19. *Increased marketing efforts (demand stimulation due to blandishing pricing effects as suggested by [Sac+12b]) and ante-dated communication of investments (attracting interest of business partners) may be preferred over unsustainable introductory prices for Internet services in the long run.*

Conclusion #3.20. *Active decisions of users should be avoided when applying moderate price increases. Subtle price increase automatisms may be used in order to circumvent situations requiring a decision by the customer and having the potential to lead to a loss of revenue (see Section 3.3 and 3.6.1).*

Conclusion #3.21. *For **low competition** in a telecom market, significant price increases may be preferred over small changes to avoid revenue loss or stagnation. Small, but adequate, price reductions are advised (immediately revenue-effective), if necessary. In contrast, under **high competition**, more significant price cuts are advised due to market penetration effects and flattening revenue curves.*

Conclusion #3.22. *First-degree (eg, through the functioning of virtual quality classes, see Section 3.3) and third-degree price discrimination (see customer segments in [Sac+12a]) is recommended for network services. Teaser prices should be applied only to price-sensitive customers.*

Conclusion #4.1. *WTP is according to literature [Cow07] beta distributed. For the case of telecommunications products however, BE Internet (low cost, no quality guarantees) may represent a notable exception from an otherwise beta-distributed relationship. Further empirical testing is required.*

Conclusion #4.2. *Customer segments (attaching users to the best matching offer from an available set of choices) and typical demand-centric WTP figures (demand for an offer with a given price and quality without any choice), as addressed in [Cow07], are likely different concepts following different demand distribution curves. In other words, the demand for each combination of price and quality from a set of alternatives, ie, customer segments, needs to be separated from the overall purchasing trends, ie, WTP.*

Conclusion #4.3. *For the **normally-distributed customer segments** (in terms of demand being attached to offered quality and price combinations) the peak can be expected at medium quality and price levels (on a relative perspective between poorest and best offers).*

Conclusion #4.4. *The peak in the the normally distributed **customer segments** curve may, due to steeper or flatter price increase curves, slightly shift towards lower- or higher-quality classes respectively.*

Conclusion #4.5. *Within a single assessment context or trial, QoE perception may follow a logarithmic shape between the poorest and best quality level. Across trials, this perspective may, however, lead to inconsistencies that may hamper the meta analysis and commercialisation. A globally applicable QoE metric has not yet emerged, but would be required for practical utilisation of QoE information.*

Conclusion #4.6. *For a single-class service provisioning, the revenue $\mathcal{R}$ for network services continuously increases in demand, and, thus, theoretically goes to ∞ for infinite capacity ($C = \infty$). Hence, the provisioning of the a QoS class Q yielding the individually highest demand is naïvely preferable.*

Conclusion #4.7. *Whenever no sufficient premium demand exists, ie, very few customers utilising the capacity C under acceptable bargaining powers, a substantial extension of the customer base may provide cost advantages over competitors due to scaling effects.*

Conclusion #4.8. *For a single-class service provisioning, VoD services—under the given tariffing design and controlled conditions—should be provisioned with intermediate quality.*

Conclusion #4.9. *While price increases (raised maximum price $p_{\max}$ and price steps) first yield higher profits, this effect flattens with increasing price differences.*

Conclusion #4.10. *For a single-class service provisioning, high demands and low capacity shift profit optimum towards low-quality services (lower x values and the corresponding quality classes Q_x).*

Conclusion #4.11. *For a single-class service provisioning, the premium segment (high quality, high price offers) is unattractive even under unconstrained capacity.*

Conclusion #4.12. *Multi-class service provisioning yields higher revenues and, thus, revenues for capacity C than a corresponding single-class service provisioning. The effect becomes more distinctive under tight capacity constraints, where low quality offers are rendered more attractive in the single-class case, but not with the multi-class design.*

Conclusion #4.13. *For multi-class service provisioning, more expensive tariffs typically yield higher revenues for a given capacity C .*

Conclusion #4.14. *Combined cellular and fixed line NSPs have a clear **incentive to offload** traffic from cellular to Wi-Fi networks using their fixed line backbone networks.*

Conclusion #4.15. *Combined cellular and fixed line NSPs have **no incentive** to protect a QoS level above bulk traffic requirements in the unlicensed spectrum, ie, an equilibrium at a high utilisation of unlicensed spectrum exists.*

Conclusion #4.16. *Combined cellular and fixed line NSPs can cannibalise the unlicensed spectrum to materialise on their **strategic advantage** of controlling a portion of the available licensed spectrum.*

Conclusion #4.17. *Combined cellular and fixed line NSPs have **no incentive** of protecting a QoS level above the most challenging service requirements in the licensed spectrum.*

Conclusion #4.18. *For the common Internet QoE-differentiation or any paid prioritisation is legally disallowed in the EU. However, “special services” can make use of such practices, based on the interpretation of the EP and CoEU that uses legal loopholes for a softer handling of NN in practice.*

Conclusion #4.19. *Paid QoE-differentiation appears to be unlikely in the USA but also in the EU. Depending on the convergence of the legal situation and the functioning of the market, niches or separate services may be offered. Single class QoE optimisation will not be targeted by current NN rulings, but may be of limited benefit to end users.*

Conclusion #4.20. *Globally, the transparency that allows customers to evaluate services properly, seems to be commonly understood as one of the key principles of NN and remains in full force, eg, in the UK and the USA.*

Conclusion #4.21. *From a structural VN point of view, a side payments paradox exists in analogy to [ACK10], which favours the sender of side payments, typically the CP, over the receiver, typically the NSP.*

Conclusion #4.22. *The economic feasibility of applying side payments for network services seems questionable from a game-theoretic point of view. While paradox effects cannot be confirmed in this context (where receivers of side payments would obtain less revenue and less profit in total), the model clearly points to the absence of relevant revenue or profit gain potentials when considering a large-scale market introduction of side payments. Niche cases, like applying side payments as surcharges for specific contents or players, may still be economically attractive, as the used model does not capture such cases.*

Conclusion #5.1. *Utility and QoE are disparate concepts.*

Conclusion #5.2. *Utility data are inherently required for further optimising the operations of communications networks.*

Conclusion #5.3. *Separate QoE results, as local QoS-to-QoE mappings, need to be concatenated using context-specific adaptations in order to reflect the user’s appreciation across service types and test scenarios.*

Conclusion #5.4. *User utility has to integrate both purchasing decisions (ie, monetary aspects) and QoE considerations.*

Conclusion #5.5. *As the NSP utility does only indirectly relate to QoE measurements, a separate quantification is required.*

Conclusion #5.6. *The NSP utility directly relates to WTP and demand figures, which are insufficiently available—an approximation from QoE is required.*

Conclusion #5.7. *The user utility $\mathcal{U}^u$ can slightly be increased when applying an intelligently controlled degradation. Benefits due to low prices for low qualities mitigate the potential for user utility gains.*

Conclusion #5.8. *The NSP utility $\mathcal{U}^n$ can be increased substantially when offering multiple alternatives to customers. Whenever focusing on a single offer, $\mathcal{U}^n$ resembles the limited gain potential of $\mathcal{U}^u$.*

Conclusion #5.9. *Following QoE literature, the interdependence of QoS metrics jointly affecting the QoE outcome is most likely characterised by specifically parameterised weighting factors or functions.*

Conclusion #5.10. *The utility-aware network flow path computation requires a considerate handling of complexity increases due to precision advancements.*

Conclusion #5.11. *The demand for network video quality is substantially more sensitive to packet loss increase than bitrate reductions (see price curves compensating the higher quality sensitivity as given in (5.22)).*

Conclusion #5.12. *The utility converges to zero even for moderate packet loss levels. High packet losses may easily create a negative utility, which reflects the highly negative quality experience.*

Conclusion #5.13. *The NSP utility $\mathcal{U}^n$ is highly sensitive to packet losses and moderately sensitive to bitrate and associated bandwidth variations.*

Conclusion #5.14. *Work-conservation, as required for an efficient network resource usage, is optimally provided by DRR and WFQ.*

Conclusion #5.15. *A CO extent of around 100% is required in order to assure a realistic reliability, ie, 99.99% [Tel02], for providing quality-assured network services.*

Conclusion #5.16. *BE requires assigned minimum resource guarantees, eg, a DRR Quantum of 25%, for avoiding a complete erosion of the BE Internet product segment at exceptional load levels. Any residual capacity, which cannot be sold effectively for a premium price, are returned to the BE segment.*

Conclusion #5.17. *Quality-assured network services need to be separated from basic connectivity demands in order to on the one hand protect BE Internet and on the other hand provide quality gains for “premium” customers.*

Conclusion #6.1. *In noisy test environments, eg, when subjects are unfamiliar with the purchasing system or the video content is too dissimilar for the chosen evaluation period (as is probably in our case), market entrance pricing effects can neither be confirmed nor disproved.*

Conclusion #6.2. *In the 2015 trial, subjects compensated average price increases with corresponding quality reductions.*

Conclusion #6.3. *In the 2015 trial, subjects did not actively responded to price cuts, which lowered the absolute spending proportionally to the applied price cuts.*

Conclusion #6.4. *Based on the 2015 trial data, it can be confirmed that prices and purchasing situations affect quality perception.*

Conclusion #6.5. *MOS ACR-5 ratings are less sensitive to QoS alterations, when subjects pay for quality upgrades. Both the 2012 trial and 2015 trial confirm this claim.*

Conclusion #6.6. *While QoE ratings obtained in a purchasing situation, represented as MOS on ACR-5 scale, can be approximated using logarithmic curves, the quality of the fit is lower at QoS and price extrema than expected for QoE ratings without a purchase.*

Conclusion #6.7. *The VoD and likely the network quality markets are of supraregional or even global character.*

Conclusion #6.8. *The offering of a larger set of “inadequate” quality levels—products that are unlikely purchased due to their unappealing quality—might entice customers to spend more money.*

Conclusion #6.9. *For future WTP trials, the inherent price bias of WTP trials deserves careful handling and moderation. Strategies that allow subjects to both keep an eye at the price levels and to perceive the network video quality might be beneficial (as in the 2012 trial).*

Conclusion #6.10. *Market entrance pricing effects for network quality markets are likely, but deserve further testing in large-scale campaigns or with the help of meta analysis (if applicable). We further recommend a bigger number of quality classes as in the 2012 trial (or comparable strategies) in order to diffuse the range of available product offers and to centre the attention on the perceived product valuations. A low number of quality classes supports strategies to avoid active redeceptions, which may lower the outcomes of campaigns.*

Nomenclature

Symbols

α	Consumers' quality surplus
β	Network resource quantities
$\mathcal{R}$	Revenue
π	Profit
C	Capacity of the network
q	Quality of Service (continuous)
x	Quality of Experience (continuous)
Q	Discrete quality class
$\mathcal{Q}$	Set of all considered discrete quality classes $\{Q_i, Q_{i+1}, \dots\}$
d	Service demand (quantities)
t	Time / Point in time / Measurement round index
ζ	Network resource usage per-service demand d_i per-time t , s.t. the desired q or Q_i in t
Z	Overall resource demand over all service demands $\{d_i, d_{i+1}, \dots\}$, eg, $\sum_{\{d_i, d_{i+1}, \dots\}} \zeta(Q_i)$ for a single service class with quality class Q_i
G	A Value Network Graph representing a specific VN
E	Set of nodes in G (reflecting the entities of the corresponding VN)
R	Set of relationships in G (reflecting the business interactions of the corresponding VN)
$\mathcal{L}$	Set of labels where each label L is attached to an edge of G
φ	(Initial) Value of business interaction or resource (without considering costs or value adjustments due to competitive forces)
ξ	Fungibility of a resource type
I	Investment costs
p	Price / Charged fee
ι	Industry rivalry

δ	VN dependency indicator
Δ	VN metric that describes the dependency of an entity on a given VN (as a result of all δ_i with $i \in \{s, c, r\}$)
ρ	The bargaining power of an entity in a VN as opposite of δ
T	Coalition of entities (in a game-theoretic sense)
am	Amplitude
cv	Convergence time
F	Fast Convergence
R	Regular Convergence
S	Slow Convergence
X	Irregular Convergence
l	Packet loss
b	Bandwidth / Bitrate

Glossary

AC

“Admission control (AC) limits the number of flows in the network by denying access to new ones if the network risks to be overloaded. It has been proposed for the Internet in [She95].” [MMC06]. 181

AS

Number of connected IP routing prefixed with routing policy to the Internet [RFC 1930]. vii, 15, 175, 176, 178, 181, 261

Domain

A domain refers to ASes being managed by a single corporation. Deutsche Telekom may for example handle various globally distributed but connected ASes, which together represent a single domain. xviii, 3, 6, 9, 10, 28, 34, 132, 173–182, 229, 232, 261

IC

Interconnection (IC) refers to the connection of various Domains within a network, eg, via transit or peering agreements. 7, 8, 33, 35, 36, 42, 56, 66, 174–178, 181–183, 231, 232

NP

The complexity class *NP* of nondeterministic polynomial time problems; the set of *NP*-hard problems are hardest to be computationally solved. 138

Quantum

The traffic quantum is used for weighting the class-based traffic in the DRR or comparable scheduling discipline. The quantum represents the share of resources assigned to a class, where unused resources are made available to other services classes in work-conserving disciplines such as DRR or WFQ. 204–208, 210

Acronyms

3GPP Third Generation Partnership Project. 177, 275
ACR Absolute Category Rating. 19, 82, 93, 115, 131, 184, 186, 191, 217, 220, 223–226, 252, 257
ADAPTISE ADmission control and resource Allocation for adaPtive mulTImedia SEr-vices. xviii, 179, 180, 202, 213
AF Assured Forwarding. 175
ALTQ Alternate Queueing. 203
ANOVA Analysis of Variance. 99, 111, 263
ANOVA RM ANOVA with Repeated Measures. 111, 222
API Application Programming Interface. 37
ARCEP Autorité de Régulation des Communications Électroniques et des Postes. 277
ARPU Average Revenue Per User. 2, 30, 110, 114, 133
ASPW Average Service Preference Weighting. 189, 190, 192, 200
ATM Asynchronous Transfer Mode. 17
AVC Advanced Video Coding. 94
BDM Becker–DeGroot–Marschak Method. 81
BE Best Effort. vii, 3, 6, 7, 14, 23–27, 81, 113, 119, 125, 128, 147, 148, 152, 153, 174, 175, 203, 205, 207, 208, 210, 230, 233, 253, 256
BEREC Body of European Regulators of Electronic Communications. 152, 153
BGP Border Gateway Protocol. xviii, 175, 176
BIQINI BACCARDI IMS QoS Implementation. 205, 210
BM Business Model. 36, 56–58
CAPEX Capital Expenditure. 1, 30, 31, 133, 134, 143
CBQ Class Based Queuing. 204, 205
CC Closeness Centrality. 179
CDF Cumulative Distribution Function. 122, 124, 128, 129, 193, 194
CDN Content Distribution Network. 41, 65, 66, 73–75
CIF Common Intermediate Format. 186, 191, 193

CIT Critical Incident Technique. 102

CO Capacity Overprovisioning. 81, 205–207, 256

CoE Council of Europe. 153

CoEU Council of the EU. 153, 154, 255, 277

CP Content Provider. 5, 32–34, 41–43, 70, 73–75, 158–166, 168, 169, 229, 231, 255

CPE Customer Premise Equipment. 42, 141, 142, 146

CRI Composite Risk Index. 56

CRTC Canadian Radio-Television and Telecommunications Commission. 157

CS Customer Satisfaction. xvi, 4, 18, 22–24, 31, 32, 80–82, 87, 88, 115, 116, 130, 149, 166, 183, 223

DASH Dynamic Adaptive Streaming over HTTP. 216, 217, 227

Diffserv Differentiated Services Architecture. xviii, 174, 175

DRR Deficit Round Robin. 204–206, 208, 256, 261

DSCP Differentiated Service CodePoint. 174

DVD Digital Versatile Disc. 70, 71

EC European Commission. 181

EF Expedited Forwarding. 175

EGP Exterior Gateway Protocol. 175

ELA Experience Level Agreement. 183

EP European Parliament. 152–154, 255

ERMET Ericsson Risk Management Evaluation Tool. 56

ETICS Economics and Technologies for Inter-Carrier Services. v, 9, 89, 178, 181–183, 232, 283

ETSI European Telecommunications Standards Institute. 177, 273, 275, 277

EU European Union. xvii, 152–154, 157, 255, 264

FCC Federal Communications Commission. 152, 155–158, 277

FP7 Seventh Framework Programmes for Research and Technological Development. 9, 181, 283

FTW FTW Telecommunications Research Center Vienna. 8, 89, 94, 215, 269, 295

GG Gabriel Graph. 179

GML Graph Modelling Language. 64

GMPLS Generalized Multi-Protocol Label Switching. 178

GPS Generalized Process Sharing. 204

GSM Global System for Mobile Communications. 176, 264, 274

GSMA GSM Association. 176, 274

HCI Human Computer Interaction. 21

HD High Definition. 73, 74, 94, 97, 114, 131, 185–187, 191–193, 197, 200, 202, 217, 247–249

HEVC High Efficiency Video Coding. 216

HFS Hierarchical Fair Service Curve. 204

HSS Home Subscriber Server. 177

HTB Hierarchical Token Bucket. 205, 208
HTTP HyperText Transfer Protocol. 216, 264
ICN Information-Centric Networking. 9
iff if and only if. 62
IGP Interior Gateway Protocol. 175
IMS IP Multimedia Subsystem. xviii, 6, 176–178, 180–182, 205, 208, 210, 211, 213, 233, 263
Intserv Integrated Services Architecture. xviii, 174, 175
IoT Internet of Things. xi, xiii, 2, 13
IP Internet Protocol. vii, 3, 174, 176–178, 265, 267
ipfw ipfirewall. 203, 204
IPR Intellectual property. 233
IPv4 IP version 4. 174
IPv6 IP version 6. 174, 178
IPX IP exchange. xviii, 176, 177, 181, 182
ISO International Organization for Standardization. 15, 16, 273, 274
ISP Internet Service Provider. 176
ITU International Telecommunication Union. 15, 16, 18, 89, 91, 94, 215, 273, 274
JSON JavaScript Object Notation. 64
LCD Liquid Crystal Display. 97
LER Label Edge Router. 178
LSP Label Switched Path. 178
LTE Long Term Evolution. xviii, 1, 3, 6, 61, 142, 177–180, 203, 206, 213, 233
M2M Machine-to-Machine. xi, xiii, 2, 34
M3I Market Managed Multi-Service Internet. 22, 88, 89, 114, 115
MDE Model Driven Engineering. 42
MDP Media Degradation Path. 173, 174, 179–181
MLD Mean Log Deviation. 237
MOS Mean Opinion Score. 19, 82, 91–93, 99, 115, 131, 147, 184, 186, 187, 189–192, 197–200, 217, 223–226, 252, 257
MPEG Moving Picture Experts Group. 94
MPLS Multi-Protocol Label Switching. 178
NE Nash Equilibrium. 163
NGN Next Generation Networks. 80
NIP No Interaction Phase. 90, 219
NN Net Neutrality. vii, xii, xiv, xvii, xix, 4, 5, 7, 34, 67, 76, 86, 120, 146–148, 150–158, 171, 229, 230, 232, 234, 243, 255
NSP Network Service Provider. viii, xii, xviii, 1, 3–5, 8, 9, 25, 27, 32–34, 36, 41–43, 47, 52, 59–61, 66, 67, 69–76, 80, 81, 84, 86, 112, 113, 116, 132, 133, 141, 144–147, 152–166, 168, 169, 172, 173, 176, 181–185, 188–190, 192–195, 197, 198, 201, 205, 212, 213, 229–232, 235, 247, 248, 251, 252, 254–256

NVA Network Value Analysis. 35
OPEX Operational Expenditure. 31, 133, 134, 143
OS Operating System. 208
OSI Open Systems Interconnection. 15, 16
OSPF Open Shortest Path First. 175
P-CSCF Proxy Call Session Control Function. 177
PAN Personal Area Network. 34
PARQUE Pricing and Regulating Quality of Experience [WZR12]. 83–86
PCC Policy and Charging Control. 177
PCEF Policy and Charging Enforcement Function. 177
PCRF Policy and Charging Rules Function. 177, 182
PDF Probability Density Function. 99, 100, 122, 123, 126, 127
PDP Policy Decision Point. 177, 182, 210, 212
PED Price Elasticity of Demand. 81
PEP Policy Enforcement Point. 177, 210, 212
PHB Per Hop Behavior. 175
PMP Paris Metro Pricing. 82–86
PSNR Peak Signal-to-Noise Ratio. 20
PSP Progressive Second Price. 132, 137
Q-POINT QoE-driven Path Optimization Model for Multimedia Services. xviii, 179, 180, 195–198, 201, 202, 213, 234
QoD Quality of Design. xv, 21, 22
QoE Quality of Experience. vii, ix, xi–xix, 2–10, 13, 14, 16–20, 22, 23, 25, 27, 29–34, 65, 67, 71, 72, 76, 79, 80, 82, 84–93, 99–101, 115, 120, 129–133, 137, 148–150, 154, 156, 157, 171, 173, 175, 176, 179–188, 191, 192, 195–199, 202, 203, 205, 206, 208–210, 212, 215, 220, 223, 225, 229–234, 245, 251, 252, 254–257, 259
QoP Quality of Perception. xv, 21, 22
QoS Quality of Service. ix, xi, xiii, xv, 2–5, 8, 9, 13, 15–18, 20, 22, 24–27, 29–34, 67, 69, 71, 76, 79–85, 87, 91, 93, 115, 129–134, 136–139, 142–150, 153, 155, 156, 162, 172, 174–184, 186–189, 192, 193, 196–203, 205, 206, 210, 212, 223, 225, 228–234, 254–257, 259, 263
QSP Quality Selection Phase. 90, 93, 97, 101–103, 108, 219–221, 225–227
QuEEN Celtic-Plus Quality of Experience Estimators in Networks. v, 8, 31, 87, 280
RMD Resource Management in DiffServ. 175
RNG Relative Neighborhood Graph. 179
RoI Return on Investment. viii, 1, 30
RoR Ruby on Rails. 217, 233
RSD Root Square Deviation. 108
RSVP Resource Reservation Protocol. 174, 175
RSVP-TE Resource Reservation Protocol – Traffic Engineering. 178
RTP Readiness-to-Pay. 23, 87, 91, 115, 251
RTR Rundfunk- und Telekom Regulierungs-GmbH. 153, 294

SaaS Software as a Service. 33
SD Standard Definition. 88, 91, 114, 131, 185–187, 191–193, 202, 247, 249
SDN Software Defined Networking. 180
SIM Subscriber Identity Module. 142
SIP Session Initiation Protocol. 177, 179, 180
SLA Service Level Agreement. xii, xiv, 173, 176, 181, 183, 232
SNA Social Network Analysis. 179
SP Switching Point. 186, 188, 189
SPP Sending Party Pays. 181
SPR Subscription Profile Repository. 177, 182
SSIM Structural SIMilarity. 20
TCP Transmission Control Protocol. 97, 138, 178
TISPAN Telecoms & Internet converged Services & Protocols for Advanced Network. 177
UDP User Datagram Protocol. 89–91
UK United Kingdom. 2, 42, 69, 73–75, 154, 157–159, 255
UMTS Universal Mobile Telecommunications System. 61, 112
UNIZG-FER University of Zagreb, Faculty of Electrical Engineering and Computing. 9, 180, 195, 202
USA United States of America. 42, 60, 69, 73–75, 142, 151, 152, 155, 157, 158, 160, 255
USD Universal Service Directive. 152
VCG Vickrey-Clarke-Groves. 138
VN Value Network. xii, xiv, xvi, xix, 4, 7–9, 13, 14, 27, 28, 32–50, 52, 53, 55–72, 74–76, 158–160, 163, 229, 231–233, 235–238, 240, 241, 251, 255, 259, 260, 267
VNA VN Analysis. 35
VNDM Value Network Dependency Model. xvi, 41–43, 50
VNG Value Network Graph. xvi, 43, 44, 48, 50–58, 63, 64, 259
VNQ Value Network Quantification. xvi, xix, 38, 40, 42, 46, 48–50, 52, 53, 57–59, 61, 65, 75, 76, 158–160, 233, 235, 236, 238, 239, 241
VoD Video on Demand. xii, xiv, xvi, 14, 38, 41, 65–76, 89–91, 94, 114, 134, 158–160, 184, 186, 191, 193, 198, 210, 215, 216, 228, 229, 247, 251, 254, 257
VoIP Voice over IP. 3, 6, 79, 177, 206
VoLTE Voice over LTE. 3, 177, 178, 213, 233
VPN Virtual Private Network. 178
VQEG Video Quality Experts Group. 87
VQM Video Quality Metric. 191
VTI VTT Technical Research Centre of Finland. 8, 215, 227
WFQ Weighted Fair Queuing. 175, 204, 256, 261
WRR Weighted Round Robin. 203–205
WTA Willingness-to-Accept. 22–24
WTP Willingness-to-Pay. xii, xiv–xix, 4, 7, 8, 20, 22–24, 30, 31, 33, 34, 76, 79–84, 87–94,

97, 99, 103, 108–111, 113, 115, 116, 119, 120, 122, 125, 128–130, 132, 133, 137, 139, 149, 166, 173, 174, 182–185, 188–190, 192, 198, 199, 202–204, 206, 208, 210–212, 220, 223, 228, 230, 232, 234, 245, 251–253, 255, 257

XML Extensible Markup Language. 179

Own Publications

- [Rei+12] P. Reichl, P. Maillé, P. Zwickl, and A. Sackl.
'On the Fixpoint Problem of QoE-based Charging'.
In: *Proc. of the 6th International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS)*. IEEE, 2012, pp. 235–242.
ISBN: 978-1-4673-4887-4.
- [Rei+13a] P. Reichl, P. Maillé, P. Zwickl, and A. Sackl.
'A Fixed-Point Model for QoE-based Charging'.
In: *Proc. of the ACM SIGCOMM Workshop on Future human-Centric Multimedia Networking (FhMN)*. 2013, pp. 33–38.
DOI: 10.1145/2491172.2491176.
- [Rei+13b] P. Reichl, P. Zwickl, C. Löw, and P. Maillé.
'How Paradoxical is the "Paradox of Side Payments"? Notes from a Network Interconnection Perspective'.
In: *Proc. of the Wired/Wireless Internet Communications (WWIC) conference*. 2013, pp. 115–128. DOI: 10.1007/978-3-642-38401-1_9.
- [Sac+12a] A. Sackl, S. Egger, P. Zwickl, and P. Reichl. 'The QoE Alchemy: Turning Quality into Money. Experiences with a Refined Methodology for the Evaluation of Willingness-to-pay for Service Quality'.
In: *Proc. of the Fourth International QoMEX Workshop*. IEEE, 2012, pp. 170–175. DOI: 10.1109/QoMEX.2012.6263846.
- [Sac+12b] A. Sackl, P. Zwickl, S. Egger, and P. Reichl. 'The Role of Cognitive Dissonance for QoE Evaluation of Multimedia Services'.
In: *Proc. of the 2012 IEEE Globecom Workshops*. IEEE, 2012, pp. 1352–1356.
DOI: 10.1109/GLOCOMW.2012.6477779.
- [SZR12a] A. Sackl, P. Zwickl, and P. Reichl.
From Quality of Experience to Readiness to Pay for Network Quality.
Tech. rep. FTW, 2012.

- [SZR12b] A. Sackl, P. Zwickl, and P. Reichl. ‘From Quality of Experience to Willingness to Pay for Interconnection Service Quality’. In: *IFIP NETWORKING 2012 Workshops*. Vol. 7291. LNCS. Springer, 2012, pp. 89–96. doi: 10.1007/978-3-642-30039-4_11.
- [SZR13a] A. Sackl, P. Zwickl, and P. Reichl. ‘The Trouble with Choice: An Empirical Study to Investigate the Influence of Charging Strategies and Content Selection on QoE’. In: *Proc. of the 8th International Workshop on Internet Charging and QoS Technologies (ICQT’13)*. 2013, pp. 298–303. doi: 10.1109/CNSM.2013.6727850.
- [SZR13b] A. Sackl, P. Zwickl, and P. Reichl. ‘QoE Alchemy 2.0: An Improved Test Setup for the Pecuniary Bias of QoE’. In: *Proc. of the Fifth International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 2013, pp. 40–41. doi: 10.1109/QoMEX.2013.6603204.
- [Var+15] M. Varela, P. Zwickl, P. Reichl, M. Xie, and H. Schulzrinne. ‘Experience Level Agreements (ELA): The Challenges of Selling QoE to the User’. In: *Proc. of the ICC 2015 Workshops*. IEEE, 2015, pp. 1741–1746. doi: 10.1109/ICCW.2015.7247432.
- [WZR12] S. Wahlmueller, P. Zwickl, and P. Reichl. ‘Pricing and regulating Quality of Experience’. In: *Proc. of the 8th EURO-NGI Conference*. 2012, pp. 57–64. doi: 10.1109/NGI.2012.6252165.
- [ZR12] P. Zwickl and P. Reichl. ‘An Instance-based Approach for the Quantitative Assessment of Key Value Network Dependencies’. In: *Proc. of the Networking Workshops 2012*. IEEE, 2012, pp. 97–104. doi: 10.1007/978-3-642-30039-4_12.
- [ZR13] P. Zwickl and P. Reichl. ‘Graph-Theoretic Roots of Value Network Quantification’. In: *Proc. of 19th EUNICE/IFIP WG 6.6 International Workshop*. Vol. 8115. LNCS. Springer Berlin Heidelberg, 2013, pp. 282–286. doi: 10.1007/978-3-642-40552-5_26.
- [ZR15] P. Zwickl and P. Reichl. ‘Marktplätze im Umbruch’. In: ed. by C. Linnhoff-Popien, M. Zaddach, and A. Grahl. Springer, 2015. Chap. Of Values and Networks: Designing Quantitative Analysis Methods for Pluralistic Telecommunication Markets, pp. 631–641. doi: 10.1007/978-3-662-43782-7_67.

- [ZRG11] P. Zwickl, P. Reichl, and A. Ghezzi. ‘On the Quantification of Value Networks: A Dependency Model for Interconnection Scenarios’. In: *Proc. of the 7th International Workshop on Advanced Internet Charging and QoS technology (ICQT’11)*. Vol. 6995/2011. LNCS. 2011, pp. 63–74. doi: 10.1007/978-3-642-24547-3_7.
- [ZRS14] P. Zwickl, P. Reichl, and A. Sackl. ‘The 15 Commandments of Market Entrance Pricing for Differentiated Network Services: An Antagonistic Analysis of Human vs. Market’. In: *Proc. of the Measurement, Modelling, and Evaluation of Computing Systems and Dependability and Fault Tolerance (MMB & DFT’14)*. Springer, 2014, pp. 162–176. doi: 10.1007/978-3-319-05359-2_12.
- [ZSR13] P. Zwickl, A. Sackl, and P. Reichl. ‘Market Entrance, User Interaction and Willingness-to-Pay: Exploring Fundamentals of QoE-based Charging for VoD Services’. In: *Proc. of the IEEE Globecom’13*. 2013, pp. 1310–1316. doi: 10.1109/GLOCOM.2013.6831255.
- [Zwi+13] P. Zwickl, P. Fuxjaeger, I. Gojmerac, and P. Reichl. ‘Wi-Fi Offload: Tragedy of the Commons or Land of Milk and Honey?’ In: *Proc. of the 24th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) Workshops*. IEEE, 2013, pp. 148–152. doi: 10.1109/PIMRCW.2013.6707854.
- [Zwi+15] P. Zwickl, P. Reichl, L. Skorin-Kapov, O. Dobrijevic, and A. Sackl. ‘On the Approximation of ISP and User Utilities from Quality of Experience’. In: *Proc. of the Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 2015. ISBN: ISBN: 978-1-4799-8958-4.
- [Zwi15] P. Zwickl. In: *Quality of Experience: From Assessment to Application (Dagstuhl Seminar 15022)*. Ed. by K. D. Moor, M. F. P. Reichl, and M. Varela. Vol. 5. 1. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015. Chap. On QoE Monetization and Forming a Global Customer Satisfaction Metric, pp. 77–78. doi: 10.4230/DagRep.5.1.57. URL: <http://drops.dagstuhl.de/opus/volltexte/2015/5036>.

Standards

- [BT.500-13] *ITU-R BT.500-13 Methodology for the Subjective Assessment of the Quality of Television Pictures*. International Telecommunication Union (ITU). 2012.
- [BT.500-7] *ITU-R BT.500-7 Methodology for the subjective assessment of the quality of television pictures*. International Telecommunication Union (ITU). 2002.
- [BT.710-4] *ITU-R BT.710-4 Subjective Assessment Methods for Image Quality in High-definition Television*. International Telecommunication Union (ITU). 1998.
- [D SG 2/16] *Teletraffic Engineering Handbook*. ITU–D SG 2/16 & ITC Draft 2001-06-20. International Telecommunication Union (ITU), June 2001.
- [D.197] *Definition of Quality of Experience*. ITU-T D.197. International Telecommunication Union (ITU), 2004.
- [E.800-08] *Terms and Definitions related to Quality of Service and Network Performance Including Dependability*. ITU-T Rec. E.800. International Telecommunication Union (ITU), 2008.
- [E.800-94] *Terms and Definitions related to Quality of Service and Network Performance Including Dependability*. ITU-T Rec. E.800. International Telecommunication Union (ITU), 1994.
- [ES 282] *Internet Converged Services and Protocols for Advanced Networking (TISPAN), Resource and Admission Control Subsystem (RACS): Functional Architecture*. ETSI ES 282 003 v3.4.2. European Telecommunications Standards Institute (ETSI), 2010.
- [G.107] *The E-model: a Computational Model for Use in Transmission Planning*. ITU-T G.107 (02/14). International Telecommunication Union (ITU), 2014.
- [IEC 13236] *Information Technology: Quality of Service—Framework*. ISO/ IEC 13236. International Organization for Standardization (ISO), 1998.

- [IEC7498] *Information Technology – Open System Interconnection – Basic Reference Model: The Basic Model*. ISO/IEC 7498-1.
International Organization for Standardization (ISO), 1996.
- [IR.34] *Guidelines for IPX Provider Networks*. GSMA IR.34 V9.1.
GSM Association (GSMA), 2013.
- [IR.92] N. Russell, ed. *IMS Profile for Voice and SMS*. GSMA IR.92 V7.0.
GSM Association (GSMA), 2013.
- [ISO 9241] *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs), Part 11: Guidance on Usability*. ISO 9241.
International Organization for Standardization (ISO), 1998.
- [ISO/IEC 14496-10] *ISO/IEC 14496-10 – MPEG-4 Part 10, Advanced Video Coding*.
International Organization for Standardization (ISO). 2012.
- [ISO/IEC 23008-2] *ISO/IEC 23008-2:2015: Information Technology – High Efficiency Coding and Media Delivery in Heterogeneous Environments – Part 2: High Efficiency Video Coding*. International Organization for Standardization (ISO). 2012.
- [ISO/IEC 23009-1] *ISO/IEC 23009-1:2014: Information Technology – Dynamic Adaptive Streaming over HTTP (DASH) – Part 1: Media Presentation Description and Segment Formats*.
International Organization for Standardization (ISO). 2012.
- [ITU-T H.264] *H.264: Advanced Video Coding for Generic Audiovisual Services*.
International Telecommunication Union (ITU). 2014.
- [ITU-T H.265] *H.265: High efficiency video coding*.
International Telecommunication Union (ITU). 2015.
- [P.10] *Vocabulary for Performance and Quality of Service*.
ITU-T P.10/G.100 (07/06).
International Telecommunication Union (ITU), 1998.
- [P.1501] *Subjective Testing Methodology for Web Browsing*. ITU Rec P.1501 (02/14).
International Telecommunication Union (ITU), 2014.
- [P.800] *Absolute Category Rating*. ITU-T P.800 (06/98).
International Telecommunication Union (ITU), 1998.
- [P.910] International Telecommunication Union (ITU).
Subjective Video Quality Assessment Methods for Multimedia Applications.
ITU-T Rec. P.910.
- [RFC 1930] Internet Engineering Task Force (IETF). *Guidelines for Creation, Selection, and Registration of an Autonomous System (AS)*. RFC 1930.
- [RFC 2210] J. Wroclawski. *The Use of RSVP with IETF Integrated Services*. RFC 2210.
Internet Engineering Task Force (IETF), 1997.

- [RFC 2460] S. Deering and R. Hinden. *Internet Protocol, Version 6 (IPv6) Specification*. RFC 2460. Internet Engineering Task Force (IETF), 1998.
- [RFC 2474] K. Nichols, S. Blake, F. Baker, and D. Baker. *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*. RFC 2474. Internet Engineering Task Force (IETF).
- [RFC 2475] Blake, Steven and Black, David and Carlson, Mark and Davies, Elwyn and Wang, Zheng and Weiss, Walter. *An Architecture for Differentiated Services*. RFC 2475 (informational). Internet Engineering Task Force (IETF), 1998.
- [RFC 3031] *Multiprotocol Label Switching Architecture*. RFC 3031. Internet Engineering Task Force (IETF), 2001.
- [RFC 3209] *RSVP-TE: Extensions to RSVP for LSP Tunnels*. RFC 3209. Internet Engineering Task Force (IETF).
- [RFC 3290] Y. Bernet, S. Blake, D. Grossman, and A. Smith. *An Informal Management Model for Diffserv Routers*. RFC 3290 (informational). Internet Engineering Task Force (IETF), 2002.
- [RFC 3439] Internet Engineering Task Force (IETF). *Some Internet Architectural Guidelines and Philosophy*. RFC 3439. 2002.
- [RFC 3945] E. Mannie, ed. *Generalized Multi-Protocol Label Switching (GMPLS) Architecture*. RFC 3945. Internet Engineering Task Force (IETF), 2004.
- [RFC 4364] *BGP/MPLS IP Virtual Private Networks (VPNs)*. RFC 4364. Internet Engineering Task Force (IETF).
- [RFC 6733] V. Fajardo, J. Arkko, J. Loughney, and G. Zorn. *Diameter Base Protocol*. Ed. by V. Fajardo and G. Zorn. RFC 6733. Internet Engineering Task Force (IETF), 1998.
- [RFC 791] *Internet Protocol: DARPA Internet Program Protocol Specification*. RFC 791. Internet Engineering Task Force (IETF), 1981.
- [TR 102] *Human Factors (HF); Quality of Experience (QoE) Requirements for Real-time Communication Services*. ETSI TR 102 643 V1.0.1. European Telecommunications Standards Institute (ETSI), Dec. 2009.
- [TS 103] *Quality of Experience; A Monitoring Architecture*. ETSI TS 103 294 V0.0.2. European Telecommunications Standards Institute (ETSI), Oct. 2014.
- [TS 23V12] *IP Multimedia Subsystem (IMS): Stage 2*. 3GPP TS 23.228 V12.6.0, Technical Specification Group Services and System Aspects. 2014.
- [TS 23V7] *IP Multimedia Subsystem (IMS): Stage 2*. 3GPP TS 23.228 V7.16.0, Technical Specification Group Services and System Aspects. Third Generation Partnership Project (3GPP), 2008.

- [RFC 2205] L. Zhang, S. Berson, S. Herzog, and S. Jamin.
Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification.
Ed. by R. Braden. RFC 2205. Internet Engineering Task Force (IETF).
- [Utt+15] J. Uttaro et al. *Best Practices for Advertisement of Multiple Paths in IBGP.*
2015.
- [Wal+15] D. Walton, A. Retana, E. Chen, and J. Scudder.
Advertisement of Multiple Paths in BGP. 2015.

Bibliography

- [ARC10] *Internet and Network Neutrality – Proposals and Recommendations*. White Paper. Autorité de Régulation des Communications Électroniques et des Postes (ARCEP), 2010.
- [COE15] *Draft Recommendation CM/Rec(2014) of the Committee of Ministers to member States on protecting and promoting the right to freedom of expression and the right to private life with regard to network neutrality*. Council of Europe, Steering Committee on Media and Information Society, 2015.
- [CoEU 10409] *Proposal for a Regulation of the European Parliament and of the Council laying down measures concerning the European single market for electronic communications and to achieve a Connected Continent, and amending Directives 2002/20/EC, 2002/21/EC and 2002/22/EC and Regulations (EC) No 1211/2009 and (EU) No 531/2012*. 10409/1/15. CoEU, 2015.
- [ETSIW] *Workshop Report of ETSI Workshop on QoS / QoE / User Experience Focusing on Speech / Multimedia Conference Tools*. European Telecommunications Standards Institute (ETSI), 2010.
- [FCC1461] *Protecting and Promoting the Open Internet NPRM*. FCC FCC 14-61. FCC, 2015.
- [FCC1524] *Report and Order on Remand, Declaratory Ruling, And Order*. FCC 15-24. FCC, 2015.
- [FCCOI] *Open Internet Order*. FCC 10-201. FCC, 2010.
- [ABV12] I. Amigo, P. Belzarena, and S. Vaton.
'On the Problem of Revenue Sharing in Multi-domain Federations'.
In: *Proc. of the NETWORKING 2012*. Springer, 2012, pp. 252–264.
DOI: 10.1007/978-3-642-30054-7_20.

- [ACK10] E. Altman, S. Caron, and G. Kesidis.
Application Neutrality and a Paradox of Side Payments. Tech. rep. 2010.
doi: 10 . 1145 / 1921233 . 1921245.
- [AD82] G. A. Akerlof and W. Dickens.
‘The Economic Consequences of Cognitive Dissonance’.
In: *The American Economic Review* 72.3 (1982).
- [AFR97] E. W. Anderson, C. Fornell, and R. T. Rust. ‘Customer Satisfaction, Productivity, and Profitability: Differences Between Goods and Services’.
In: *Marketing Science* 16.2 (1997), pp. 129–145.
- [Ake70] G. A. Akerlof. ‘The Market for “Lemons”: Quality Uncertainty and the Market Mechanism’.
In: *Quarterly Journal of Economics* 84.3 (1970), pp. 488–500.
- [AL08] F. Agboma and A. Liotta. ‘QoE-aware QoS Management’.
In: *Proc. of the Sixth International Conference on Advances in Mobile Computing and Multimedia*. ACM, 2008, pp. 111–116.
doi: 10 . 1145 / 1497185 . 1497210.
- [All00] V. Allee. ‘Reconfiguring the Value Network’.
In: *Emerald Journal of Business Strategy* 21.4 (2000), pp. 36–39.
doi: 10 . 1108 / eb040103.
- [All08] V. Allee. ‘Value Network Analysis and Value Conversion of Tangible and Intangible Assets’.
In: *Emerald Journal of Intellectual Capital* 9.1 (2008), pp. 5–24.
doi: 10 . 1108 / 14691930810845777.
- [ALP06] D. Ariely, G. Loewenstein, and D. Prelec.
‘Tom Sawyer and the construction of value’.
In: *Journal of Economic Behavior & Organization* 60.1 (2006).
doi: 10 . 1016 / j . jeb0 . 2004 . 10 . 003.
- [Ami13] I. Amigo. ‘Technological and Economic Aspects for Quality of Service in Multidomain Alliances’.
PhD thesis. Telecom Bretagne, Université européenne de Bretagne, 2013.
- [And96] E. W. Anderson. ‘Customer Satisfaction and Price Tolerance’.
In: *Marketing Letters* 7.3 (1996), pp. 265–274.
doi: 10 . 1007 / BF00435742.
- [Atk70] A. B. Atkinson. ‘On the Measurement of Inequality’.
In: *Journal of economic theory* 2.3 (1970), pp. 244–263.
doi: 10 . 1016 / 0022 - 0531 (70) 90039 - 6.
- [ATM93] ATM Forum. *The User Network Interface (UNI), v3.1*.
Cambridge University Press, 1993. ISBN: 0-13-393828-X.

- [Bal+13] A. Balachandran, V. Sekar, A. Akella, S. Seshan, I. Stoica, and H. Zhang. 'Developing a Predictive Model of Quality of Experience for Internet Video'. In: *Proc. of the ACM SIGCOMM Computer Communication Review*. Vol. 43. 4. ACM. 2013, pp. 339–350. doi: 10.1145/2486001.2486025.
- [Bal07] P. Ballon. 'Business Modelling Revisited: The Configuration of Control and Value'. In: *INFO: The Journal of Policy, Regulation and Strategy for Telecommunications, Information and Media* 9.5 (2007), pp. 6–19. doi: 10.1108/14636690710816417.
- [Bal09] P. Ballon. 'The Platformisation of European Mobile Industry'. In: *Communications & Strategies* 75 (2009), pp. 15–34.
- [BC08] A. Biem and N. Caswell. 'A Value Network Model for Strategic Analysis'. In: *Proc. of the 41st Annual Hawaii International Conference on System Sciences*. 2008, pp. 361–367. doi: 10.1109/HICSS.2008.43.
- [BDM64] G. Becker, M. DeGroot, and J. Marschak. 'Measuring Utility by a Single-response Sequential Method'. In: *Behavioral Science* 9.3 (1964), pp. 226–232. doi: 10.1002/bs.3830090304.
- [Bec62] G. Becker. 'Irrational Behavior and Economic Theory'. In: *Journal of Political Economy* 86.169–217 (1962).
- [Ben07] Y. Benkler. *The Wealth of Networks – How Social Production Transforms Markets and Freedom*. Yale University Press, 2007. ISBN: 9780300125771.
- [Ber96] D. Bernoulli. *Versuch einer neuen Theorie der Wertbestimmung von Glücksfällen*. Leipzig. 1896.
- [BL99] R. N. Bolton and K. N. Lemon. 'A Dynamic Model of Customer's Usage of Services: Usage as Antecedent and Consequence of Satisfaction'. In: *Journal of Marketing Research* 36 (1999), pp. 171–186.
- [BM82] J. A. Bondy and U. Murty. *Graph Theory With Applications*. Fifth Printing. Elsevier Science Ltd/North-Holland, 1982. ISBN: 0444194517.
- [BO12] G. Blocq and A. Orda. 'How Good is Bargained Routing?'. In: *Proc. of the INFOCOM'12*. IEEE, 2012, pp. 2453–2461. doi: 10.1109/INFCOM.2012.6195636.
- [Bor+11] F. von Bornstaedt, M. Roettgermann, I. Korthals, F. T. Johansen, and H. Lonsethagen. 'The Sending Party Network Pays: A First Step Towards End-to-end Quality of Service'. In: *Proc. of 15th International Conference on Intelligence in Next Generation Networks (ICIN)*. IEEE, 2011. doi: 10.1109/ICIN.2011.6081074.

- [Bos07] G. Bos. 'QoS support using DiffServ'. In: *6th TSConIT* (2007).
- [BR06] S. Bessler and P. Reichl. 'A Network Provisioning Scheme Based on Decentralized Bandwidth Auctions'.
In: *Springer Telecommunications Planning: Innovations in Pricing, Network Design and Management* (2006), pp. 89–100.
doi: 10.1007/0-387-29234-9_5.
- [Bre+84] L. Breiman, J. H. Friedman, C. A. Stone, and R. A. Olshen.
Classification and Regression Trees. first. Chapman and Hall, 1984.
ISBN: 978-0412048418.
- [Cas13] T. R. Casey. 'Evolution of wireless access provisioning: Understanding and managing value system structures and dynamics'.
PhD thesis. Aalto University, School of Electrical Engineering, Department of Communications and Networking, 2013.
- [CBD86] R. G. Cummings, D. S. Brookshire, and S. W. D.
Valuing Public Goods: An Assessment of the Contingent Valuation Method. Rowman and Allanheld, 1986.
- [CDS01] C. Courcoubetis, M. P. Dramitinos, and G. D. Stamoulis.
'An Auction Mechanism for Bandwidth Allocation over Paths'.
In: *Proc. of the 17th International Teletraffic Congress (ITC)*. 2001.
doi: 10.1016/S1388-3437(01)80200-6.
- [CEL14] CELTIC QuEEN. *Deliverable D5.100: Economical Aspects, Business Models of New QoE-aware Services*. Ed. by P. Zwickl and T. Mäki. 2014.
- [Cis12] Cisco. *Cisco Visual Networking Index: Forecast and Methodology, 2011–2016*. Whitepaper. 2012.
- [CJ87] T. A. Cameron and M. D. James. 'Estimating Willingness To Pay from Survey Data: An Alternative Pre-Test-Market Evaluation Procedure'.
In: *Journal of Marketing Research* (1987), pp. 389–395.
doi: 10.2307/3151386.
- [CKA10] S. Caron, G. Kesidis, and E. Altman.
'Application Neutrality and a Paradox of Side Payments'.
In: *Proc. of the Re-Architecting the Internet (ReARCH '10)*. ACM, 2010.
- [CKY98] J. Contreras, M. Klusch, and J. Yen. 'Multi-Agent Coalition Formation in Power Transmission Planning: A Bilateral Shapley Value Approach'.
In: *Proc. of the Conference on Artificial Intelligence Planning Systems (AIPS)*. 1998, pp. 19–26.
- [Cow07] S. Cowan. *The Welfare Effects of Third-degree Price Discrimination with Non-linear Demand Functions*. Tech. rep.
Oxford University – Department of Economics, 2007.

- [CR10] M. Carbone and L. Rizzo. ‘Dummynet Revisited’. In: *ACM SIGCOMM Computer Communication Review* 40.2 (2010), pp. 12–20.
doi: 10.1145/1764873.1764876.
- [Cro07] J. Crowcroft.
‘Net Neutrality: The Technical Side of the Debate: A White Paper’. In: *ACM SIGCOMM Computer Communication Review* 37.1 (2007), pp. 49–55.
doi: 10.1145/1198255.1198263.
- [Cro08] P. Crocioni. *Net Neutrality: Few Thoughts and Many Questions*. Ofcom Whitepaper. 2008.
- [CT12] T. Casey and J. Töyli. ‘Dynamics of Two-sided Platform Success and Failure: An analysis of Public Wireless Local Area Access’. In: *Technovation* 32.2012 (2012), pp. 703–716.
doi: 10.1016/j.technovation.2012.08.003.
- [Cuz+12] A. Cuzzocrea, A. Papadimitriou, D. Katsaros, and Y. Manolopoulos. ‘Edge Betweenness Centrality: A novel Algorithm for QoS-based Topology Control over Wireless Sensor Networks’. In: *Journal of Network and Computer Applications* 35.4 (2012), pp. 1210–1217.
doi: 10.1016/j.jnca.2011.06.001.
- [CW03] C. Courcoubetis and R. Weber.
Pricing Communication Networks: Economics, Technology and Modelling. Wiley Interscience Series in Systems and Optimization. Wiley, 2003.
- [CZ09] F. Cribari-Neto and A. Zeileis. *Beta Regression in R*. Tech. rep. publisher=Department of Statistics and Mathematics, WU Vienna University of Economics and Business, 2009.
- [Dav02] C. S. Davis. *Statistical Methods for the Analysis of Repeated Measurements*. Springer, 2002. doi: 978-0-387-21573-0.
- [De +10] K. De Moor et al. ‘Linking Users’ Subjective QoE Evaluation to Signal Strength in an IEEE 802.11b/g Wireless LAN Environment’. In: *EURASIP Journal on Wireless Communications and Networking* 1 (2010).
doi: 10.1155/2010/541568.
- [DKS90] A. Demers, S. Keshav, and S. Shenkar.
‘Analysis and Simulation of a Fair Queuing Algorithm’. In: *Journal of Internetworking Research and Experience* 1 (1990).
doi: 10.1145/75247.75248.
- [Dob+14] O. Dobrijevic, A. J. Kassler, L. Skorin-Kapov, and M. Matijasevic. ‘Q-POINT: QoE-Driven Path Optimization Model for Multimedia Services’. In: *Proc. of the 12th International Conference on Wired & Wireless Internet Communications (WWIC’14)*. LNCS. Springer, 2014.
doi: 10.1007/978-3-319-13174-0_11.

- [Egg+11] C. Egger, M. Happenhofer, J. Fabini, and P. Reichl.
‘BIQINI – A Flow-Based QoS Enforcement Architecture for NGN Services’.
In: *Proc. of Testbeds and Research Infrastructures. Development of Networks and Communities*. 2011, pp. 653–667.
doi: 10 . 1007 / 978 - 3 - 642 - 17851 - 1_66.
- [Egg14a] S. Egger. ‘QoE Analysis for Interactive Internet Applications in the Presence of Delay’. PhD thesis. Graz University of Technology, Institute of Signal Processing and Speech Communication, 2014.
- [Egg14b] S. Egger. ‘QoE Analysis for Interactive Internet Applications in the Presence of Delay’. PhD thesis. Graz University of Technology, Signal Processing and Speech Communication Laboratory, 2014.
- [EH95] N. Economides and C. Himmelberg.
‘Toward A Competitive Telecommunication Industry: Selected Papers From the 1994 Telecommunications Policy Research Conference’. In: ed. by G. W. Brock. Lawrence Erlbaum Associates, Inc., Publishers, 1995. Chap. Critical Mass and Network Evolution in Telecommunications, pp. 47–63.
- [El +13] A. El Essaili, D. Schroeder, D. Staehle, M. Shehada, W. Kellerer, and E. Steinbach.
‘Quality-of-Experience driven Adaptive HTTP Media Delivery’.
In: *Proc. of the IEEE International Conference on Communications (ICC’13)*. IEEE. 2013, pp. 2480–2485. doi: 10 . 1109 / ICC . 2013 . 6654905.
- [ET12] N. Economides and J. Tåg.
‘Net Neutrality on the Internet: A Two-sided Market Analysis’.
In: *Elsevier Information Economics and Policy* 23.2 (2012), pp. 91–104.
doi: 10 . 1016 / j . infoecopol . 2012 . 01 . 001.
- [Eur02] European Parliament and the Council. ‘Directive 2002/22/EC : On Universal Service and Users’ Rights Relating to Electronic Communications Networks and Services (Universal Service Directive)’.
In: *Official Journal of the European Communities* (2002).
- [Fer04] C.-N. F. Ferrari SLP.
‘Beta Regression for Modelling Rates and Proportions’.
In: *Journal of Applied Statistics* 31.7 (2004), pp. 799–815.
doi: 10 . 1080 / 0266476042000214501.
- [Fes57] L. Festinger. *A theory of cognitive dissonance*. Stanford University Press, 1957.

- [FHT10] M. Fiedler, T. Hoßfeld, and P. Tran-Gia. ‘A Generic Quantitative Relationship between Quality of Experience and Quality of Service’. In: *IEEE Network Special Issue on Improving QoE for Network Services* 24.2 (2010). doi: 10.1109/MNET.2010.5430142.
- [FI06] P. Frank and J. Incera. ‘A Neural Network based Test Bed for Evaluating the Quality of Video Streams in IP Networks’. In: *Proc. of the 2006 Electronics, Robotics and Automotive Mechanics Conference*. Vol. 1. IEEE, 2006, pp. 178–183. doi: 10.1109/CERMA.2006.7.
- [Fin93] D. P. Finkelmann. ‘Crossing the ‘Zone of Indifference’’. In: *Marketing Management* 2.3 (1993), pp. 22–32.
- [FMS88] A. Fisher, G. H. McClelland, and W. D. Schulze. ‘Measures of Willingness to Pay versus Willingness to Accept: Evidence, Explanations and Potential Reconciliation’. In: *Amenity Resource Valuation: Integrating Economics with Other Disciplines*. Ed. by G. L. Peterson, B. L. Driver, and R. Gregory. Venture, 1988, pp. 127–134.
- [Fox11] J. Fox. *The Myth of the Rational Market: A History of Risk, Reward, and Delusion on Wall Street*. HarperBusiness, 2011. ISBN: 0060599030.
- [FP502] FP5 Project M3I, IST–1999–11429. *Deliverable 15/2 – M3I User Experiment Results*. Ed. by D. Hands. 2002.
- [FP711] FP7 ETICS. *Deliverable D3.3: Financial/Economic Dynamic Analysis*. 2011.
- [FP712a] FP7 ETICS. *Deliverable D3.4: Business, Legal and Socioeconomic Impacts*. 2012.
- [FP712b] FP7 ETICS. *Deliverable D4.4: Final ETICS Architecture and Functional Entities High Level Design*. Ed. by P. Zwickl and H. Weisgrab. 2012.
- [Fre06] L. Freeman. *The Development of Social Network Analysis*. Empirical Press, 2006. ISBN: 1-59457-714-5.
- [Fre79] L. Freeman. ‘Centrality in Social Networks: Conceptual Clarification’. In: *Social Networks* 1.3 (1979), pp. 215–239. doi: 10.1016/0378-8733(78)90021-7.
- [Fri07] M. Friedman. *Price Theory*. Aldine Transaction, 2007. ISBN: 020230969X.
- [FS48] M. Friedman and L. Savage. ‘The Utility Analysis of Choices Involving Risk’. In: *The Journal of Political Economy* LVI.4 (1948), pp. 279–304.

- [Fu+13] B. Fu, D. Staehle, G. Kunzmann, E. Steinbach, and W. Kellerer. 'QoE-aware Priority Marking and Traffic Management for H.264/SVC-based Mobile Video Delivery'. In: *Proc. of the Eighth ACM Workshop on Performance Monitoring and Measurement of Heterogeneous Wireless and Wired Networks*. ACM, 2013, pp. 173–180. doi: 10.1145/2512840.2512865.
- [Ful+98] E. Fulp, M. Ott, D. Reininger, and D. Reeves. 'Paying for QoS: An Optimal Distributed Algorithm for Pricing Network Resources'. In: *Proc. of the IEEE sixth international Workshop on Quality of Service (WQoS98)*. 1998, pp. 75–84. doi: 10.1109/IWQOS.1998.675223.
- [GA01] J. Gordijn and H. Akkermans. 'E3-value: Design and Evaluation of e-Business Models'. In: *IEEE Intelligent Systems* 16.4 (2001), pp. 11–17.
- [Gee+10] D. Geerts et al. 'Linking an integrated framework with appropriate methods for measuring QoE'. In: *Proc. of the Second International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE. 2010, pp. 158–163. doi: 10.1109/QOMEX.2010.5516292.
- [GGB94] J. B. Gotlieb, D. Grewal, and S. W. Brown. 'Consumer Satisfaction and Perceived Quality: Complementary or Divergent Constructs?'. In: *Journal of Applied Psychology* 79.6 (1994), pp. 875–885.
- [Ghe11] A. Ghezzi. 'A Reference Framework for Strategy Analysis in the Mobile Telecommunications Industry'. PhD thesis. Politecnico di Milano, Department of Management, Economics and Industrial Engineering, 2011.
- [Ghe12] A. Ghezzi. 'A Proposal of Business Model Design Parameters for Future Internet Carriers'. In: *Proc. of the Networking 2012 Workshops*. Springer LNCS, 2012, pp. 72–79. doi: 10.1007/978-3-642-30039-4_9.
- [GKT02] A. Gelman, J. N. Katz, and F. Tuerlinckx. 'The Mathematics and Statistics of Voting Power'. In: *Statistical Science* (2002), pp. 402–435.
- [GLS09] R. Gulati, D. Lavie, and H. Singh. 'The Nature of Partnering Experience and the Gains from Alliances'. In: *Strategic Management Journal* 30.11 (2009), pp. 1213–1233.
- [GMS00] R. Gibbens, R. Mason, and R. Steinberg. 'Internet Service Classes under Competition'. In: *IEEE Signal Processing Magazine* 18.12 (2000), pp. 2490–2498. doi: 10.1109/49.898732.
- [GNZ00] R. Gulati, N. Nohria, and A. Zaheer. 'Strategic Networks'. In: *Strategic Management Journal* 21 (2000), pp. 203–215.

- [Góm+13] G. Gómez, J. Lorca, R. García, and Q. Pérez. 'Towards a QoE-Driven Resource Control in LTE and LTE-A Networks'. In: *Journal of Computer Networks and Communications* 2013.2 (2013), pp. 1–15.
- [GS69] R. Gabriel and R. Sokal. 'A New Statistical Approach to Geographic Variation Analysis'. In: *Systematic Zoology* 18.3 (1969), pp. 259–78. doi: 10.2307/2412323.
- [Gub+13] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami. 'Internet of Things (IoT): A vision, architectural elements, and future directions'. In: *Future Generation Computer Systems* 29.7 (2013), pp. 1645–1660. doi: 10.1016/j.future.2013.01.010.
- [Ham94] D. M. Hamby. 'A Review of Techniques for Parameter Sensitivity Analysis of Environmental Models'. In: *Environmental Monitoring and Assessment* 32.2 (1994), pp. 135–154. doi: 10.1007/BF00547132.
- [Han91] M. W. Hanemann. 'Willingness to Pay and Willingness to Accept: How Much Can They Differ?'. In: *The American Economic Review* 81.3 (1991), pp. 635–647.
- [Har68] G. Hardin. 'The Tragedy of the Commons'. In: *Science* 162.3859 (1968), pp. 1243–1248. doi: 10.1126/science.162.3859.1243.
- [HB13] H. Haddadi and O. Bonaventure, eds. *Recent Advances in Networking*. Vol. 1. ACM SIGCOMM. ACM, 2013.
- [HD95] N. Hanna and R. H. Doge. *Pricing: Policies and Procedures*. Prentice Hall, 1995. ISBN: 978-0333694138.
- [Hei04] M. Heidelberger. *Nature from Within: Gustav Theodor Fechner and his Psychophysical Worldview*. Trans. by C. Kloor. University of Pittsburg Press, 2004.
- [Hir65] A. Hirschmann. 'Obstacles to Development: A Classification and Quasi-Vanishing Act'. In: *Economic Development and Cultural Change* 13.4 (1965).
- [HKH05] C. Homburg, N. Koschate, and W. D. Hoyer. 'Do Satisfied Customers Really Pay More? A Study of the Relationship Between Customer Satisfaction and Willingness to Pay'. In: *Journal of Marketing* (2005), pp. 84–96. doi: 10.1509/jmkg.69.2.84.60760.
- [Hog07] C. Hogendorn. 'Broadband Internet: Net neutrality versus Open Access'. In: *International Economics and Economic Policy* 4.2 (2007), pp. 185–208. doi: 10.1007/s10368-007-0084-6.

- [Hoß+11] T. Hoßfeld, S. Biedermann, R. Schatz, A. Platzer, S. Egger, and M. Fiedler. 'The Memory Effect and its Implications on Web QoE Modeling'. In: *Proc. of the 23rd International Teletraffic Congress*. International Teletraffic Congress. 2011, pp. 103–110. ISBN: 978-0-9836283-0-9.
- [HS89] H. Hakansson and I. Snehota. 'No Business is an Island: The Network Concept of Business Strategy'. In: *Scandinavian Journal of Management* (1989), pp. 187–200. doi: 10.1016/0956-5221(89)90026-2.
- [HSP2.0] *Wi-Fi CERTIFIED Passpoint™ (Release 1) 6 Deployment Guidelines*. Wi-Fi Alliance, Hotspot 2.0 Technical Task Group, 2012.
- [HW05] L. He and J. Walrand. 'Pricing and Revenue Sharing Strategies for Internet Service Providers'. In: *Proc. of the IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies 1* (2005), 205–216 vol. 1. doi: 10.1109/INFCOM.2005.1497892.
- [IM94] M. Ivaldi and D. Martimort. 'Competition under Nonlinear Pricing'. In: *Annales d'Economie et de Statistique* (1994), pp. 71–114.
- [ISM13] K. Ivešić, L. Skorin-Kapov, and M. Matijašević. 'Admission Control for Adaptive Multimedia Services Based on User and Service Related Knowledge'. In: *Proc. of the EUROCON*. IEEE, 2013. doi: 10.1109/EUROCON.2013.6625019.
- [ISM14] K. Ivešić, L. Skorin-Kapov, and M. Matijašević. 'Cross-layer QoE-driven Admission Control and Resource Allocation for Adaptive Multimedia Services in LTE'. In: *Journal of Network and Computer Applications* (2014). doi: 10.1016/j.jnca.2014.09.010.
- [Jai06] R. Jain. 'Internet 3.0: Ten Problems with Current Internet Architecture and Solutions for the Next Generation'. In: *Proc. of the Military Communications Conference (MILCOM'06)*. IEEE, 2006. doi: 10.1109/MILCOM.2006.301995.
- [Jar88] J. Jarillo. 'On Strategic Networks'. In: *Strategic Management Journal* 9.1 (1988), pp. 31–44. doi: 10.1002/smj.4250090104.
- [JG10] S. Jordan and A. Ghosh. 'A Framework for Classification of Traffic Management Practices as Reasonable or Unreasonable'. In: *ACM Transactions on Internet Technology* 10.3 (2010), pp. 12–23. doi: 10.1145/1852096.1852100.

- [JK06] F. Jouault and I. Kurtev. 'On the Architectural Alignment of ATL and QVT'. In: *Proc. of the 2006 ACM Symposium on Applied Computing*. ACM, 2006, pp. 1188–1195.
- [JKA06] M. Jacobides, T. Knudsenb, and M. Augierc. 'Benefiting from Innovation: Value Creation, Value Appropriation and the Role of Industry Architectures'. In: *Research Policy* 35.8 (2006), pp. 1200–1221. doi: 10.1016/j.respol.2006.09.005.
- [JL10] J. Joskowicz and J. Lopez Ardao. 'A General Parametric Model for Perceptual Video Quality Estimation'. In: *Proc. of the International Workshop on Communications Quality and Reliability (CQR)*. IEEE, 2010. doi: 10.1109/CQR.2010.5619912.
- [JS14] S. Jordan and G. Schaffer. 'User and ISP Rights of Device Attachment and Device Management'. In: *ACM Transactions on Internet Technology* 13 (2014), 6:1–6:20. doi: 10.1145/2513227.
- [JSS03] A. Juttner, I. Szabó, and A. Szentesi. 'On Bandwidth Efficiency of the Hose Resource Management Model in Virtual Private Networks'. In: *Proc. of the 22nd Annual Joint Conference of the IEEE Computer and Communications (INFOCOM)*. Vol. 1. IEEE, 2003, pp. 386–395. doi: 10.1109/INFCOM.2003.1208690.
- [JW08] R. Jain and J. Walrand. 'An Efficient Mechanism for Network Bandwidth Auction'. In: *Proc. of the Network Operations and Management Symposium Workshop (NOMS'08)*. IEEE, 2008. doi: 10.1109/NOMSW.2007.38.
- [KA03] N. J. Keon and G. Anandalingam. 'Optimal Pricing for Multiple Services in Telecommunications Networks Offering Quality-of-Service Guarantees'. In: *IEEE/ACM Transactions on Networking (TON)* 11.1 (Feb. 2003). doi: 10.1109/TNET.2002.808409.
- [Kah12] D. Kahneman. *Thinking, Fast and Slow*. Penguin, 2012. ISBN: 978-0374533557.
- [Kel03] F. Kelly. 'Fairness and Stability of End-to-End Congestion Control'. In: *European Journal of Control* 9.2 (2003), pp. 159–176. doi: 10.3166/ejc.9.159-176.
- [Kes97] S. Keshav. *An Engineering Approach to Computer Networking*. Addison-Wesley, 1997. ISBN: 978-0201634426.
- [Kil08] K. Kilkki. 'Quality of Experience in Communications Ecosystems'. In: *Journal of Universal Computer Science* 14.5 (2008), pp. 615–624.

- [Kil12] K. Kilkki. *An Introduction to Communications Ecosystems*. 2012. ISBN: 978-1478253303.
- [Kim06] W. Kim. 'Pricing Model for Internet-based Multi-class Video-on-Demand (VoD) Services'. In: *Telecommunication System* 33 (2006), pp. 317–331. doi: 10.1007/s11235-006-9017-x.
- [Kle76] L. Kleinrock. *Queuing Systems*. Vol. 2 Computer Applications. Wiley, 1976. ISBN: 978-0471491118.
- [KR10] A. Kirmani and A. R. Rao. 'No Pain, No Gain: A Critical Review of the Literature on Signaling Unobservable Product Quality.' In: *The Journal of Marketing* (2010), pp. 66–79. doi: 10.1509/jmkg.64.2.66.18000.
- [KRD12] C. Keimel, A. Redl, and K. Diepold. 'The TUM High Definition Video Datasets'. In: *Proc. of the Fourth International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 2012, pp. 97–102. doi: 10.1109/QoMEX.2012.6263865.
- [Kri91] A. Krishna. 'Effect of Dealing Patterns on Consumer Perceptions of Deal Frequency and Willingness to Pay'. In: *Journal of Marketing Research* (1991), pp. 441–451. doi: 10.2307/3172784.
- [KS84] J. L. Knetsch and J. A. Sinden. 'Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value'. In: *Quarterly Journal of Economics* 99.3 (1984), pp. 507–521.
- [KT79] D. Kahneman and A. Tversky. 'Prospect Theory: An Analysis of Decision under Risk'. In: *Econometrica* 47.2 (1979), pp. 263–292. doi: 10.2307/1914185.
- [Lag+12] K. U. R. Laghari, O. Issa, F. Speranza, and T. Falk. 'Quality-of-Experience Perception for Video Streaming Services: Preliminary Subjective and Objective Results'. In: *Proc. of the Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC)*. IEEE, 2012, 1–9.
- [LC12] K. U. R. Laghari and K. Connelly. 'Toward Total Quality of Experience: A QoE Model in a Communication Ecosystem'. In: *IEEE Communications Magazine* 50.4 (2012), pp. 58–65. doi: 10.1109/MCOM.2012.6178834.

- [LKC12] K. U. R. Laghari, I. Khan, and N. Crespi. ‘Quantitative and Qualitative Assessment of QoE for Multimedia Services in Wireless Environment’. In: *Proc. of the 4th Workshop on Mobile Video (MoVid)*. Chapel Hill, North Carolina: ACM, 2012, pp. 7–12. doi: 10.1145/2151677.2151680.
- [LL01] M. Lemley and L. Lessig. ‘The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era’. In: *UCLA Law Review* 48 (2001), p. 925.
- [LMP12] P. Le Callet, S. Möller, and A. Perkis, eds. *Qualinet White Paper on Definitions of Quality of Experience. European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003)*. Whitepaper. Qualinet (COST Action IC 1003), 2012.
- [LS98] A. A. Lazar and N. Semret. *Design and Analysis of the Progressive Second Price Auction for Network Bandwidth Sharing*. Tech. rep. Columbia University, 1998.
- [LY85] R. I. Lerman and S. Yitzhaki. ‘Income Inequality Effects by Income Source: A New Approach and Applications to the United States’. In: *The Review of Economics and Statistics* 67.1 (Feb. 1985), pp. 151–7. doi: 10.2307/1928447.
- [Mar06] M. de Marin de Montmarin. ‘A Result Similar to the Odlyzko’s ‘Paris Metro Pricing’’. In: *Applied Economics* 38.15 (2006), pp. 1821–1824. doi: 10.1080/00036840500426967.
- [Mar07] G. M. Maranell, ed. *A Sourcebook for Behavioral Scientists*. 2007.
- [MCN07] O. Markaki, D. Charilas, and D. Nikitopoulos. ‘Enhancing Quality of Experience in Next Generation Networks through Network Selection Mechanisms’. In: *Proc. of the 18th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2007. doi: 10.1109/PIMRC.2007.4394198.
- [Met15] F. Metzger. ‘Evaluating Reliable Streaming in Mobile Networks’. PhD thesis. University of Vienna, Faculty of Computer Science, 2015.
- [MKP98] R. Madhavan, B. Koka, and J. Prescott. ‘Networks in Transition: How Industry Events (Re)Shape Interfirm Relationships’. In: *Strategic Management Journal* 19 (1998), pp. 439–459.
- [MLT12] C. Müller, S. Lederer, and C. Timmerer. ‘An Evaluation of Dynamic Adaptive Streaming over HTTP in Vehicular Environments’. In: *Proc. of the Fourth Workshop on Mobile Video*. ACM. 2012, pp. 37–42. doi: 10.1145/2151677.2151686.

- [MMC05] R. Martin, M. Menth, and J. Charzinski.
‘Comparison of Border-to-border Budget based Network Admission Control and Capacity Overprovisioning’.
In: *Proc. of the 2005 NETWORKING 2005*. Springer, 2005, pp. 1056–1068.
doi: 10 . 1007 / 11422778 _85.
- [MMC06] M. Menth, R. Martin, and J. Charzinski.
‘Capacity Overprovisioning for Networks with Resilience Requirements’.
In: *ACM SIGCOMM Computer Communication Review* 36.4 (2006),
pp. 87–98. doi: 10 . 1145 / 1151659 . 1159925.
- [Mon90] K. B. Monroe. *Pricing: Making Profitable Decisions*. McGraw-Hill, 1990.
ISBN: 978-0072528817.
- [MSW09] J. Musacchio, G. Schwartz, and J. Walrand.
‘A Two-sided Market Analysis of Provider Investment Incentives with an Application to the Net-neutrality Issue’.
In: *Review of Network Economics* 8.1 (2009).
- [MV95] J. MacKie-Mason and H. Varian. ‘Public Access to the Internet’. In:
ed. by B. Kahin and J. Keller. MIT Press, 1995. Chap. Pricing the Internet.
ISBN: 978-0262611183.
- [Nel70] P. Nelson. ‘Information and Consumer Behavior’.
In: *Journal of Political Economy* 78.2 (1970), pp. 311–329.
- [Ngo99] P.-V. Ngobo. ‘Decreasing Returns in Customer Loyalty: Does It Really Matter to Delight the Customers?’
In: *Advances in Consumer Research* 26.1 (1999), pp. 469–476.
- [Nis+07] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani.
Algorithmic Game Theory. Vol. 1.
Cambridge University Press Cambridge, 2007. ISBN: 978-0521872829.
- [NJ04] A. Norrman and U. Jansson. ‘Ericsson’s Proactive Supply Chain Risk Management Approach after a Serious Sub-supplier Accident’.
In: *International Journal of Physical Distribution & Logistics Management* 34.5 (2004), pp. 434–456. doi: 10 . 1108 / 09600030410545463.
- [Nor10] W. B. Norton. *A Business Case for Peering in 2010*. White Paper.
DrPeering International, 2010.
- [NR09] L. Nussbaum and O. Richard.
‘A Comparative Study of Network Link Emulators’.
In: *Proc. of the 2009 Spring Simulation Multiconference*.
Society for Computer Simulation International. 2009, p. 85.

- [NR98] R. Normann and R. Ramirez. *Designing Interactive Strategy: From the Value Chain to the Value Constellation*. John Wiley & Sons, 1998.
ISBN: 978-0471986072.
- [NS08] M. Nagarajan and G. Sošić. ‘Game-theoretic Analysis of Cooperation among Supply Chain Agents: Review and Extensions’.
In: *European Journal of Operational Research* 187.3 (2008), pp. 719–745.
DOI: 10.1016/j.ejor.2006.05.045.
- [NS11] P. Neelakanta and R. Sardenberg. ‘Consumer Benefit versus Price Elasticity of Demand: A Nonlinear Complex System Model of Pricing Internet Services on QoS-centric Architecture’.
In: *Netnomics* 12.1 (2011), pp. 31–60.
DOI: 10.1007/s11066-010-9055-0.
- [NW05] E. Ngai and F. Wat. ‘Fuzzy Decision Support System for Risk Analysis in E-commerce Development’.
In: *Decision Support Systems* 40 (2005), pp. 235–255.
DOI: 10.1016/j.dss.2003.12.002.
- [Od99] A. Odlyzko. ‘Paris Metro Pricing for the Internet’.
In: *Proc. of the First ACM Conference on Electronic Commerce*. ACM. 1999, pp. 140–147. DOI: 10.1145/336992.337030.
- [OFC10] *Traffic Management and ‘net neutrality’*. Ofcom Discussion Document. Ofcom, 2010.
- [OFC11] *Ofcom’s approach to net neutrality*. Ofcom Whitepaper. Ofcom, 2011.
- [Ogu95] A. Ogu. ‘Rethinking self-regulation’.
In: *Oxford Journal of Legal Studies* (1995), pp. 97–108.
- [OS12] O. Oyman and S. Singh.
‘Quality of Experience for HTTP Adaptive Streaming Services’.
In: *IEEE Communications Magazine* 50.4 (2012), pp. 20–27.
DOI: 10.1109/MCOM.2012.6178830.
- [Pal10] N. Palan. ‘Measurement of Specialization—The Choice of Indices’.
In: *FIW* (2010).
- [PB93] J. W. Payne, J. R. Bettman, and E. J. Johnson. *The adaptive decision maker*. Cambridge University Press, 1993.
- [PD10] P. Pathak and R. Dutta. ‘Using Centrality-based Power Control for Hot-spot Mitigation in Wireless Networks’. In: *Proc. of the IEEE Globecom*. IEEE, 2010. DOI: 10.1109/GLOCOM.2010.5683097.

- [Pin+12] M. H. Pinson et al. 'The Influence of Subjects and Environment on Audiovisual Subjective Tests: An International Study'. In: *IEEE Journal of Selected Topics in Signal Processing* 6.6 (2012), pp. 640–651. doi: 10.1109/JSTSP.2012.2215306.
- [Pin11] C. Pinheiro. *Social Network Analysis in Telecommunications*. John Wiley & Sons, 2011. ISBN: 978-0470647547.
- [Pir+11] G. Piro, L. A. Grieco, G. Boggia, F. Capozzi, and P. Camarda. 'Simulating LTE cellular Systems: An Open-source Framework'. In: *IEEE Transactions on Vehicular Technology* 60.2 (2011), pp. 498–513. doi: 10.1109/TVT.2010.2091660.
- [PL13] A. Pradayrol and D. Levy. *European Telecom Operators: 4G – Going Faster, but Where?* White Paper. Arthur D. Little and Exane BNP Paribas, 2013.
- [PMH06] A. Perkis, S. Munkeby, and O. I. Hillestad. 'A Model for Measuring Quality of Experience'. In: *Proc. of the 7th Nordic Signal Processing Symposium (NORSIG)*. IEEE, 2006, pp. 198–201. doi: 10.1109/NORSIG.2006.275209.
- [Por79] M. E. Porter. 'How Competitive Forces Shape Strategy'. In: *Harvard Business Review* 102 (1979).
- [Por80] M. Porter. *Competitive Strategy: Techniques for Analyzing Industries and Competitors: with a new Introduction*. Free Press, 1980.
- [Por85] M. E. Porter. *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press, 1985.
- [PR06] J. Peppard and A. Rylander. 'From Value Chain To Value Network: Insight for Mobile Operators'. In: *European Management Journal* 24.2–3 (2006), pp. 128–141.
- [PS15] C. Piesbergen and N. Sarubin. 'Erfassung von kognitiver Dissonanz nach Kaufentscheidungen'. In: *Preprint* (2015).
- [PW04] M. Pinson and S. Wolf. 'A new Standardized Method for Objectively Measuring Video Quality'. In: *IEEE Transaction on Broadcasting* 50.3 (2004), pp. 312–322. doi: 10.1109/TBC.2004.834028.
- [Qua+00] G. Quadros, A. Alves, E. Monteiro, and F. Boavida. 'How Unfair can Weighted Fair Queuing be?' In: *Proc. of the Fifth IEEE Symposium on Computers and Communications (ISCC)*. IEEE, 2000, pp. 779–784. doi: 10.1109/ISCC.2000.860738.

- [Raw09] J. Rawls. *A Theory of Justice*. Harvard university press, 2009.
doi: 978-0674000780.
- [Rei+10] P. Reichl, S. Egger, R. Schatz, and A. D’Alconzo. ‘The Logarithmic Nature of QoE and the Role of the Weber-Fechner Law in QoE Assessment’. In: *Proc. of the International Conference on Communications (ICC 2010)*. IEEE, 2010. doi: 10.1109/ICC.2010.5501894.
- [Rei07] P. Reichl. ‘From ‘Quality-of-Service’ and ‘Quality-of-Design’ to ‘Quality-of-Experience’: A Holistic View on Future Interactive Telecommunication Services’. In: *Proc. of the 15th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. IEEE, 2007, pp. 1–6. doi: 10.1109/SOFTCOM.2007.4446062.
- [Rei10] P. Reichl. ‘From Charging for Quality of Service to Charging for Quality of Experience’. In: *Annals of Telecommunications* 65.3–4 (2010), pp. 189–199.
doi: 10.1007/s12243-009-0144-8.
- [RH06] P. Reichl and F. Hammer. ‘Charging for Quality-of-Experience—a New Paradigm for Pricing IP-based Services’. In: *Proc. of the Second ISCA Tutorial and Research Workshop on Perceptual Quality of Systems*. Citeseer, Berlin, Germany, 2006.
- [RHS03] P. Reichl, D. Hausheer, and B. Stiller.
‘The Cumulus Pricing Model as an Adaptive Framework for Feasible, Efficient, and User-friendly Tariffing of Internet Services’.
In: *Computer Networks* 43 (2003), pp. 3–24.
doi: 10.1016/S1389-1286(03)00252-4.
- [RNR08] M. Ries, O. Nemethova, and M. Rupp. ‘On the Willingness to Pay in Relation to Delivered Quality of Mobile Video Streaming’.
In: *Proc. of the International Conference on Consumer Electronics*. IEEE, 2008. doi: 10.1109/ICCE.2008.4588054.
- [Rot88] A. E. Roth. *The Shapley Value: Essays in Honor of Lloyd S. Shapley*. Cambridge University Press, 1988. ISBN: 978-0521021333.
- [RS04] L. E. Raileanu and K. Stoffel. ‘Theoretical Comparison between the Gini Index and Information Gain Criteria’.
In: *Annals of Mathematics and Artificial Intelligence* 41.1 (2004), pp. 77–93.
doi: 10.1023/B:AMAI.0000018580.96245.c6.
- [RS90] F. Reichheld and E. Sasser. ‘Zero Defections: Quality Comes to Services’.
In: *Harvard Business Review* 68 (1990), pp. 105–111.
- [RT06] J. Rochet and J. Tirole. ‘Two-Sided Markets : A Progress Report’.
In: *RAND Journal of Economics* 35.3 (2006), pp. 645–667.
doi: 10.1111/j.1756-2171.2006.tb00036.x.

- [RTR] *Positionspapier Netzneutralität*. White Paper. RTR, 2013.
- [Sal+13] M. Salo, M. Baldauf, P. Fröhlich, and S. Suetter. 'Peak Moments of Physical Mobile Interaction Techniques'. In: (2013).
- [Sch+13] R. Schatz, T. Hoßfeld, L. Janowski, and S. Egger. 'From Packets to People: Quality of Experience as a New Measurement Challenge'. In: *Data traffic monitoring and analysis*. Springer, 2013, pp. 219–263. doi: 10.1007/978-3-642-36784-7_10.
- [Sci41] T. Scitovsky. 'A Note on Welfare Propositions in Economics'. In: *Review of Economic Studies* IX (1941), pp. 77–88.
- [Sen+13] S. Sen, C. Joe-Wong, S. Ha, and M. Chiang. 'A Survey of Smart Data Pricing: Past Proposals, Current Plans, and Future Trends'. In: *ACM Computing Surveys* 46.2 (2013). doi: 10.1145/2543581.2543582.
- [SFC10] J. Shaikh, M. Fiedler, and D. Collange. 'Quality of Experience from User and Network Perspectives'. In: *Annals of Telecommunications–Annales des Télécommunications* 65.1-2 (2010), pp. 47–57. doi: 10.1007/s12243-009-0142-x.
- [Sha48] C. E. Shannon. 'A Mathematical Theory of Communication'. In: *The Bell System Technical Journal* 27 (1948), pp. 379–423, 623–656. doi: 10.1145/584091.584093.
- [Sha53] L. S. Shapley. 'Contributions to the Theory of Games'. In: ed. by W. Kuhn and A. Tucker. Vol. 2. *Annals of Mathematical Studies* v. 28. Princeton University Press, 1953. Chap. Value for n-Person Games, pp. 307–317. ISBN: 978-0691079370.
- [She95] S. Shenker. 'Fundamental Design Issues for the Future Internet'. In: *IEEE Journal on Selected Areas in Communications* 13.7 (1995), pp. 1176–1188. doi: 10.1109/49.414637.
- [Sho+94] J. F. Shogren, S. Y. Shin, D. J. Hayes, and J. B. Kliebenstein. 'Resolving Differences in Willingness to Pay and Willingness to Accept'. In: *The American Economic Review* (1994), pp. 255–270.
- [Sko07] L. Skorin-Kapov. 'A Framework for Service-Level End-to-End Quality of Service Negotiation and Adaptation'. PhD thesis. University of Zagreb, Faculty of Electrical Engineering and Computing, 2007.
- [SRC84] J. Saltzer, D. Reed, and D. Clark. 'End-to-End Arguments in System Design'. In: *ACM Transactions in Computer Science* (1984). doi: 10.1145/357401.357402.

- [SRV04] M. Samir, G. Rubino, and M. Varela.
‘Performance Evaluation of Real-time Speech through a Packet Network: A Random Neural Networks-based Approach’.
In: *Performance evaluation* 57.2 (2004), pp. 141–161.
doi: 10.1016/j.peva.2003.10.007.
- [SS14] A. Sackl and R. Schatz.
Modelling QoE and WTP in the context of VoD services. Tech. rep. FTW, 2014.
- [SS69] L. S. Shapley and M. Shubik. ‘On Market Games’.
In: *Journal of Economic Theory* 1.1 (1969), pp. 9–25.
doi: 10.1016/0022-0531(69)90008-8.
- [SSW08a] G. Schwartz, N. Shetty, and J. Walrand.
‘Impact of QoS on Internet User Welfare’.
In: *Internet and Network Economics* (2008), pp. 716–723.
doi: 10.1007/978-3-540-92185-1_77.
- [SSW08b] N. Shetty, G. Schwartz, and J. Walrand.
‘Modeling the Impact of QoS on Internet User Welfare and ISP Incentives’.
In: *Working Paper, Berkeley University* (2008).
- [SSW10] N. Shetty, G. Schwartz, and J. Walrand. ‘Internet QoS and Regulations’.
In: *IEEE/ACM Transactions on Networking (TON)* 18.6 (2010), pp. 1725–1737.
- [ST97] T. Sandler and J. Tschirhart. ‘Club theory: Thirty years later’.
In: *Public choice* 93.3-4 (1997), pp. 335–355.
doi: 10.1023/A:1017952723093.
- [Sta+10] N. Staelens et al. ‘Assessing Quality of Experience of IPTV and Video on Demand Services in Real-Life Environments’.
In: *IEEE Transactions on Broadcasting* 56.4 (2010), pp. 485–466.
doi: 10.1109/TBC.2010.2067710.
- [Ste61] S. S. Stevens. ‘To Honor Fechner and Repeal His Law’.
In: *Science* 133.3446 (1961), pp. 80–86. issn: 00368075.
doi: 10.1126/science.133.3446.80.
- [Sti10] J. E. Stiglitz.
Freefall: America, Free Markets, and the Sinking of the World Economy. W. W. Norton & Company, 2010. isbn: 0393075966.
- [Sto11] T. Stockhammer. ‘Dynamic Adaptive Streaming over HTTP: Standards and Design principles’.
In: *Proc. of the Second Annual ACM Conference on Multimedia Systems*. ACM, 2011, pp. 133–144. doi: 10.1145/1943552.1943572.

- [Str12] Strohmeier, Dominik AND Jumisko-Pyykko, S. AND Raake, Alexander. 'Toward Task-dependent Evaluation of Web-QoE: Free Exploration vs. "Who Ate What?"' In: *Proc. of the 2012 IEEE Globecom Workshops*. IEEE, 2012. doi: 10.1109/GLOCOMW.2012.6477771.
- [STY86] O. Stark, J. E. Taylor, and S. Yitzhaki. 'Remittances and Inequality'. In: *Wiley The Economic Journal* 96.383 (Sept. 1986), pp. 722–740. doi: 10.2307/2232987.
- [Sul+12] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand. 'Overview of the High Efficiency Video Coding (HEVC) Standard'. In: *IEEE Transactions on Circuits and Systems for Video Technology* 22.12 (2012), pp. 1649–1668. doi: 10.1109/TCSVT.2012.2221191.
- [SV02] M. Shreedhar and G. Varghese. 'Efficient Fair Queuing using Deficit Round Robin'. In: *ACM Transactions on Networking* 4.3 (2002). doi: 10.1109/90.502236.
- [SW03] M. Siller and J. Woods. 'Improving Quality of Experience for Multimedia Services by QoS arbitration on a QoE framework'. In: *Proc. of the 13th Packed Video Workshop 2003*. Citeseer, 2003.
- [Tee10] D. J. Teece. 'Business Models, Business Strategy and Innovation'. In: *Long Range Planning* 43.2–3 (2010), pp. 172–194. doi: 10.1016/j.lrp.2009.07.003.
- [Tee86] D. J. Teece. 'Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy'. In: *Research Policy* 15.6 (1986), pp. 285–305. doi: 0.1016/0048-7333(86)90027-2.
- [Tel02] T. Telkamp. 'Traffic Characteristics and Network Planning'. In: *Proc. of the Internet Statistics and Metrics Analysis Workshop*. 2002.
- [Tim98] P. Timmers. 'Business Models for Electronic Markets'. In: *Electronic Markets* 8.2 (1998). doi: 10.1080/10196789800000016.
- [Tou80] G. Toussaint. 'The Relative Neighborhood Graph of a Finite Planar Set'. In: *Pattern Recognition* 12 (1980), pp. 261–268. doi: 10.1016/0031-3203(80)90066-7.
- [TS03] B. Teitelbaum and S. Shalunov. 'What QoS Research Hasn't Understood About Risk'. In: *Proc. of the ACM SIGCOMM Workshop on Revisiting IP QoS: What have we learned, why do we care? (RIPQoS'03)*. ACM, 2003, pp. 148–150. doi: 10.1145/944592.944603.

- [Tuf02] B. Tuffin. 'Revisited Progressive Second Price Auction for Charging Telecommunication Networks'.
In: *Telecommunication Systems* 20.3–4 (2002), pp. 255–263.
doi: 10.1023/A:1016545228543.
- [Tuf04] B. Tuffin. 'Charging the Internet Without Bandwidth Reservation: An Overview and Bibliography of Mathematical Approaches'.
In: *Journal of Information Science and Engineering* 19 (2004), pp. 1–22.
- [Van01] A. Van Moorsel. 'Metrics for the Internet Age: Quality of Experience and Quality of Business'.
In: *Proc. of the Fifth International Workshop on Performability Modeling of Computer and Communication Systems, Arbeitsberichte des Instituts für Informatik, Universität Erlangen-Nürnberg, Germany*. Vol. 34. 13.
Citeseer. 2001, pp. 26–31.
- [Vid03] Video Quality Experts Group.
Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II (FR_TV2). 2003.
- [VM44] J. Von Neumann and O. Morgenstern.
Theory of Games and Economic Behavior. Princeton University Press, 1944.
ISBN: 978-0691130613.
- [Wah12] S. Wahlmueller. 'Pricing and Regulating Quality of Experience'.
MA thesis. Graz University of Technology, Institute of Signal Processing and Speech Communication, 2012.
- [WAR11] S. Wong, E. Altman, and J. Rojas-Mora.
'Internet Access: Where Law, Economy, Culture and Technology Meet'.
In: *Elsevier Computer Networks* 55.2 (2011), pp. 470–479.
doi: 10.1016/j.comnet.2010.09.004.
- [Wei+07] H. Weigand, P. Johannesson, B. Andersson, M. Bergholtz, A. Edirisuriya, and T. Ilayperuma.
'Strategic Analysis Using Value Modeling—The c3-Value Approach'. In: *Proc. of the 40th Annual Hawaii International Conference on System Sciences*. Vol. 40. HICSS '07 9. IEEE, 2007.
- [Whi95] D. White. 'Application of Systems Thinking to Risk Management: A Review of the Literature'. In: *Management Decision* 3.10 (1995), pp. 35–45.
doi: 10.1108/EUM00000000003918.
- [Wil76] R. Willig. 'Consumer's Surplus Without Apology'.
In: *American Economic Review* 66.4 (1976), pp. 589–597.
- [WL03] D. Wilson and L. Lockshin. 'Using the Thiel Index Coefficient to Analyse Variation in Wine Consumption'.
PhD thesis. University of South Australia, 2003.

- [Wu+09] W. Wu, A. Arefin, R. Rivas, K. Nahrstedt, R. Sheppard, and Z. Yang. 'Quality of Experience in Distributed Interactive Multimedia Environments: Toward a Theoretical Framework'. In: *Proc. of the 17th international conference on Multimedia (MM'09)*. ACM, 2009, pp. 481–490. doi: 10.1145/1631272.1631338.
- [Wu03] T. Wu. 'Network Neutrality, Broadband Discrimination'. In: *Journal of Telecommunications and High Technology Law* 2 (2003), p. 141. doi: 10.2139/ssrn.388863.
- [YH08] K. Yamagishi and T. Hayashi. 'Parametric Packet-layer Model for Monitoring Video Quality of IPTV Services'. In: *Proc. of the IEEE International Conference on Communications (ICC'08)*. IEEE, 2008, pp. 110–114. doi: 10.1109/ICC.2008.29.
- [Yit83] S. Yitzhaki. 'On an Extension of the Gini Inequality Index'. In: *International Economic Review* 24.3 (Oct. 1983), pp. 617–13. doi: 10.2307/2648789.
- [Yu95] E. Yu. 'Modelling Strategic Relationships for Process Reengineering'. PhD thesis. University of Toronto, Dept. of Computer Science, 1995.
- [Zin+10] T. Zinner, T. Hoßfeld, T. N. Minash, and M. Fiedler. 'Controlled vs. Uncontrolled Degradations of QoE – The Provisioning-Delivery Hysteresis in Case of Video'. In: *Proc. of the New Dimensions in the Assessment and Support of Quality of Experience (QoE) for Multimedia Applications*. Tampere, June 2010.
- [ZR94] G. Zlotkin and J. S. Rosenschein. 'Coalition, Cryptography, and Stability: Mechanisms for Coalition Formation in Task Oriented Domains'. In: *Proc. of the AAAI*. Vol. 432. 1994, p. 437.
- [Zwi11] P. Zwickl. 'Visual Design and Analysis Support for Answer Set Programming'. MA thesis. Vienna, Austria: Vienna University of Technology, Faculty of Informatics, 2011.