universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Determinants of Residential Internet Access in Austria

Verfasser

Philipp Sandner, Bakk.

angestrebter akademischer Grad

Magister der Sozial- und Wirtschaftswissenschaften
(Mag. rer. soc. oec.)

Wien, im November 2009

Studienkennzahl lt. Studienblatt:	A 140
Studienrichtung lt. Studienblatt:	Diplomstudium Volkswirtschaft
Betreuerin:	Privatdoz. Dipl.-Ing. Dr. Christine Zulehner

Acknowledgment

I would like to thank the economics division of RTR GesmbH, which gave me the opportunity to write this thesis during a several months lasting apprenticeship. Not only the data I was allowed to work with but also the working environment in the division was more than perfect and was a crucial factor for the successful ending of this work. I especially want to thank Anton Schwarz and Elisabeth Dornetshumer, who provided me with suggestions and help throughout the writing process. Finally, many thanks to Christine Zulehner for her helpful comments on this thesis.

The opinions, conclusions and results expressed are entirely those of the author and do not represent those of RTR GesmbH.

Table of Contents

List of Tables

List of Figures

1. Introduction

"The future is already here, it's just unevenly distributed."
-William Gibson (science fiction author)

The Internet is one of the few inventions of the last 100 years that has changed our daily life in a revolutionary way. It has never been so easy to communicate and obtain information as it is now. For many Austrians the so called web 2.0 is an essential part of their life and it is just unimaginable for them to have no access to the World Wide Web. Internet itself is a crucial prerequisite for the development from an industrial to a knowledge based economy. Consequently, EU and Austrian authorities promote Internet diffusion and have set eager aims for themselves and their policies to boost penetration rates.[1]

To guarantee the efficiency of these strategies and to broaden Internet access, it is crucial to know the determinants of Internet use[2] to tackle the so called digital divide which is a synonym for differences in penetration rates between different groups or entities of society. The Internet could be a powerful tool to facilitate the access to information and participation of disadvantaged groups in political and social life. Nevertheless, it is important to realize that the World Wide Web also incorporates the risk of reinforcing social inequalities as those who already have a large amount of resources benefit first and most of new technologies and media. This phenomenon is called the “Matthew effect” coined by the sociologist Robert Merton (1968), according to the Gospel of Matthew: “For to everyone who has, more shall be given.” (Matt. 25:29, New American). To avoid such developments and benefit from the positive effects of the Internet, a detailed analysis of its determinants and the digital divide with its various dimensions is necessary. This thesis should be considered to be a contribution towards an understanding of the gaps in and determinants of Internet use which is essential to promote Internet diffusion and its opportunities for everyone.

From a technical point of view everyone in Austria could have an Internet access but not everyone has as 65.4 percent[3] of all households are online which is above the EU 27 average of 60 percent but lags behind forerunner countries like Norway (84%), Sweden

---

[1] See Austria’s current government program p. 63: “[…] bis 2013 soll die Versorgung der Bevölkerung mit Zugängen von zumindest 25 Mb/s erreicht sein.“ or EU ministerial Riga declaration (2006).
[2] The existence of a residential Internet access in the respondent’s household was collected in all surveys and not the household’s actual Internet use. However, the physical possibility to enter the World Wide Web is a strong indication for home Internet use. Hence I will use “Internet use” and “Internet usage” as synonyms for residential Internet access.
[3] RTR (2009)

(84%) or the Netherlands (86%)[4]. A brief look at the data reveals that there are significant differences in Internet penetration rates among households depending on the filter used e.g. age, income, sex etc. Hence strategies for a further increase of the Internet penetration rate have to address this problem which requires an in-depth analysis of the major gaps in Internet use. This thesis is an attempt to contribute some findings about the sociodemographic determinants of Internet use and to test some alternative hypotheses known from literature that could also explain the differences in Internet penetration rates based on Austrian data from 2006 and 2009.

In the beginning, the penetration rates are computed for several characteristics and categories to find possible differences just through a descriptive analysis. Furthermore the logit model – named basic model – estimates the influence of sociodemographic, spatial and job-related variables on residential Internet access. All used survey samples contain sufficient observations to include additional variables that could explain differences in home Internet use. The last part of this thesis deals with a multinomial logit model – named advanced model – which examines the influence of the already used independent variables on the choice of access type e.g. DSL, CATV, Narrowband and mobile Broadband[5].

The aim is to find determinants for the choice and examine possible developments over time. Earlier papers indicate that various sociodemographic variables are significant. Especially the role of mobile Broadband in this case is very interesting as Austria is one of the leading countries in the distribution of mobile Broadband in the EU[6]. 27 percent of all Internet using households entered the World Wide Web via UMTS/HSDPA according to the survey in 2009 and there is a forecast of steady growth for this year[7]. This fact makes the analysis even more interesting because it is a unique chance to examine the effect of mobile Broadband on Internet penetration rates.

One of the aims of the thesis is to test for alternative approaches to explain the Internet use of households. A special focus is dedicated to the so called spatial gap. In literature and in the political discussions concerning Internet penetration it is often stated that there is a significant difference between rural and urban areas due to an underdeveloped ICT infrastructure and fewer providers of Internet services which lead to disadvantages for rural inhabitants and companies as well as to a lower Internet penetration rate. The idea is to categorize all observations according to their residence into urban, suburban and rural

---

4 http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-QA-08-046/EN/KS-QA-08-046-EN.PDF (19th August, 2009)
5 Broadband access is defined by a bandwidth of at least 512 kBit/s, an always-on characteristic and usually a flat rate.
6 European Commission (2008)
7 http://www.kurier.at/geldundwirtschaft/319664.php (17th June, 2009)

households and add these variables in the model to check for significant influence on residential Internet access. Households interviewed in the survey were asked for their ZIP code and therefore it is possible with the Geographic Information System (GIS) tool to allocate the number of inhabitants of each community to the household's answers. A crucial point of this thesis is to find a proper filter to distinguish between urban, suburban and rural communities. As there are different options to define rural-urban communities, various models are computed with alternative measures and their cut-off points are shifted to check for robustness. For further discussion on this topic see chapter 4.3.

As I create spatial variables based on subjective criteria to measure a possible spatial influence, the representativeness of these variables is questionable. Both RTR surveys are representative for Austria on national and state level which is not a guarantee that my variables based on my rural urban definitions are representative as well. Hence I compute my model with survey data from an annual EU survey – named EU-SILC[8] – which is representative for Austria on community level to check my findings for robustness and representativeness.

Another alternative approach to explain differences in home Internet access could be suburban areas. Smaller towns with a metropolitan neighbourhood often have the same supply with IT services and are part of the urban market just because they are close-by. A variable measuring this effect will be part of the basic model.

It is known from Psychology and earlier papers (see Agarwal et al., 2005) that Internet use is influenced by peer groups. People adopt their habits to those of their peer group. The challenge obviously is to define a proper peer group for each household and get available data to test this hypothesis. ZIP code and state area data could be an option but it is doubtful if peer groups are formed on this level. For further discussion see chapter 3.

The relationship between Internet use and job position is of interest as it could deliver explanatory power to the model. The idea is that job-related Internet use is positively correlated with private use. Once a person has discovered the advantages of the World Wide Web he or she does not want to live without them at home. Respondent's job position was collected for the 2009 survey but not their professional Internet use. Therefore job position is a proxy for job-related Internet usage.

The price of the available Internet services could also have an effect on Internet use. In urban areas with a more competitive market the providers are forced to lower prices to attract consumers while on the countryside the incumbent, Telekom Austria (TA), is sometimes challenged by none or only one alternative provider. Especially the introduction of mobile Broadband forced market prices to drop in recent years as more alternative providers

[8] For more information on the EU Community Statistics on Income and Living Conditions (EU-SILC) see section 5.2.

are available in rural areas. However, price differences between alternative operators, which offer their services mainly in densely populated areas, and TA were substantial in 2006. The different possibilities how to measure the influence of costs on penetration rates are discussed in section 4.5.2.

All these ideas and different approaches are discussed in detail in the remainder of this thesis which is structured as following: The first part provides the reader with a description of the theories and dimensions of the digital divide (chapter 2). A section about peer group influence (chapter 3) completes the theory part. Chapter 4 takes a look at the available data and presents ways how to implement the ideas and theories into a statistical model and the problems generated by attempting this. The applied part of my thesis examines the survey data with the help of a logit model (chapter 5) and a multinomial logit model (chapter 6). Furthermore, the estimation results are presented. In the end, I summarize my findings and discuss their implications (chapter 7).

2. Digital Divide

The following chapter provides the reader with an overview of the scientific literature about different aspects of the so called “digital divide” which is a popular catchphrase used to describe the gap in Internet use between countries or groups of people.

2.1 Definition and Origin

It is uncertain who introduced the term “digital divide” into the scientific and political discussion about the lack of Internet use by disadvantaged groups of a society. Krings & Riehm (2006) give a brief overview of the term’s history and point out that it became prominent in the second half of the 90s when the US National Telecommunications and Information Administration (NTIA) used the term in the title of their second study concerning Internet use in the USA named “Falling through the Net II. New Data on the Digital Divide” (1998). Furthermore there is no official definition of digital divide. The OECD (2001) stated that “the term ‘digital divide’ refers to the gap between individuals, households, businesses and geographic areas at different socio-economic levels with regard both to their opportunities to access information and communication technologies (ICTs) and to their use of the Internet for a variety of activities.”. An alternative explanation of digital divide is that “the term is used to describe a situation where a discrete sector of the population suffers significant and possibly indefinite lags in its adoption of Information and Communication Technology (ICT) through circumstances beyond its immediate control.” (Warren, 2007). It follows from this fairly general definition that the digital divide has several dimension. The following Table provides an overview of different aspects of the digital divide:

Table 1: Different aspects of the digital divide

Unit of Observation	**Citizens** Individuals/Households	**Businesses** and Organisations	**Regional units** e.g. countries
Independent variables (examples)	▪ Age ▪ Gender ▪ Income ▪ Education ▪ Location ▪ Ethnicity	▪ Sector ▪ Number of employees ▪ Turnover ▪ Location	▪ Location ▪ GDP/capita ▪ Size ▪ Population ▪ Language
Indicators (examples)	▪ Access to and/or usage of ICT & Internet ▪ Skills in using ICT ▪ ICT infrastructure		

Source: Selhofer & Hüsing (2002)

As the title of my thesis indicates, the focus is on determinants of residential Internet *access* but the term "digital divide" can also refer to gaps in PC use, a lack of PC skills or awareness issues[9] can also be meant with the term "digital divide". An easy to read overview about all the various forms of the digital divide was written by Van Dijk (2005) who provides the reader with a good introduction to the broad topic.

In my thesis I will follow Fong et al. (2001) who divided the dimensions of the term into:

i. Differences in Internet use between developed and less-developed countries
ii. Differences between children and adults in the adoption of Internet
iii. Differences between types of people within a country
iv. Differences between types of areas

The last three of these differences will be subject of further discussion and hypothesis tests. The theoretical part of this thesis is structured according to Fong's classifications. Foremost, I want to give the reader a general introduction into the topic of digital divide and the literature published so far.

[9] According to the Community Survey on ICT usage in households and by individuals (2008), the main reasons for not having home Internet are a perceived lack of need (38%), costs for equipment (25%) and access (21%) as well as a lack of skills (24%). However, this thesis deals with the reasons of having Internet rather than with the reasons of *not* having Internet.

Fong et al. (2001) found a gap in Internet use within groups of different socioeconomic status and type of areas and used the phrase “double digital divide”. Due to the introduction of Broadband and the gap of adoption within different socioeconomic groups Smolenski (2000) was concerned about a third digital divide. In Austria the percentage of Narrowband households dropped from 32.6 percent in 2006 to 5.9 percent according to the 2009 survey hence these concerns may turn out to be for no reason in the long run. The decrease in Narrowband penetration is a good example of radical changes in ICTs within a short period of time.

Various papers dealing with the digital divide can be found but almost all of them are several years old. At the beginning of the 21st century the scientific community in general seemed to be aware of the problem and a lot of research focussed on the gap between Internet non-users and users.

Selhofer & Hüsing (2002) introduced a measure for the digital divide. The Internet and PC use of certain “risk groups” formed by gender, age, income and education[10] are compared to the average of the society. This approach can be challenged in many ways but it provides the reader with a good intuition of PC and Internet diffusion across countries. According to Selhofer & Hüsing (2004), Austria improved its Digital Divide Index (DIDIX) from 47 (e.g. an average of 47 percent of all risk groups have the same Internet and PC use than the average of society) to 63 which is the second place in the EU 15. Especially in the risk group “age” which consists of people aged 50 years or older the gap closed drastically. As mentioned earlier these results are based on data from 2004. A more recent dissertation dealing with the determinants of Internet use in Austria calculated a DIDIX of 74.1 percent for 2007 and found that the gender gap has decreased while the age, income and education gaps are persistent (Donat, 2008).

In its annual digital competitiveness report the European Commission (2009) reports a digital divide and calculates an index which measures the disparity in regular Internet use between several risk groups and the average of the total population. A value of one implies equality of the risk group usage with the average population usage. The average EU risk index in 2008 is 0.66 and some groups are less frequent users than others. While women, unemployed and rural inhabitants have rates above 0.8, elderly people and low educated persons have substantial lower rates. Austria’s score for the listed risk group does not differ significantly from the EU average. In its conclusions, the Commission reports that education and age are the major determinants of Internet use.

[10] According to the authors, disadvantaged groups of society are people aged 50 years or older, people who finished formal education at an age of 15 or below, people who are part of the lowest quartile of income distribution and respondents who are female. For further details see Selhofer & Hüsing (2002).

The NTIA studies beginning in 1995 and 1998 reported that “the most disadvantaged groups in the USA were rural poor, rural and city ethnic minorities; young households; and female-headed households.” (Warren, 2007). The studies’ titles themselves give a very good intuition of the US development until 2004. In “Falling through the Net: Defining the Digital Divide” (NTIA, 1999) a widened digital divide measured by several sociodemographic characteristics is reported. In 2000 according to NTIA US America is on its way “toward digital inclusion”. The next version of the report about the US Internet use is named “A Nation online” (NTIA, 2002) while the penetration rate was slightly above 50 percent. The 98 pages report itself does not even contain the term “digital divide”. The most recent publication by the NTIA (2004) is “A nation online: Entering the Broadband Age” is focusing on the diffusion of Broadband in rural and urban areas.

The optimistic view of these studies is questioned by Chen & Wellman (2003) who examine the existence of the digital divide in eight countries[11] and maintain that the sociodemographic gap is still there and wide: “Within countries, the uneven diffusion of the Internet appears along familiar lines of social inequality such as socio-economic status, gender, age, geographic location, and ethnicity.” The authors stress out that the digital divide is not only about the physical possibility to have an access but also about the awareness of the World Wide Web’s opportunities and technical knowledge. For further discussion of these aspects the paper is recommended to interested readers.

According to a study by the Future Foundation (2004), 51 percent (24 million people) of the UK population was digitally excluded and had no Internet access at home in 2004. They predicted that one third (23 million people) of the Britons will be still excluded in 2025 which was a bad and too pessimistic estimation as last year 71 percent[12] of the British population already had an Internet access according to Eurostat.

A different point of view about the digital divide in the US was published by the General Accounting Office (2001) where they admitted that their findings point towards a digital divide but “it is often the case that individuals with greater education and income are the first to adopt new technologies, and individuals in rural areas are the last to be reached by the deployment of new telecommunications infrastructure. Since the Internet is still in a relatively early stage of commercial deployment, these socioeconomic and geographic differences in Internet usage are not surprising and may not be long lasting.”

Three years later, Chen & Wellman (2004) did only find little signs of a declining digital divide which was “substantial in almost all countries[11]”. They noted that there was

---

[11] USA, UK, Germany, Italy, Japan, Korea (Rep.), China, Mexico
[12] http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-QA-08-046/EN/KS-QA-08-046-EN.PDF (19th August, 2009)

simultaneously both: an increase of the Internet penetration rate and a widening of the digital divide.

A more detailed but less formal paper about social inequalities in Australia and the impact of Internet diffusion was published by Willis & Tranter (2002): The Internet itself has the potential to alter social inequalities in one way or another. On the one hand, Internet could serve as a tool for information acquisition and distribution to overcome the digital divide in society. On the other hand, the already advantaged parts of society can make better use of the Internet and existing patterns of inequalities are reinforced. The majority of papers back the second argument. Willis & Tranter (2002) suggest that it is important to realise that "ICT facilitates both the reinforcement and cross-cutting of social inequalities."

A paper dealing with the Austrian digital divide was published by Brandtweiner & Donat (2007). Eurobarometer data[13] was used and their results of a logit model suggest an age, income and job position gap while region was not identified as a determinant of Internet use. The paper contains an in-depth analysis of the non-user and user groups. Another sociological approach was chosen by Kriengs & Riehm (2006) when they reviewed the literature and its solutions for the digital divide. The paper is a good introduction into the topic and recommended for readers without mathematical background.

In recent years a steady growth of the Internet penetration and the introduction of cheap broadband products were observed which led to a competitive market and a significant different market situation[14]. Hence in a fast developing field like ICT the published results and hypotheses are not long lasting and have to be subject of evaluation with recent data which is the aim of this thesis.

---

[13] Eurobarometer surveys conduct citizen's opinion about EU policies and several other topics. In May and June 2003 Internet usage besides others were part of Eurobarometer. The full questionnaire is available at: http://www.za.uni-koeln.de/data/en/eurobarometer/questionnaires/ZA3905_bq_en.pdf (9th July, 2009)

[14] For a detailed analysis of the price developments in the Broadband market see: http://www.rtr.at/de/komp/KonsultationBBMarkt2007/Untersuchung_Breitbandmarkt.pdf (1st September, 2009)

2.2. Determinants of Internet Use

All studies trying to find and explain determinants of Internet usage start with a model consisting of several sociodemographic variables (see Chaudhuri, Kenneth & Horrigan (2005); NTIA (1999, 2000, 2002, 2004); Chen & Wellman (2003); OECD (2001); Fong et al. (2001); Willis & Tranter (2002); Lenhart et al. (2003)).

Typically several of the following characteristics are included in the model:

- Income
- Education
- Gender
- Age
- Ethnicity
- Geographical location
- Household size
- Language

Respondent's ethnicity was not included in the surveys hence this characteristic will not be in the basic model as well as the language because only the Austrian situation is examined.

After this brief presentation of the papers yet published dealing with various dimensions of the digital divide, the next step is to go further into detail and discuss three relevant gaps – namely spatial, sociodemographic and grey – and evidence for them in literature.

2.2.1. Spatial Gap

The difference in Internet use between rural and urban population is of special interest in this thesis. In literature and political discussions the need for additional funds to promote rural Internet penetration is often argued with a geographical divide. Two reasons are usually mentioned for the importance to focus on the rural – urban digital divide: First, a spatial gap in Internet use may exacerbate existing inequalities in society (see Drabenstott (2001); Forestier (2002)). Only those with an access to the World Wide Web can benefit from Internet's advantages. Public access as found in libraries or universities is a weak substitute for home access because instant availability of major online activities like writing emails, searching for information or social networks[15] is crucial for the user's will to participate. Second, Internet itself has potential to diminish the rural - urban gap especially the "rural penalty" consisting of less competitive markets, higher transitions costs and greater spatial distance between businesses and consumers (Malecki (2003); Whitacre (2005)). The flip side of that coin is that the disadvantages for non-Internet households increase as more and more services are solely provided online. One of the main questions of this thesis is to answer if there is a significant difference in usage which can be explained by a spatial gap.

In 2006 when the first survey was conducted the incumbent TA had a nationwide Broadband coverage of about 90 percent of all households[16]. In addition several other providers participated in the market as well. Although today a diminishing part of the population enters the World Wide Web via Narrowband, back in 2006 Narrowband had a penetration rate of 32.6 percent. Hence it is reasonable to assume that in addition to 90 percent Broadband coverage plus additional coverage of small local providers and the option to choose Narrowband which affords only a POTS access every household in Austria had the chance to have an Internet access in 2006. Therefore a spatial gap can not be argued with the non-availability of Internet access in rural communities.

Vienna has the highest Internet penetration rate in Austria with 72.3 percent of all households (RTR, 2009) and is obviously considered to be urban. Federal states with a predominant rural structure and a low population density such as Burgenland and Lower Austria have the lowest rates with around 60 percent. Thus there seems to be a spatial gap that cannot be justified with a lack of Internet availability. So why is Internet usage not as developed in rural areas as it is in urban communities? Maybe rural people are less interested in having Internet or are not aware enough of World Wide Web's various advantages. This is unlikely the case because due to spatial distance the importance of quick access to information and online services such as online banking or writing emails increases. In major cities public institutions e.g. universities and libraries or customer friendly companies[17] offer free Internet access while in countryside the own home Internet connection is the only option in most cases. Lower wages and population density in the countryside attract less market players which may lead to non-competitive markets and higher prices. Hence costs could be a determinate for the difference in Internet usage in 2006. At the end of 2007, TA offered a triple play package including fixed line, mobile telephone and Broadband service for a nation wide price of € 19.90. Alternative providers

---

[15] For a detailed description of the major online activities in Austria see RTR (2009): „Der österreichische Breitbandmarkt aus Sicht der Nachfrager im Jahre 2009"
[16] http://www.telekomaustria.com/presse/news/2006/0526-pk-neuestruktur.php (29th May, 2009)
[17] E.g. Starbucks, McDonalds and many others:
see http://freewave.at/hotspots or
http://www.starbucks.at/de-at/_About+Starbucks/_Press+Room/Wien+28+September+2005.htm or
http://www.wirtschaftsblatt.at/home/schwerpunkt/itnews/372334/index.do (29th May, 2009)

were forced to reduce their prices as well to challenge the incumbent. Thus in 2009 variance in prices should be significantly lower as in 2006. One hypothesis test of this thesis is to check for a relation between price and residential Internet access. A difficult challenge would be to create a price basket for every region and collect even the prices from small local providers hence a proxy is used. For further discussions on this topic see chapter 4.5.2.

A lot of scientific literature concerning the spatial gap can be found. It seems that in the early stages of Internet diffusion the geographical divide had a major influence which declined in recent years. For a brief overview of the US development see Vlosky & Poku (2002). In 1999, the NTIA described the situation in the US as follows: "Regardless of income level, Americans living in rural areas are lagging behind in Internet access. Indeed, at the lowest income levels, those in urban areas are more than twice as likely to have Internet access as those earning the same income in rural areas." (NTIA, 1999). Three years later, GAO (2001) specifically examined the situation of Internet usage, no matter if it was Narrow- or Broadband access, in five types of US areas and did not find any significant differences. Six of the eight countries[18] examined by Chen & Wellman (2004) had a declining regional gap. Lenhart et al. (2003) published a study dealing with so called American on- and offliners which reveals a gap in Internet use between urban, suburban and rural areas across the country besides various others sociodemographic determinants mentioned later in the thesis. A study carried out in Western Australia suggests little evidence for a gap in Internet usage when the results are checked for costs, the need to communicate for work and education purposes (Madden et al., 2003). In Canada the difference in Internet access between rural and metropolitan areas was also declining with a higher growth rate for remote Internet access in rural homes than in urban communities (Fong et al., 2001). Similar results were published for the US state of Ohio where differences in Internet activity were caused by urban - rural patterns besides explanatory effects of existing Internet infrastructure and educational institutions (Grubesic, 2002). The Austrian situation was examined by Brandtweiner & Donat (2007) when they constructed a logit model to find determinants of Internet use. The region variables turned out to have no significant influence while age, income and job position had. Their data was collected in 2003. Donat (2008) included an East – West region variable in her model of Internet usage which turned out to be significant. The hypothesis was that inhabitants of Western Austria have worse infrastructure therefore use the Internet less frequent than Western citizens.

A more recent dissertation about the geographical digital divide concludes that it has shifted from spatial dimension to one of high – speed access (Whitacre, 2005). Smolenski (2000) argued similar when suggesting a third digital divided. The European Commission

---

[18] USA, UK, Germany, Italy, Japan, Korea (Rep.), China, Mexico

started a Broadband initiative to promote Broadband deployment in rural areas with additional funds[19]. The European Commission (2008) addressed the spatial gap in Broadband use in their 14th Implementation Report of the Telecommunications Regulatory Package. In Austria the rural DSL coverage rate is 11.4 percent lower than the national one and results for Cable coverage even suggest a 21 percent gap. As the first part of this thesis deals with Internet and not Broadband use, considerations of this possible aspect of spatial difference in residential Internet access rates will follow in chapter 6 when a more advanced model is constructed.

The Digital Divide Index (DIDIX) mentioned earlier includes gender, age, income and education. Due to data non-availability, spatial issues are not consider but named "crucial" (Selhofer & Hüsing, 2004) for a more detailed analysis which this thesis is supposed to be. Other publications questioning the influence of a spatial gap besides yet presented ones are Hindman (2000) and Mills & Whitacre (2003). Hindman (2000) concludes from a US national survey that age, income and education predict Internet usage much stronger than a spatial variable. In a similar study, Mills & Whitacre (2003) explained 63 percent of the spatial gap with household attributes like income and education which indicates that sociodemographic characteristics account for an absolute majority in differences of residential Internet usage. Hence Whitacre (2005) promotes initiatives that address inequities in rural income and education levels which will have the largest impact in reducing the digital divide in the long run. The problem is that progress will not be measurable in a five years term which makes it less attractive for politicians. Thus in the short run, it is more favourable for policy makers to subsidise the improvement of rural infrastructure to raise Internet usage. Scientific opinions about the effects and dimensions of a spatial gap are differing. Therefore it is interesting to take a look at the Austrian situation in 2006 and 2009 especially at the change over time as mobile Broadband penetration rates increased by 600 percent in three years.

A crucial question is how to measure the rural – urban gap. Every household was asked for its ZIP Code in the survey to which – via the GIS tool - the exact number of inhabitants is allocated. The challenge is to find a proper number to filter between urban and non-urban communities. Smaller towns around an urban centre often have the same number of Internet Service Providers (ISP) and products to choose from due to their lower spatial distance thus they are part of the urban competitive market. A possible way to deal with this problem is to assign these towns to the urban fraction or to create a third category named "suburb" as Lenhart et al. (2003) did. A different way to define which community is urban or rural is via population density. Rural communities usually have a lower population density

---

[19] For further information see: http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/09/35&format=HTML&aged=0&language=EN&guiLanguage=en#fn2 (8th July, 2009)

than urban cities. This approach was already chosen in a report in 2008 for the European Commission which examined the development of Broadband access in all member states with a focus on the urban - rural gap. The coverage rate of DSL in urban, suburban and rural areas was computed. These categories were defined via population density[20].

The spatial variable will be included in the basic model to check for its influence and significance. Results have to be subject of robustness tests e.g. a slight change of the cut-off number mustn't cause drastic changes in regression's outcome. For further discussion see chapter 4.3.

The next section will discuss published results concerning sociodemographic influence on the digital divide so far.

2.2.2. Sociodemographic Gaps

The main determinates of Internet use are sociodemographic variables according to earlier papers (see NTIA (1999 – 2004); Fong et al. (2001); GAO (2001); Chen & Wellman (2003); Lenhart et al. (2003)). Almost every study related to this topic reveals an uneven distribution of Internet use within all sociodemograpic characteristics. Non-Internet users are often already part of a disadvantage group of society e.g. less educated, elderly or earn lower wages. The main problem is that social exclusion is followed by digital exclusion which turns to new social exclusion as Warren (2007) points out. Policies addressing the social exclusion and therefore the digital divide are beyond the scope of this thesis. Interested readers see Whitacre (2005) and Warren (2007).

In the following section all relevant sociodemographic attributes, which are included in the model, are discussed.

2.2.2.1 Income

Internet access requires a PC or Laptop, e-literacy and a connection to an ISP. Although a significant drop in ICT supply prices in recent years made the necessary tools to enter the World Wide Web affordable for about everyone and led to a higher computer penetration, it still requires a certain amount of disposable income to pay for the monthly charge and

---

[20] The methodological part of the report is published here: http://ec.europa.eu/information_society/eeurope/i2010/docs/benchmarking/broadband_methodology_06_2007.pdf (8th July, 2009)

possible maintenance of the hardware. Hence the expectation is that Internet use is more likely in high income groups.

Scientific literature identifies income as the most important factor among all sociodemographic variables with a persistent influence on Internet usage over years. In the USA 86.1 percent of the income group of $ 150,000 or higher were online whereas only 31.2 percent of all the persons with an income of $ 15,000 or less used the Internet in 2003 (NTIA, 2004). Similar patterns are found by Chen & Wellman (2003) in all of the eight examined countries[21]. There is a linear relation between income and Internet use consequently Fong et al. (2001) calls low income “a substantial barrier to Internet access”. Even the increase in using rates is uneven distributed among income groups. Disadvantages groups did not catch up but the gap widened (Lenhart et al., 2003). European studies document the same trends in almost all countries at different levels (see Selhofer & Hüsing: 2002, 2004). An OECD report (2001) notes that there is a substantial income gap in Internet use but at least rates of increase for the lowest income groups have been higher in the past than for high income households. An inter-country analysis discovered that the income effect on Internet activity has greater impacts in countries with lower Internet usage rates (IURs) than in countries with higher ones (Beilock & Dimitrova, 2003). As more and more individuals realise the benefits of being online, residential Internet access is not an exception or luxury but ordinary for a broad majority thus the income gaps diminishes.

All these observations are based on data from an early stage of Internet diffusion thus it is the question once again if it is a “natural” pattern of diffusion where educated and high-earning individuals adopt first to new developments whereas disadvantaged groups lag behind or if it is a substantial and persistent gap. Recent survey data from 2006 and 2009 provides the opportunity to take a look at the Austrian situation and development.

As income data is considered to be sensitive, respondents sometimes refuse to give a proper answer. Unfortunately there is a significant number of missing values that has to be dealt with. Education could be a proxy for income due to a high correlation between these two variables. Another option could be an imputation procedure. For further discussion on this problem see chapter 4.1.1

[21] USA, UK, Germany, Italy, Japan, Korea (Rep.), China, Mexico

2.2.2.2 Education

The Internet is a quick and reliable source of information[22] for people all over the world especially for students because most of its content is free and available 24/7. Homework, presentations and projects are done more effectively and in less time with the help of the World Wide Web. Hence a positive relation between a person's education and Internet use can be assumed. Well-educated people are the first to adopt new technologies like the Internet and are more likely to have a white collar job which is well paid and includes the need to use the Internet. As argued before, all these characteristics have a positive influence on the probability of residential Internet access.

US data supports these predications as Lenhart et al. (2003) reports that educational level and student status were the most important independent determinants in their model. Only 15.5 percent of US population with less than a high school degree uses the Internet while 88 percent of the people with a bachelor degree or beyond are online (NTIA, 2004). This positive relation between high educational level and home Internet access is confirmed in an OECD Report (2001) which further notes that within groups of identical income those with a higher education have higher rates of Internet access. The general picture in Europe is the same but the education gap is compared to gender, grey and income gap the widest according to DIDIX (Selhofer & Hüsing, 2004) with approximately 27 e.g. people who finished compulsory school at the age of 16 are only 27 percent as likely to use ICT as the average of society. In Austria the gap was not as wide as the EU DIDIX average indicated. Donat (2008) computed a score of 67.7 percent. A UK study suggests that almost 45 percent of the digital excluded persons have no higher level of education (Future Foundation, 2004). Fong et al. (2001) reports a substantial education gap for Canada and the USA. Australian data reveals an influence of educational level on Internet adoption and use as well (Willis & Tranter, 2002).

2.2.2.3 Household size

Children are curious and open-minded when it comes to the adoption of new ICTs. Internet is essential for them to fulfil their school tasks and keep in touch with their friends via online communities. These days it has to be feared that kids without residential Internet access are in a certain way disadvantaged when it comes to making new friends or communicate with

[22] In a recent study Wikipedia got better reviews than Online Brockhaus in several categories and even Microsoft Encarta is history:
http://www.welt.de/webwelt/article1431467/Wikipedia_siegt_gegen_Online_Brockhaus.html
(1st September, 2009)

other class mates. Hence children can be a major motivation for parents to get a PC and an Internet access. Smolenski (2000) published that 50 percent of all 40,000 respondent in a survey agreed with the statement that a home computer is essential for a child to succeed in school. These considerations suggest that there should be a positive relation between the number of persons living in a household and Internet use. Agarwal et al. (2005) predicted household's Internet access with several sociodemographic variables including four dummy variables for children aged 11 or younger, children aged between 12 and 17, children in both (<11 and 12 – 17) age groups and a general variable if the respondent is parent or guardian of any children under age 18. In all of his several different models only the general parent variable has significant influence. It has to be mentioned that the other determinants like income, education, peer effects[23] or ethnicity are highly significant and have greater impact on Internet use than the presence of children in a household. OECD (2001) backs the positive influence of household size and notes that families with children have the highest Internet access rates among all households. In their study it is approximately twice more likely for couples with children under 18 to have residential Internet access than for single person households. The situation in Canada and the USA was examined by Fong et al. in 2001. US households without Internet access or even a PC have a mean household size of 2.1 while digital included households contain of an average of 2.8 persons. NTIA (2002) reported that families with children under 18 are more likely to enter the World Wide Web (62 %) than families with no children (53 %).

The number of persons living in the same household as the respondent will be included in the model. Especially in an advanced model it is important to find out which influence the sociodemographic variables have on the choice of Internet access type e.g. ADSL, CATV, Mobile Broadband, etc. Household size could have explanatory power as a first look at the data indicates significant differences in the choice varying with the number of people in a household.

2.2.2.4 Gender

Many substantial differences in sociodemographic status e.g. income and education between men and women exist and as these characteristics influence Internet use, the presence of a gender gap in Internet access has to be assumed. In an early stage of Internet diffusion the gap seemed increasing and persistent but declined over time. Several reasons for an early gender gap are mentioned. As technology itself is a product of social relations, diffusion of new technologies favours particular social groups, such as men (Edwards, 1995). Gender

[23] For further discussion of peer effects see chapter 4.

inequalities in professions and industries producing content and technologies of the Internet led to a primarily male orientation of the World Wide Web (Bimber, 2000). He further notes that many media exhibit gender differences in quality and quantity of use for example in total number of watched television news or hours reading a daily newspaper. Hence a gap in Internet use between sexes is not surprising. Shashanni (1997) examined 115 female and 87 male college students and determined that male students were more experienced, confident and interested to use the Internet than their female colleagues.

In the 90s gender effects in Internet use did exist but since Internet usage became less expensive and skill-intensive, the gap declined. The empirical results back these assumptions. Bimber (2000) determined that one-half of the existing gender gap is explained by gender-specific phenomena while sociodemographic characteristics account for the rest. His results are based on survey data from 1996, 1998 and 1999. OECD (2001) expressed concerns about a small gender gap but mentioned equality in Internet usage in the USA. Similar results were published by Ono & Zavodny (2003) who suggest that women were significantly less likely to use the Internet than men but this gap disappeared by 2000. Indeed, NTIA (2002) and GAO (2001) found no differences in Internet usage based on sex for 2000 and 2001 in the USA. No gender influence was revealed by Lenhart et al. (2003) when they examined a sample of approximately 3,500 American adults. One year later, Chen & Wellman (2004) supported these findings but revealed gender gaps in other countries including the UK and Germany. In Australia, differences in Internet use between sexes seemed to decline over time as well (Willis & Tranter, 2002).

Results for Europe indicate a declining gap. Hüsing & Selhofer (2004) published a time series of their DIDIX for the 15 EU member states where sex has the best score by far with 87 in 2002[24] which means that there is almost no gender divide. Austria is ranked third with an above the average score of 93.

In 2006, according to the RTR survey 62 percent of the male respondents but only 44 percent of all women were online. In comparison to 2009 it seems that the gap still exists and is persistent as 71 percent of all men used the Internet while 56 percent of their female counterparts did so. Interesting gender patterns concerning Internet activities were found as well. Hence, gender differences and their development over time will be examined in this thesis. A gender dummy variable is added to the basic model and checked for significance.

[24] In fact, the gender DIDIX increased from 80 in 1997 to 84 in 2000 to an all time high of 87 in 2002.

2.2.3 Grey Gap

A person's age actually is a sociodemographic characteristic like income or education. Since literature often mentions the difference in Internet use between age groups as grey gap and Fong et al. (2001) listed it as a separate dimensions of the digital divide, I will deal with the age gap on its own. Age is considered to be a special sociodemographic dimension.

Older people adapt slower to new developments in ICTs. As argued in section 2.2.2.3 children playfully approach new ICTs and are less worried about making mistakes while learning how to deal with new technology. Unfortunately for elderly people the opposite is true. Although Internet access became less expensive and skill-intensive it still requires a PC or a laptop which costs at least several hundred Euros and basic IT knowledge is necessary. For elderly people these requirements often turn out to be substantial barriers to Internet use. The World Wide Web's benefits for old age pensioners maybe are not as big as for people aged 50 or younger. Online Gaming, Online Communities or Online Shopping is predominantly favoured by younger users (RTR, 2009). Thus it is very remarkable that older people recognize the importance of computers for their children and grandchildren to succeed according to Smolenski (2000) although they lag behind in Internet use.

The recent RTR report (2009) revealed that only 40 percent of the people aged between 60 and 69 use the Internet while the Austrian average is 65.5 percent. In the age group 70 + the Internet penetration is even lower with 15 percent. There are substantial differences in pattern of Internet use as well. People aged 60 or older are significant less interested in Online Communities, Online Banking and Online Games than younger users.

The existence of the grey gap is called persistent over time in several papers. OECD (2001) mentioned that in all member states age patterns played a role in determination of Internet usage and were similar across countries. The most active Internet using group appeared to be 35 to 45 aged users. NTIA (2002) reported that children and teenagers use the Internet more than any other age group in the USA. In Canada the Internet use rate is increasing with age and has its peek at the 25 to 44 years group (57 %). It drops to 17 percent for the 65 years or older group (Fong et al., 2001). A substantial grey gap in the USA was documented by the UCLA Center for Communication Policy (2003): 97 percent of the children aged 18 or young accessed the Internet while only 34 percent of the 65 plus generation did. Lenhart et al. (2003) suggest that about a quarter of all non-Internet users are 65 or older while just 14 percent are 18 to 29. 56 percent of the non-user group maintain that they will not go online in the future whereas people aged 50 or older represent 71 percent of this group. An Australian paper noted that there is not only a grey gap in Internet use but also in early adoption of the Internet (Willis, 2002). The majority of people aged 60 or older are retired and dropped out of the labour force which leads to less income and a smaller budget for consumption goods. A smaller disposable income reduces the probability to go online as argued in section 2.2.2.1. Hence old age pensioners are double disadvantaged by income and age. A British study revealed that the age group 65 plus represents a relative majority of approximately 45 percent in the category “Digitally excluded below the poverty line (= 60 % of median income)”. Whereas the group of digitally included people below the poverty line contains higher proportions of people aged 16 to 54 compared to the digitally excluded population (Future Foundation, 2004). Once again a look at DIDIX provides a good impression of the European Situation. Selhofer & Hüsing (2002) suggest a widening of the grey gap from 1997 to 2000 in the EU 15. Austria has one of the lowest scores with 22 in cross country comparison while the European average is 41 in 2000. The difference in Internet usage between age groups is declining a little bit by 2002 when the EU age score improves to 53 which was identical with the general EU DIDIX[25].

As one of the aims of this thesis is to examine whether there is a grey gap in Austria and take a look at its development, it is is necessary to find a definition of who is old and who is not. The age of every respondent[26] was collected in the 2006 and the 2009 survey. Several options can be considered: The addition of just the simple age variable consisting of the respondent’s age would capture a linear effect on Internet use. The average Austrian retirement age could be taken as a cut off point for a dummy variable which would measure the effect old and not old but not a linear effect. Selhofer & Hüsing (2002) define the disadvantaged group of elderly people with 50 years or older and for Future Foundation (2004) everyone above the age of 55 is considered to be old. Another option is the creation of age groups and include them via dummy variables e.g. 16 – 29 is a dummy variable, 30 – 39 another one, etc. A special focus concerning the developments of the grey gap should be on mobile Broadband. People aged 60 to 69 have an interestingly high mobile Broadband penetration rate of 25 percent exceeding the age groups of 40 to 49 and 50 to 59. In case the grey gap is decreasing over time and considering the high growth rates of mobile Broadband penetration in recent years this would be a strong indication for a positive role of mobile Broadband in promoting Internet for elderly people.

Further discussion on this challenge and arguments for and against each option of an age variable will be provided in chapter 4.2.

---

[25] General DIDIX consists of the average of all scores in four risk groups (Gender, Age, Education, Income). For further details see Selhofer & Hüsing (2002, 2004).

[26] Respondents needed to be at least 16 years of age.

2.2.4. Inter-country Gap

This thesis is about determinants of Internet use in Austria. Hence differences in Internet use between developed and less-developed countries are not subject of interest for this topic. For a brief overview of the problem see OECD (2001) or Chen & Wellman (2004). An exploratory model of inter-country Internet diffusion is provided by Beilock & Dimitrova (2003). A more general view on the topic is chosen by Wunnava & Leiter (2009) when they suggest seven hypotheses about inter-country Internet diffusion and a model with determinants.

3. Peer Group Theory

One of the main learning methods is imitation. Humans tend to adopt their behaviour to those of their adjacencies and imitate certain patterns of daily life. Hence it is reasonable to assume that there is a positive relation between a person's Internet use and its surrounding community's Internet use. If my family, my friends or colleagues from work are Internet users, there is an immediate incentive for me to go online. The idea is that individual choice is influence by other community members. In a region with educated, young and well paid people who themselves are more likely to be online the social influence will pressure other people in this region to go online as well. This influence can be based on (1) learning from others e.g. tips on use from friends or (2) network externality e.g. greater benefits from being online when others are also online or (3) just pressure to conform.

Different aspects of this peer effect are mentioned in literature. A solid overview of research so far and an attempt to measure the effect was published by Agarwal et al. (2005). These findings suggest a strong group effect on Internet use besides the usual sociodemographic influences. His approach to test for group effects will be subject of further discussion later. Wellman et al. (2001) backed the peer group theory as he reported that Internet use in particular e-mail is a supplement to face-to face and telephone contacts without decreasing or increasing it. People living physically close are contacted up to three times more often than friends who live at a distance. A study dealing with PC diffusion found that peer effects played a major role in the household's decision to buy a computer (Goolsbee & Klenow, 2002) and indicated a potential explanatory gain from adding a group variable. Availability of local Internet content such as localized information was found to be an important motive for using the Internet (Kraut et al., 1996). Another possible explanation is local recruitment where it has been observed that recruiters use the Internet to hire people who live nearby instead of broaden the employee search in terms of geography (Niles & Hanson, 2003). These local level network externalities can be simultaneously both driven by peer effects and may stimulate them.

While there are clear indications for the existence of this peer group effect, the challenge is how to measure it in the data. Agarwal et al. (2005) used US Data from 2003 and formed groups according to the FIPS code[27] which represents the county the respondent resides. An individual's peer group is defined as everyone else living in the same FIPS area. The explanatory variable for each individual is computed as the total number of Internet

[27] The Federal Information Processing Standard (FIPS) for counties uniquely identifies counties in the USA and certain US possessions. For a complete list see http://www.census.gov/geo/www/fips/fips65/data/national.txt (1st September, 2009)

users (excluding the respondent) in a FIPS area divided by total respondents (excluding the respondent). Due to data constraints it was not possible to create smaller groups. Thus a higher value of the FIPS variable for an individual i means that it is more likely for this individual to come in touch with people who are online themselves. Individual's sociodemographic variables are included in the model as well.

Certain problems have to be considered in their model. First of all, there is the possibility of self-selection where people who are online choose to live in an area where others are also online. Similarly, infrastructure constraints like higher costs or unavailability of Internet services e.g. in rural areas do apply to the whole community and not solely to a single person. In both cases there would be a correlation between the individual's choice and group choice without the existence of peer effects. Agarwal et al. (2005) deal with these problems via instrumental variables.

Three models are calculated: Agarwal et al. (2005) started with a linear probability regression model where the dependent variable is the probability of user i to go online. Sociodemographic characteristics and the FIPS variable are included. The results show that peer effects have significant influence after accounting for sociodemographic differences. Two potential problems arise with this model: (1) if individuals' decision to enter the World Wide Web is dependent on others from the same FIPS area than the FIPS variable cannot vary exogenously. (2) Unobserved variables could influence the results as well. Unobserved FIPS level characteristics such as price or Internet availability are correlated with the region variable e.g. urban or rural residence. It is argued that it is difficult to think of an advantage or barrier to Internet use which is uncorrelated with this variable. If so, these unobserved effects would bias the coefficient of the region variable and not the FIPS coefficient. Unobserved individual characteristics that are correlated with FIPS but not with observed individual characteristics are the motivation to use an instrumental variable regression model. FIPS level demographic variables such as median income and age are drawn from the 2000 US census data and used as instruments for the regression. The key finding is that FIPS continues to be positive and significant. A logistic regression model confirms these results.

Can this approach be used to check for peer effects in this thesis with Austrian data? Every respondent was asked for the household's ZIP code ("Postleitzahl"). In case the presented approach from Agarwal is chosen, there is the need to calculate the average Internet usage in each ZIP area to include it in the regression. The conducted surveys are representative for the Austrian population but do not include enough observations for each ZIP area to compute proper averages. In fact a lot of ZIP areas have no observations as there are approximately 2150 ZIP codes[28] in Austria but only 4000 surveyed households in

---

[28] For a complete list of all ZIP codes see http://www.post.at/783.php (1st September, 2009)

2006 and 3000 respondents in 2009. Unfortunately the average Internet usage for each ZIP area is not available from other sources thus alternatives have to be considered. One option would be to define a federal state e.g. Vienna or Lower Austria as a peer group. It is difficult to argue why an inhabitant of Southern Lower Austria should feel pressure to conform to the behaviour of an inhabitant of a village near the Czech border while villages just a few kilometres outside of Vienna are not part of the Viennese peer group. Another idea for the creation of smaller and more reasonable groups would be to group the individuals in geographical terms of quarters. Some states have traditional quarters like “Weinviertel” in Lower Austria while others have not. Therefore this approach needs to be executed carefully but provides the necessary flexibility to deal with the problem of smaller villages next to big cities. In such a case the village is included in the city’s peer group.

Donat (2008) interviewed 529 persons about their media using habits including questions concerning Internet usage of the respondent’s relatives, friends or colleagues. In her model of Internet usage in Austria she found significant influences of the social environment which is a strong indication for peer effects. Unfortunately no such information was collected in the NASE surveys. Further discussion of this problem is presented in chapter 4.4.

Another interesting approach is chosen by Whitacre (2005) trying to measure the influence of social networks. The regional access rate is included in his model as Agarwal et al. (2005) did. In addition, a critical mass term is created. Mahler & Rogers (1999) presented their idea of a different pattern of diffusion in the concept of “critical mass”, defined as the minimal number of adopters for an innovation, after which the further rate of adoption is self-sustaining. The initial rate of adoption is lower than in normal patterns of diffusion but once a critical point in diffusion is reached; the rate of adoption is much higher. For some innovations this concept is more applicable than for others. Fax machines or video conferencing follow the critical mass approach as the interactive nature create network externalities which affect the utility of potential adopters *and* previous adopters. Cellular phones connect to an existing base of phone users hence do not follow this approach. According to Mahler & Rogers (1999), network externalities play an even bigger role in interactive innovations like the Internet as for “normal” innovations. Whitacre’s general logit model includes sociodemographic variables, a spatial variable, a regional density variable and a non-linear critical mass variable which is calculated as the squared regional density variable. It has to be noted, that the critical mass term is dropped due to insignificance and its reduction in explanatory power of the regional density term. As the critical mass term depends on the existence of a regional density term its yet questionable if it will be part of the model. Further information will be given in chapter 4.4.

4. Data

This chapter is all about data and descriptive statistics to get a feeling for the samples my thesis is built upon. Yet mentioned challenges concerning the creation of variables will be discussed here. First I will give an overview of the data and later I will discuss the relevant variables to measure the mentioned gaps and the peer group effect.

My analysis is based upon two surveys conducted on behalf of the Austrian National Regulatory Authority for Broadcasting and Telecommunications (RTR). The 2006 survey was carried out by "Market Institut" in the months of November and December when 4020 households were selected via random quota according to household size, region, number of inhabitants in residential town and age of the household's head. In 2009 "Institut für empirische Sozialforschung" (IFES) was commissioned to conduct the survey for the RTR via random digit dialling and weighted according to the household statistics of the Austrian micro census. The target subject was the person[29] in charge of the Internet access related decisions. Both samples are representative for Austria and gathered with computer assisted telephone interviews (CATI). The answering household members needed to be at least 16 years old. My thesis will only deal with non-business respondent data hence all conclusions and results are concerning this group.

Once in two years consumer surveys ("Nachfrageseitige Erhebung") are needed to get a market overview and an insight in consumer preferences to start a new market analysis process. RTR published the results in reports[30] on their website where interested readers can get further information about the market analysis process. In both surveys respondents were asked questions concerning their Internet use, habits, provider and related costs, etc. In the end respondent's sociodemographic characteristics were collected.

---

[29] In the remainder I will write about the "individual" or "household" attributes, decisions etc. using these two terms as synonyms for the respondent and his or her household.

[30] 2009: http://www.rtr.at/de/komp/BerichtNASE2009 (26th August, 2009)
2006: http://www.rtr.at/de/komp/KonsultationBBMarkt2007/Untersuchung_Breitbandmarkt.pdf (26th August, 2009)

A first glance at the data from both surveys is provided in Table 2:

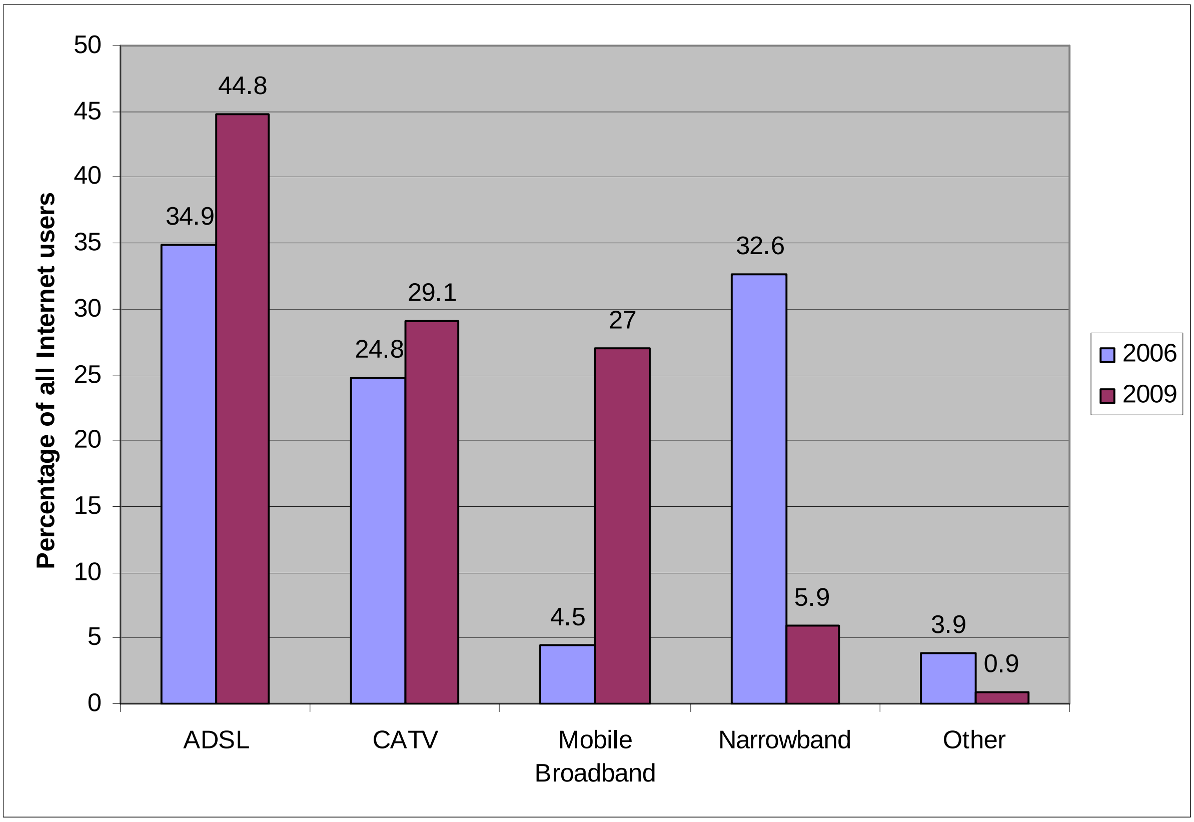
Table 2: Descriptive overview of survey data

	2006		2009	
Variable	*N*	*Mean*	*N*	*Mean*
Age	4020	52.2	3001	46.1
Income	1834	€ 2081.6	2329	€ 2105.5
Household size	4020	2.53	3001	3.1
Internet use	4020	52 %	3001	65.4 %
Sex (Ref. group: men)	1610	40.7 %	1429	48.2 %

Source: RTR survey data

In over three years the Internet usage rate increased to 65.4 percent which is about the same as the European average. The average age, income and household size is a bit larger in the 2009 sample than it is in 2006. Interesting to note and source of a challenge is the high share of NAs in the income category. Almost 58 percent of all respondents refused to declare their income in 2006 while in the second survey 22 percent did so. Obviously people are not willing to reveal their income due to various reasons e.g. the fear of envy is often mentioned in the discussion. The number of observations for this sociodemographic characteristic is too small to get significant results and requires further action. One possible option would be to exclude income from the model and add a proxy e.g. education. Its high correlation with income is obvious because well paid jobs require a better education than blue collar jobs thus education would be a good proxy. But on the other hand education itself would not be a variable in the model and the influence of both major sociodemographic characteristics could not be tested separately although literature suggests strong and significant influence of both on Internet access rates. An imputation procedure for income is the alternative option. As income is assumed to be a function of gender, age and education it should be possible to get appropriate results from an OLS regression to justify an imputation which estimates income data for the missing values. Hence income could be part of the model. For further discussion see the following sociodemographic variables section.

Figure 1: Internet access type



Source: RTR survey data

The shifts in Internet access types are typical for the fast and drastic changes in ICT usage in recent years. The increase of mobile Broadband from 4.5 to 27 percent and the drop in Narrowband usage deserve special attention. Austria is the leading EU member state in mobile Broadband distribution[31] thus results concerning the effects of mobile Broadband on various gaps are of interest for other countries and policy makers. In particular the high mobile Broadband penetration rate in the age group 60 to 69 could be a possible explanation for a diminishing grey gap. The Narrowband penetration rate dropped from 32.6 percent to 5.9 in 2009. Hence Broadband has become the new standard access type as prices have dropped and Broadband is available almost everywhere in Austria.

[31] Austria with a rate of 11.4 percent of the total population actively using mobile Broadband is ahead of Finland (9.1 %) and Portugal (8.3%) while the EU average is 2.8 percent. For further details see European Commission (2008).

4.1 Sociodemographic variables

Several sociodemographic characteristics will be part of the model via different variables. As discussed in the section 2.2.2 results in the literature indicate that income, education, gender and household size are the main sociodemographic determinants of Internet use. The following part of the thesis presents in which form they are included in the model.

4.1.1 Income

It is not possible to add the income variable to the model with the available NASE data due to data incompleteness. There are two possible approaches towards the problem: A proxy instead of the original variable or an imputation procedure. In consideration of the fact that income is a major sociodemographic characteristic and part of every model dealing with this topic, imputation is the method of preference in the beginning and in case of unsatisfying OLS results, the proxy option is chosen.

First, I compute an OLS regression with income on the left side and the available sociodemographic characteristics which may influence the household's income on the right. In 2006 age, education, gender and household size are the dependent variables while in 2009 the degree of employment and the job position of the respondent are included additionally. In both models the age variable is squared as well to measure the reinforced age effect on income. The logarithm of income is computed to transform a possible exponential relationship to a linear one. For the 2006 data this approach improved the results whereas in 2009 the opposite was the case thus the log age is only used for 2006. Hence the interpretation of the coefficients has to be done with caution depending on the year.

>Insert Table 21 about here<

The parameter of interest is the R^2 respectively the adjusted R^2 which includes a penalty for expanding the number of explanatory variables without a good reason. In 2006 the model does a by far better job in explaining income with an adjusted R^2 of 0.46 than in 2009. All variables are highly significant and so is the model itself. The signs of the coefficients are as expected with a positive influence on income. An increase in age, household size and education leads to a higher income. Only the squared age has a negative sign but no influence because its coefficient is equal to zero.

The 2009 model is extended by additional variables to improve the initially low R^2 with moderate success as it is about 0.29. The coefficients of the major variables are positive like in 2006. Being a civil servant or an employee has no significant influence on income but all the other jobs position have. With leading employee as the only exception, the job coefficients are negative which is reasonable because the reference group is supposed to be a higher income category – namely self-employed – compared to being retired or unemployed. Age squared has a negative influence again but is insignificant anyway.

Both models delivered fairly good results to justify an imputation with these independent variables. *Stata* provides the *impute* command which automatically generates an imputed variable of income which is added to the basic model.

4.1.2 Education

Fortunately, education is not needed as a proxy for income and hence it can account for educational effects on residential Internet access rates in the model. The highest completed school degree determinates the education group the respondent is categorized to. Four categories of education are included in the 2006 data set: Compulsory school, some years of college without graduation, college graduation ("Matura") and university degree. In 2009, one more group was created which includes all respondents with a completed apprenticeship. People from the apprenticeship group were part of the compulsory group in the earlier survey thus it can be assumed that effects from the compulsory group are split up between these two groups. Table 3 contains the Internet penetration rate for each education group.

Table 3: Residential Internet access rates in percent by education groups

School type	2006	2009
Compulsory School	35.34	52.63
Apprenticeship	NA	72.38
Some Years of College	60	76.17
College Graduation	74.56	69.7
University degree	85.9	69.5

Source: RTR survey data

As expected there is a significant difference of almost 50 percent between people with compulsory school qualifications and people with a university degree in 2006. A positive relation between Internet usage rate and education is evident. Due to an increase in the Internet penetration rate over three years, access broadens and the education gap is diminishing. In 2009 the only persistent gap is between the lowest education group and all other categories. People with a college graduation or a university degree actually had a lower Internet usage rate as respondents with some years of college without a graduation. But the difference is small (7 %) and the rates are above the national average rate (65.4 %). So the assumption of an education gap is reinforced at least between the lowest group and all other groups. For each education group a dummy variable is formed and added to the model with compulsory school as the reference group. Education will most likely have significant explanatory influence on the dependent variable.

4.1.3 Household size

In each survey the respondent's household size was collected. This variable containing the number of people living in the same household as the respondent measures the linear effect of household size on Internet use. A positive influence is assumed as argued earlier. Descriptive statistics are listed in Table 4. Households without Internet have a significant lower size with 2 respectively 2.22 persons in 2009 than online households with about 3 or 3.5 members which is higher than sample average hence there is a persistent gap over time.

Table 4: Average household size by Internet use

Variable	Internet Use	2006	Sample Average	2009	Sample Average
Household Size	0	2	2.53	2.22	3.12
	1	3.06		3.54	

Source: RTR survey data

4.1.4 Gender

A dichotomous variable with female respondents as the reference group measures the influence of being male on residential Internet access rates and is part of every model dealing with Internet usage. Men tend to have higher Internet usage rates and gender specific Internet habits have been observed in the past. For further information see RTR (2009). Descriptive analysis reveals that men have a significantly higher Internet penetration rate in both surveys. But the observed gender gap decreased over three years. The gender variable should have high explanatory power in the model.

Table 5: Residential Internet access rates in percent by gender

Gender	2006	2009
Men	62.3	76
Women	43.7	63.2

Source: RTR survey data

So far, I presented the differences in Internet use for each sociodemographic variable and found substantial gaps for each of them in both surveys. The following Table compares the percentages for these characteristics between on- and offliners and over the years.

Table 6: Gender, education and income averages by Internet use

Variable	Internet Use	2006	Sample share	2009	Sample share
Gender (Ref. group: men)	0	31.46	41	37.28	48
	1	49.81		53.41	
Income	0	€ 1,506	€ 1,997	€ 1,538	€ 2,136
	1	€ 2,463		€ 2,421	
Compulsory School	0	68.8	50	25.26	15
	1	33.14		10.25	
Apprenticeship	0	NA	NA	41.5	39
	1	NA		37.7	
Some years of College without graduation	0	16.24	20	15.3	16
	1	23.16		16	
College graduation ("Matura")	0	9.68	19	8.33	17
	1	27.39		21.65	
University degree	0	3.07	11	5.03	13
	1	18.24		16.56	

Source: RTR survey data

The percentages in the Table are the shares of Internet users for each characteristic e.g. 33.14 percent of all Internet users have a compulsory school certificate as their best completed educational qualification. The comparison with the overall share of this category in the sample gives an impression of how good Internet use separates the groups. For income, the sample average is computed. The results for household size were already presented in section 4.1.3.

Due to the apprenticeship group in 2009, the compulsory group is smaller and effects are split up between these two groups whereas in 2006 all effects are included in one group. Compulsory school graduates are overrepresented in the Internet non-user group which suggests a positive relation between Internet use and education as argued earlier. The results in 2009 are not as straightforward as in 2006 because the majority of compulsory school members are allocated to the apprenticeship group where the Internet penetration rate is about the same as the sample share. Moreover, no significant difference between on- and offliners can be found in this group whereas t-tests for all other groups turned out to be significant. A merger of both groups into one would deliver results similar to those of 2006. All other education groups are underrepresented in the non-user group respectively overrepresented in the online group. The only insignificant difference is found in the “Some years of College without graduation” group in 2009. A t-test confirms the significant differences between the Internet users and non-users for all characteristics except for the “Apprenticeship” and “Some years of college without graduation” group in 2009.

Gender differences in residential Internet access rates are substantial. Men are underrepresented in the offliner group in both years as their rate is clearly below the average share of men in the sample. Hence this finding indicates a gender gap in Internet use.

Higher household income appears to increase Internet usage as the average income of onliners is significant higher than those of offliners. An income gap seems to be present in both years.

All major assumptions stated in the chapter 1 and 2 of the thesis concerning sociodemographic gaps are backed by the descriptive analysis of the data. Differences in Internet use in education, household size, gender and income are substantial and persistent over time. Thus it is essential to include these variables in the model.

4.2 Age Variable

2006 survey data from “Market Institut” only contained age groups and not the exact age for every respondent while “IFES” in 2009 delivered it. As I want to compute a descriptive statistic the mean of each age group is used as the age for every group member. It is important to note that in both surveys the person who is partly or fully involved in the decision process concerning Internet access responded thus only the household was selected randomly via a certain method but not the person. Consequently the age data could be biased as the focus of these surveys was the household’s preferences and not the individual’s ones. Despite the fact that results from the following descriptive analysis can only be examined with caution, they provide a good insight into and feeling for the problem of the grey gap.

First, the average age of Internet users and non-users is computed.

Table 7: Average age by home Internet access

Internet use	2006	Sample average	2009	Sample average
0	59.2	52	55.8	46
1	45.6		41.4	

Source: RTR survey data

Onliners are significantly younger than offliners. Internet use as a filter for age reveals the gap especially when compared to the average age in each sample. This first impression of the data supports the assumptions made earlier that elderly people are less likely to use the Internet than younger people. For a more detailed analysis of the data I will form age groups for both samples starting with a dummy variable for 16 to 24 years old respondents. Due to data constraints the second group includes people aged between 25 and 29. Furthermore, each dummy variable contains 10 years of age e.g. the third group is 30 to 39, then 40 to 49 etc. The last group is 70 years and older. The following Table provides an overview as seen before:

Table 8: Residential Internet access rates in percent by age groups

Age group	2006	2009
16 – 24	73.9	85.7
25 – 29	69.9	84.2
30 – 39	77	81.6
40 – 49	73.4	76.4
50 – 59	58.5	67.3
60 – 69	36.6	39.8
70 +	7.9	21

Source: RTR survey data

The message from these numbers is obvious: The older people are the less they use the Internet. The development in both samples is similar. The only differences are higher rates for all groups in the 2009 sample due to a higher overall penetration rate. People aged between 16 and 40 have constant and high Internet usage rates over years. Rates drop for age groups above 40 with a drastic decrease in the end of the scale for the group 70 plus. All rates for groups under 50 are clearly above the sample average Internet use. Once again the age gap assumption is backed.

In a final step I calculated the share of each age group for the Internet user and non-user samples as done earlier. Even more interesting than the differences between the on- and offliner shares is the comparison with the sample share of each category.

Table 9: Share of age groups in percent for on- and offliners

Age group	Internet use	2006	Sample share	2009	Sample share
16 – 24	0	2.5	4.1	5.1	13.7
	1	6		17.6	
25 – 29	0	3.1	4.6	4.2	7.4
	1	6.4		9	
30 – 39	0	7	14.2	8.8	17.3
	1	21.5		21.3	
40 – 49	0	14.6	26.1	11.7	19.9
	1	37.7		24.1	
50 – 59	0	12.9	14.8	12.7	15
	1	17.2		16.1	
60 – 69	0	20.4	15.5	27.9	15.5
	1	11.3		9.4	
70 +	0	39.5	20.7	29.6	11.2
	1	3.5		2.5	

Source: RTR survey data

Results from Table 9 are as expected. All differences within an age group between Internet users and non-users are significant. While younger people are overrepresented among onliners in comparison to their sample share, the opposite is true for elderly people. In 2006 people aged 50 and older stand for 51 percent of the sample but represent 72.8 percent of the Internet non-user which is a strong indication that points towards a grey gap. Similar results are found in the 2009 sample. The 50 plus group share in the survey sample is 41.2 percent and stands for 70.2 percent of all offliners. The gap actually decreased a bit over the last three years. All available age data indicates a wide and persistent gap hence this variable has to be part of the model.

4.3 Spatial Variable

A focus of this thesis is on the spatial dimension of Internet use and a suggested regional gap. As there are various possible dimensions of this spatial influence several different variables will be included in the model to test for all of them.

The main task is to find a cut off point to create a dummy to measure the possible differences in Internet use between urban and rural communities. As there is no scientific definition of what has to be considered to be a city and what not, it is the author's choice to define a number of inhabitants as a filter. A single cut-off point for a unique and reasonable definition is hard to find and difficult to argue. In case a number of 50,000 inhabitants is chosen to be the criterion for the dummy variable, why not 40,000 or 60,000? In the end, it is a subjective decision which has to be subject of robustness tests. Slight changes of the urban-rural criterion could lead to drastic changes in the outcome of the regression. If this is the case, one has to take a closer at the data and try to find an explanation. Due to serious doubts towards a single cut-off point, a second additional one is defined to deepen the analysis and guarantee more detailed results. The first cut-off point will be somewhere between 5,000 and 15,000 inhabitants. All cities with a population below the defined figure will be included into one dummy. The second dummy will represent all communities with a population between this point and 50,000 and a third will include all cities larger than 50,000. Further discussion on the cut-off point is following.

A brief look at the population of Austria's cities discovers a huge gap between Vienna with its 1.7 million inhabitants and Graz (250,000), the second largest city. Hence it is interesting to check for additional advantages of Vienna besides being a city with a population larger than 50,000. Reasons for a "Vienna mark up" in Internet penetration could be caused by the most competitive market in the country and the unique population density. Vienna will be the reference group for all federal states hence their effects are estimated relative to Vienna and significant findings indicate substantial differences between Vienna and other states.

Another idea is to test whether smaller towns surrounding bigger cities are part of the urban market and enjoy the advantages of the city like cheaper services and a larger range of available products while actually being considered to be rural. The ZIP codes of all communities surrounding the cities included in the 50,000 plus dummy can be obtained via GIS and I will form a variable to check for the influence of an urban neighbourhood. A similar approach was already chosen for an in-depth analysis of the Broadband market in Austria a few years ago.

Instead of the rural-urban dummy variables or just as an additional explanatory factor the number of inhabitants for each community as a continuous variable can be added.

Depending on the results, maybe the squared number of inhabitants could be included in the model as well. Another possibility to check for regional effects is to create dummies for the nine Austrian federal states ("Bundesländer") with Vienna as the reference group as argued earlier which could reveal structural state differences in residential Internet access rates.

The following section contains discussions on the cut-off point and descriptive statistics based on these urban-rural definitions.

As mentioned earlier, the definition of the cut-off point for the city size dummy variables is up to a certain extent a subjective choice and crucial for this thesis. The first variable includes small towns hence its cut-off point should be somewhere between 5,000 and 15,000 people. The second boundary is around 50,000 inhabitants. I took a closer look at the data and computed the number of observations for several possible combinations of cut-off points. Results are presented in Table 10.

Only nine cities[32] in Austria have more than 50,000 inhabitants and 73 cities more than 10,000. The average community size in Austria is about 4,300. Thus the absolute majority of all observations from both surveys are clearly below the 50,000 inhabitants cut-off point.

Table 10: Number of observations for different "number of inhabitants" cut-off points

Size		5,000 - 40,000	5,000 - 50,000	10,000 – 50,000	15,000 - 50,000
	<	320	320	845	1,776
2006	=	2,368	2,413	1,568	957
	>	1,209	1,164	1,164	1,164
	<	207	207	776	1205
2009	=	1,697	1,750	1,181	752
	>	981	928	928	928

Source: RTR survey data

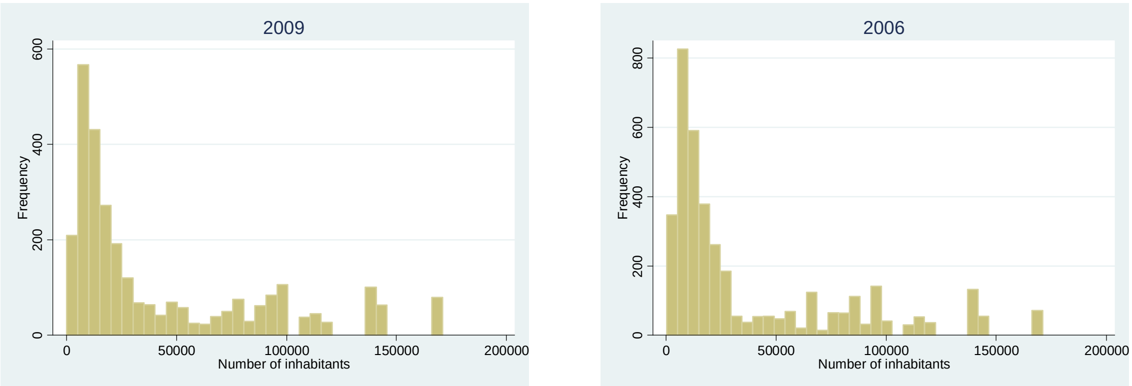
There is almost no difference in the number of observations if the second boundary is changed to 40,000 thus no further discussion about this boundary is needed and all cities with a population larger than 50,000[33] are included into the large city variable. In 2009 about one third and in 2006 a quarter of the respondents reside in one of the largest cities in

[32] Vienna, Graz, Linz, Salzburg, Innsbruck, Klagenfurt, Villach, Wels and Sankt Pölten. For a full list of the largest cities in Austria see http://en.wikipedia.org/wiki/List_of_cities_and_towns_in_Austria (26th August, 2009)

[33] Please note that all Viennese districts are assigned to the 50,000 plus variable although five districts have a population below this cut-off point.

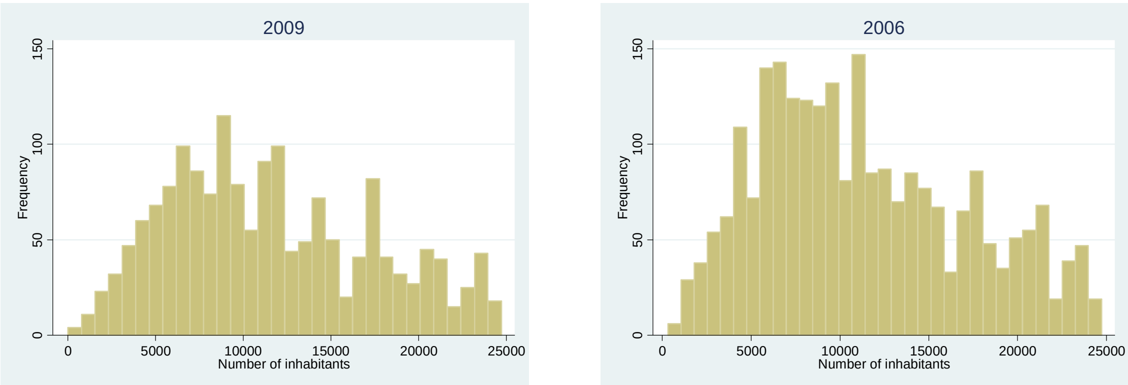
Austria. The determination of what has to be considered to be a small town is not as simple as in the first case. Histograms with the overall distribution of all observations and the number of observations of respondents living in communities with a population of 25,000 people or less are provided in Figure 2 and Figure 3. The majority of all observations are distributed around 10,000. A cut-off point of 5,000 people would create a dummy variable which only contains a small fraction of all respondents (2006: 8.2 %; 2009: 7.2%). Moreover, the middle size town variable would account for about 60 percent in both surveys which cast serious doubts on reasonable interpretations of the regression results. A boundary of 15,000 people includes too many observations into the small town variable. Furthermore a community with 15,000 inhabitants cannot be considered to be a small village. Approximately 500 respondents are transferred from the middle size town category to the small town category in both samples when the boundary is set to 10,000 instead of 5,000 which is a good trade-off between a proper definition of a small town and a sufficient number of observations in all three categories.

Figure 2: Distribution of all observations



Source: RTR survey data

Figure 3: Distribution of observations in communities with less than 25,000 inhabitants



Source: RTR survey data

However, it can be argued that 10,000 is to high for what people usually call a village or small town. The small amount of observations under 5,000 can probably be explained with data incompleteness. Around 115 people did not report their ZIP code and a significant number of all respondents misstated at least one number of their ZIP code that is why I had to reduce the four digit ZIP code to a three digit one thus the last digit of each ZIP code was deleted e.g. 5020 was reduced to 502. It was not possible to just simply exclude all observations with an incorrect ZIP code because this would have reduced the sample size significantly. ZIP codes are distributed according to geographical criteria therefore similar ZIP codes indicate low spatial distance. A three digit ZIP code is not unique anymore so I merged all identical codes into one and summarized the number of inhabitants of each community. In fact I merged smaller neighbouring towns into one ZIP code and summarized their inhabitants. This procedure led to a higher sample average and more observations living in 5,000 plus towns but it was possible to assign a community size to almost all observations besides those who did not report a ZIP code at all. Hence the price for a comfortable sample size which allows larger models and more testing is a bias in the urban-rural variables.

The following descriptive statistics and models are computed with the 10,000 and 50,000 cut-off points and afterwards all results are checked for robustness with the 5,000 and 15,000 categories.

Both surveys are representative for Austria on national and regional level. As I compute penetration rates for the rural-urban dummies which are defined via subjective criteria of the author its uncertain if these dummies are representative for the Austrian population in the given categories. Hence the results are compared with those from the annual EU Community Statistics on Income and Living Conditions (EU-SILC) 2006 survey which collected the Internet access rates for households by various sociodemographic criteria among many things including the number of inhabitants in a community[34]. This survey is representative for Austria on community, city, state and federal level thus it is an excellent chance to check the NASE findings with data from a different source. Statistik Austria kindly provided 50 percent of the full sample which are more than enough observations (N = 6,000) to compute a proper model. Unfortunately, data for 2009 is not available.

Another possible measure for an urban-rural gap in residential Internet access is the population density of a community. With the area of each town from the GIS, the population density can be calculated and is assigned to each observation. Hence an additional variable with the population density of each respondent's town is created.

[34] For more information on EU-SILC and its methodology see http://www.statistik.at/web_de/frageboegen/private_haushalte/eu_silc/index.html (27th July, 2009)

The EU Broadband coverage report[35] defined an urban city as a community with 500 or more inhabitants per squared kilometre (inhabitants/km^2). A suburban city had a population density between 100 and 500 inhabitants/km^2 and a rural town less than 100 inhabitants/km^2. No detailed explanations were provided why exactly these cut-off points are used and not others. As I have to check for robustness, I shifted these points slightly and documented the change in the number of observations. The aim is to create roughly even distributed groups.

The EU cut-off points turn out to be good starting points as the three groups are quite balanced. Table 12 contains the result for slight shifts of the cut-off points. In 2006 the small town group is the largest in front of the metropolitan group. It is not reasonable to reduce the smaller cut-off point to 50 as the middle town group increases drastically in both years. Histograms illustrate the distribution of the observations under 300 inhabitants/km^2 and back this conclusion. A shift of the second cut-off point to 800 causes a moderate increase of the mid-size community group and leads to complete balanced groups in 2009. The 2006 groups are approximately even distributed only the metropolitan group lags behind a bit.

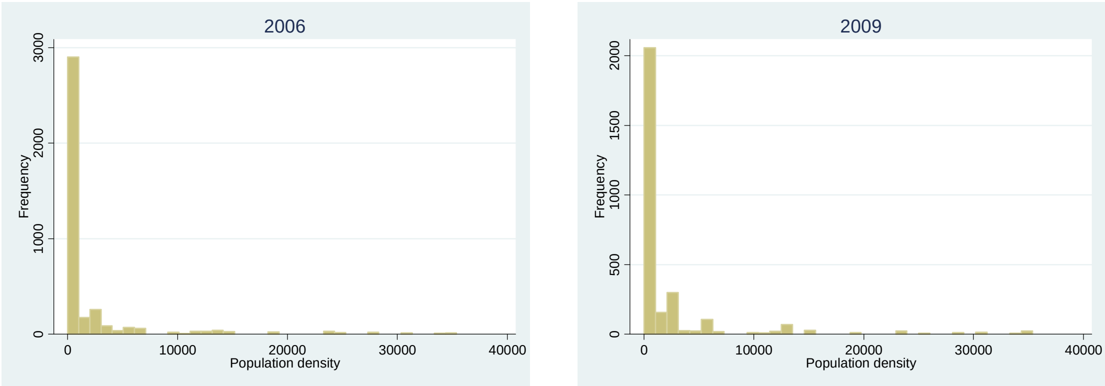
Table 11: Number of observations for different "population density" cut-off points

Area in inhabitants/km^2		*100 - 500*	*50 - 500*	*100 - 800*
	<	1,482	586	1,482
2006	=	1,084	1,980	1,321
	>	1,326	1,326	1,089
	<	975	348	975
2009	=	817	1,444	982
	>	1,145	1,145	980

Source: RTR survey data

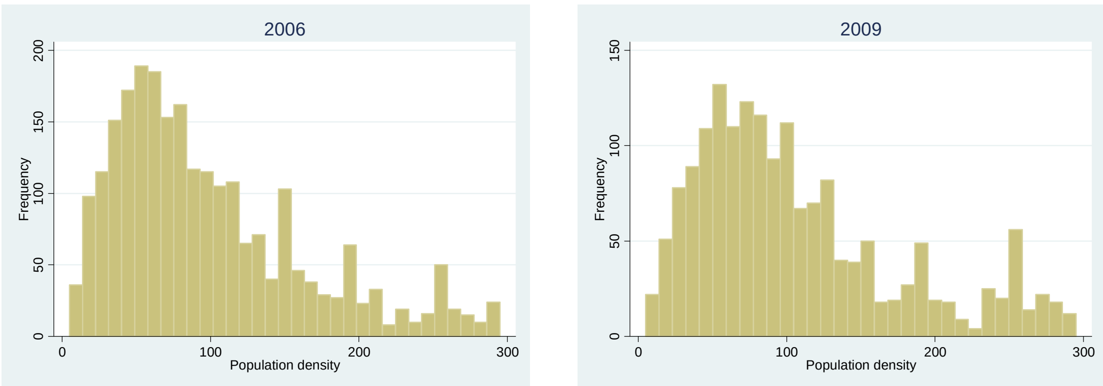
[35]For further information on the methodological part of the report see http://ec.europa.eu/information_society/eeurope/i2010/docs/benchmarking/broadband_methodology_06_2007.pdf (14th July, 2009)

Figure 4: Distribution of all observations



Source: RTR survey data

Figure 5: Distribution of observations in communities with less than 300 inhabitants/km^2



Source: RTR survey data

Hence 100 respectively 800 inhabitants/km^2 are chosen as cut-off points for the population density variables and all following descriptive statistics and regressions are computed with these groups. Robustness checks of the regression results are done with the other two pairs of points.

In the following section, penetration rates for each group of spatial variables are computed and discussed.

Table 12: Residential Internet access rates in percent for all spatial variables

Spatial Variables	*2006*	*2009*
Population Size		
<10,000	49.9	59.2
10,000 – 50,000	55.4	73.5
> 50,000	52.9	62.3
Population Density		
< 100	51.5	60.8
100 – 800	55.1	61.7
>800	52.6	90
States		
Vienna	54.8	70.5
Lower Austria	55	57.8
Upper Austria	59.1	66.3
Burgenland	41.6	58.6
Styria	42.8	58.3
Carinthia	46.2	60.9
Tyrol	54.7	61.6
Vorarlberg	46.8	59.6
Salzburg	51.2	63.7
Suburban communities	55.5	80

Source: RTR survey data

Results from the descriptive analysis are not as expected because the Internet penetration in large cities is not higher than in mid-size communities. In both samples the penetration rate is the highest for the 10,000 to 50,000 people category while the metropolitan variable ranks second. In 2006 the differences between the three categories are small but significant whereas in 2009 especially the gap between middle size towns' Internet use and metropolitan Internet use is wide. A possible explanation is hard to find. Maybe the small number of large communities over 50,000 is the problem. Vienna's rates in both samples are higher than the metropolitan average and represent the majority in the large city sample (2006: 61.6%; 2009: 59.7%). Hence the penetration rate in the other eight cities has to be lower than the high Viennese average. For example, Graz's rate in the 2006 sample is only 38.2 percent. Compared to the national mean, the 2006 urban community result is just average while the 2009 rate is below the Austrian penetration rate of 65.4 percent.

The comparison of the descriptive results of the 2006 RTR NASE data with the 2006 EU-SILC survey numbers as an indication whether the rural-urban dummies are representative for Austria or not delivers similar results. The penetration rate for communities under 10,000 inhabitants is 53 percent with the EU-SILC data while in the NASE sample it is 49.9 percent. EU-SILC provides statistics for towns larger than 10,000 and 100,000 inhabitants. Surprisingly the rate for cities with more than 100,000 is with 41 percent lower than the rate for communities with more than 10,000 inhabitants (49 %). Vienna has a penetration rate of 59 which is 4.2 percent higher than the result from the NASE data. The EU-SILC statistics indicate the same as the NASE results: Middle size towns have a higher average Internet penetration rate than large cities and small towns under 10,000 are not lagging behind.

Results for the Internet penetration rates of population density groups back our spatial gap assumption[36]. For the 2006 data the differences are small and similar to the number of inhabitants approach. The middle size town group ranks first with a slightly over the average Internet usage rate (52 %). The over 800 inhabitants/km^2 group's rate is just above average while the small town group lags behind a bit. The 2009 results are impressing because the large city group has a 90 percent penetration rate while the other two groups did not even reach the survey average of 65.4 percent. At this point population density seems to be the better approach to measure the spatial gap. In the regression part of the thesis models with both approaches will be computed.

The suburban variable created to test the effect of neighbouring a large city while not being considered to be metropolitan has an average population size of about 21,000 in both surveys and a sample size of 355 (2009) respectively 425 (2006) observations. The results are impressing. In both samples the rate is above average especially the 2009 result with a rate of 80 percent is astonishing.

At state level, significant increases in three years can be observed. Lower Austria is the only exception with almost no increase which lead to a drop back from a forerunner state to the last place in 2009. Vienna and Upper Austria have the highest penetration rates while states with a predominantly rural structure such as Burgenland and Styria lag behind. Hence the state variables can account for differences in Internet use in the model and indicate regional differences.

No EU-SILC data is available for penetration rates by population density hence no further check for representativeness is possible. But EU-SILC provides a different interesting measure for the degree of urbanization: A combination of population density and size of each

[36] Note that exceptional small towns can have an extraordinary population density due to the high concentration of people on a small area. Survey data was checked for these outliers and none were found.

town. According to the Eurostat definition a high degree of urbanisation is given when the population density in a community exceeds at least 50,000 inhabitants and more than 500 inhabitants/km^2. A moderate degree of urbanisation is characterised by more than 50,000 inhabitants and a population density between 101 and 500 inhabitants/km^2. All other communities form the non-urbanised category. It is easy to create dummy variables following these definitions to have an additional chance to check for an urban-rural gap. EU-SILC data itself can be used for a robustness check once again as data also includes this measure.

The spatial dimension of the digital divide will be tested via several variables. First, three dummy variables defined by population size, density or degree of urbanity are created, a suburban dummy variable tries to measure the effect of surrounding a large city while being considered rural due to a small number of inhabitants and at last a dummy variable for each state is part of the model with Vienna as the reference group.

4.4 Peer Group effect

As mentioned earlier, the challenge is how to measure the peer group effect. Agarwal et al. (2005) had access to FIPS data from the 2000 US census and survey data collected by the Pew Charitable Trust. The aim is to include the peer group's penetration rate to check for significant influence as the argument is that people's decision depends on their peer group's attitude towards Internet use. This requires data from the smallest entity available which would be ZIP code areas in Austria. Actually a peer group is defined as "a group of friends that a certain person will try to impress to get their bond, social status, and interests."[37] Thus the ZIP code area is only a proxy for a peer group. In fact, data concerning real peer groups for juveniles or adults would be needed but due to unavailability the FIPS respectively the ZIP code is used. Unfortunately the Internet penetration rate for each ZIP code area is not available and with the obtained survey data it is impossible to calculate the penetration rate for each area due to a lack of observations.

Instead an alternative possibility could be to compute the rate for the sum of various ZIP code areas e.g. the traditional quarters of a state. ZIP areas overlap the quarter borders thus it is impossible to create quarter data with ZIP data without serious inaccuracies. Moreover, for some states there are no traditional quarters and smaller reasonable entities than the state itself would have been required which is an almost impossible task. The peer group effect cannot be tested on state level where data is available due to the fact that it

[37] Siegler, Robert (2006): "How Children Develop, Exploring Child Develop Student Media Tool Kit & Scientific American Reader to Accompany How Children Develop" in New York: Worth Publishers.

cannot be argued that someone from the south of Lower Austria feels pressure to conform to the behaviour of an inhabitant of Northern Lower Austria. A peer group is definitely smaller than a state's population which can be up to 1.7 million people in Austria. State dummy variables check for regional differences in Internet use in the model anyway as argued in the section before. But state data is useless when it comes to test for peer effects.

Due to the lack of proper data, it is not possible to include peer effects in the model. Real ZIP area data from a representative survey with a sufficient sample of respondents is needed for further research on peer effects in Internet use. Results from earlier research and Donat (2008) indicate significant peer group influence even when checked for sociodemographic impacts. Hence peer effects in Austria are a potential focus of future research.

4.5 Hypothesis testing

Two interesting assumptions mentioned in chapter 1 are subject of further discussion in this thesis. The first deals with the relation between job position and Internet use where literature suggests that job related Internet use increases the probability that the employee has a residential Internet access as well. The second hypothesis concerns the effect of costs. The demand of Internet as a normal good should increase as prices drop thus a negative relation between the prices of available Internet services and Internet use is assumed. Variables to test both hypotheses are created and added in the model. Further discussion is following.

4.5.1 Job - Internet use Relation

Only "IFES" in 2009 collected the respondent's employer-employee relationship e.g. self-employed, civil servant, employee etc. Being an old age pensioner, being a housewife/houseman, being in maternity leave or jobless was collected as well. Some categories are merged into one due to a too small number of observations for those categories or strong similarities between them[38]. For each category a dummy variable is created and included in the model. Table 13 contains penetration rates for these categories. The idea is that employees who experience the advantages of the World Wide Web during work time are not willing to spend their leisure time without a residential Internet access.

---

[38] Employees, skilled engineering worker („Facharbeiter") and semiskilled workers („Angelernter Arbeiter") are in the employee variable. Farmers („Landwirt") and keeping household („Im Haushalt tätig") form the farmer/household variable and apprentice ("In Ausbildung/Lehrling") and maternity leave ("Karenz") are in the apprentice/maternity variable.

Hence being a member of a job group where Internet use can be assumed e.g. employees or civil servant should have a positive impact on the Internet penetration rate.

Table 13: Residential Internet access rates in percent by job position

Position in Labour Market	N	Penetration rate
Self-employed	268	80.6
Civil servant	206	82.7
Leading employee	292	87.9
Employee	1038	78.15
Farmer/Household	128	55
Old age pensioner	774	33.6
Unemployed	50	48.6
Apprentice/Maternity	276	94.3

Source: RTR survey data

The lowest rate in Internet use is found among old age pensioners. As being a pensioner is not a profession of its own this effect is actually the yet mentioned grey gap effect from chapter 2. The same effect just in the opposite direction may be observed for the apprentice/maternity leave category[39] which has the highest rate with 94.3 percent. Primarily teenagers start their work life as an apprentice and Internet is wide spread in this age group. Hence this effect is probably caused mainly from age and not from profession. Results for the other job groups are as expected. Employees, farmers and housewives/housemen have a significant lower penetration rate than Internet related jobs like leading employee or civil servant. Especially unemployed individuals have a very low rate with 48.6 percent. The job variable should have some explanatory power in the model.

An additional question concerning the level of employment was collected as well. It could make a difference whether an individual works full time or is just fractionally employed ("geringfügig beschäftigt"). Usually a full time employee deals with more important matters than a fractionally employed one hence Internet use is more likely to be found in this group. This dimension of the job – Internet usage relation is kind of a supplement to the job group variable approach. Only respondents who were classified as civil servant, leading employee, employee, worker or as an apprentice were asked this question.

[39] Only 51 of the 276 observations in this category are on maternity leave thus an absolute majority works as an apprentice.

Table 14: Residential Internet access rate in percent by level of employment

Level of Employment	N	Percent
Full Time	1,178	84.4
Part Time	338	82
Fractionally employed	141	73.4

Source: RTR survey data

The difference in Internet use between full time and part time employed respondents is small while fractional employed people have a significant lower penetration rate. Hence this additional variable could have explanatory power and will be part of the basic model.

4.5.2 Cost – Internet Use relation

Internet access can be considered to be a normal good from an economist's perspective. As prices decrease for various Internet services, demand should increase. Hence a negative correlation between costs for Internet access and Internet penetration rate is assumed. As mentioned earlier, the challenge here is to find a proper measurement for Internet costs. The actual expenses of each respondent are available from survey data.

In 2006 the Internet market was significant different from today's market. The share of Narrowband Internet households was about one third while today hardly anyone uses this technology to go online. Mobile Broadband Internet access was in an early stage of diffusion. The incumbent TA had significant higher prices although challenged by various local and nationwide market players. A fixed line telephone access was compulsory to get an ADSL access from TA thus in addition to € 19.90 for the Internet service the consumer had to pay another € 15.98 for the POTS[40]. Smaller providers offered significant cheaper products e.g. a double play package from Tele2 for € 27.90. As there are several hundred small ISPs in each state which sometimes only operate on regional level it is almost impossible to assign each observation the lowest price for the available options. Moreover it is the question if it is reasonable to use the household's expenses. On the one hand, individuals are sometimes too lazy or uninformed to change their Internet provider even in face of significant lower prices. The effect of lower costs might not be measured with the available data. On the other hand, it is bold to assign a consumer the cheapest available product without considering the different product characteristics e.g. a Narrowband access is cheaper than a Broadband one but for obvious reasons. A brief look at the historical prices of all ISPs for 2006 reveals that alternative providers had significant lower prices thus the availability of alternative Internet

[40] Prices refer to November 2006.

services can be used as a dummy for costs. The assumption is that lower costs due to alternative cheaper products increase the probability that an individual has a residential Internet access. An alternative ISP dummy variable is created with the availability data of all major providers and used as a proxy for Internet costs.

The variations in price level were significantly higher in 2006 than those in 2009. In 2007, mobile Broadband prices dropped and the incumbent, TA, reacted with a triple play package including Broadband access for € 19.90 which initiated a general decrease in prices of all ISPs. The diffusion of mobile Broadband and the introduction of flat rates for Broadband products reinforced this effect. These days, consumers can get Broadband access basically everywhere for about € 20 from at least one provider thus costs are not a determinant of Internet access anymore and will not be part of the 2009 model.

For about 50 percent of all households an alternative ISP product was available and the Internet penetration rate for those households was with 57.5 percent significantly higher than for households without an alternative. This is an indication for the positive effects of decreased cost but results have to be examined with caution. Unobserved effects can cause the higher penetration rate as no product characteristics are included e.g. better Internet products from alternative providers.

In the next chapter the basic model for both years is introduced and estimation results are presented.

5. Basic Model

First I will present the basic theory behind my model and my motivation to use it. A short model description is following. In the end of this chapter I present the regression results and their interpretation.

5.1 Theoretical framework and model description

I chose a Logit model to test for influences on home Internet access because the depending variable is dichotomous and it is not obvious why the underlying distribution is rather normal then logistic as Internet access is a qualitative characteristic not a quantitative. The logit model is quite common in literature when it comes to examine the deployment of Internet or Broadband access (see Donat (2008); Chaudhuri (2005) and Whitacre (2005)).

Suppose that Xβ is the linear combination of several characteristics of Internet users then the logistic model specifies the probability of having Internet with

$$prob(INTERNET) = \frac{e^{X\mathrm{B}}}{1+e^{X\mathrm{B}}}$$

Hence the probability of not having Internet access is

$$prob(NO\_INTERNET) = 1 - prob(INTERNET) = \frac{1}{1+e^{X\mathrm{B}}}$$

Maximum likelihood is used as the method of estimation. The likelihood function is formed as

$$L = \prod_i \frac{e^{X_i\beta}}{1+e^{X_i\beta}} \prod_j \frac{1}{1+e^{X_j\beta}}$$

where i refers to those who have Internet and j to those who have not. The maximum likelihood estimator of β is computed when maximizing this likelihood with the respect to the vector β. For the *n*th individual the probability of having residential Internet access is estimated as

$$\frac{e^{X_n\mathrm{B}^{MLE}}}{1+e^{X_n\mathrm{B}^{MLE}}}$$

Therefore the odd-ratio for the logit model is

$$\frac{prob(INTERNET)}{prob(NO\_INTERNET)} = e^{X\mathrm{B}}$$

All regressions and descriptive analysis in this thesis are computed with *Stata* which has the *logistic* command where the output already includes the odd-ratios. The interpretation of the odd-ratio is not as simple as the interpretation of the linear regression output. One unit of change of the independent variable increases the probability of the event by the factor $e^{X\beta}$. In case an independent variable is categorized into dummies the effect is always relative to the reference group.

First a logit model is presented for both years where residential Internet access is the dependent variable and all other yet mentioned sociodemographic and spatial variables are independent. The 2006 model includes an alternative ISPs variable as a proxy for Internet costs as well hence

$$\Pr_i(INTERNET = 1) = \alpha + \sum_{j=1}^{6} \beta_j S_{ji} + \sum_{j=1}^{12} \gamma_j G_{ji} + \sum_{j=1}^{6} \delta_j A_{ji} + \Theta P_i$$

where $\Pr_i$ is the probability that individual or household i has a residential Internet access. α is the constant, β_j are the coefficients of the sociodemographic variables while γ_j are the ones for the spatial variables. δ_j represent the estimates of the age variables and Θ is the coefficient for the cost influence. No error term is needed as the value of the dependent variable in the logit model is generated via a chance mechanism included in the estimated probabilities.

S_{1i} to S_{6i} represent the sociodemographic influences which are income, household size, gender and education. The reference group for education is compulsory school hence the model computes the probabilities of all other educational categories in relation to compulsory school. Men are the reference group in the gender category. The spatial dimension of the digital divide is measured in G_{1i} to G_{12i} which contains the categories formed by number of inhabitants, states and the suburban communities. Vienna and the small town variable are the references groups for the states and the number of inhabitants respectively population density dummies. A_{ji} includes all age groups with the 16 to 24 group being the reference group. The influence of costs on the decision of individual i is tested with the last term P_i.

The model for the 2009 data is very similar. The only difference is an additional educational group (apprenticeship) and instead of the cost influence of P_i, the influence of employment E_{ji} is added to the model. E_{1i} to E_{7i} represent the eight job categories with self-employed as the reference group. Being full and part time employed are the last two terms with fractionally employed as the reference group. Θ_j are the coefficients for the employment variables.

$$\Pr_i(INTERNET=1) = \alpha + \sum_{j=1}^{7} \beta_j S_{ji} + \sum_{j=1}^{12} \gamma_j G_{ji} + \sum_{j=1}^{6} \delta_j A_{ji} + \sum_{j=1}^{9} \Theta_j E_{ji}$$

The procedure is the same for both models: I will estimate the model and present the results. Afterwards the influence of the variables is compared over the years. The robustness of the results is checked with the help of the alternative urban-rural definitions and the EU-SILC data. All Tables will also include McFadden's pseudo R-square which is the default goodness-of-fit criteria in the *logistic* command with sample weights to compare the different models. The aim is to find the best model to explain residential Internet access rates and examine possible gaps.

5.2. Estimation results

First I will present the results for 2006 with number of inhabitants and population density as a filter for rural and urban communities. In each section the models with the alternative categories are also discussed right away to check for robustness. Afterwards the model is computed with the EU-SILC data and the degree of urbanity variables. The same procedure is used for the 2009 model except for the EU-SILC robustness check due to data unavailability.

As I use sample weights[41] to guarantee the representativeness of my results *Stata* uses *robust standard errors* which means that instead of the conventional estimator of variance the Huber-White standard errors are used for all regressions.

[41] Called *pweight* in *Stata*.

5.2.1. 2006 Logit Model

This model includes all available sociodemographic characteristics, different spatial variables depending on which definition is chosen, a suburb variable to measure possible effects of a metropolitan neighbourhood and a dummy indicating the availability of alternative ISPs which serves as a proxy for costs.

Previous to the first estimation, I took a look at the correlations between the federal states, the urban-rural and the suburban variables. In case these variables would be high correlated, a possible spatial divide could be difficult to measure with a model including all the spatial variables because the effect could be split up. The only substantial correlations for the 2006 NASE spatial variables were computed between Vienna and the large city measures e.g. the variables for communities with more than 50,000 inhabitants, a population density above 500 respectively 800 inhabitants/km^2 and high urbanised communities according to the Eurostat definition. This finding is not surprising at all. The correlation coefficient is around 0.7 for all of these measures. However, this is the only substantial correlation whereas all the other computed correlations are clearly under 0.5 respectively -0.5. Hence it should be possible to measure a geographical divide with a model containing all the spatial variables.

5.2.1.1. Population size

Table 22 presents the coefficients for all models with number of inhabitants as the filter.

>Insert Table 22 about here<

All models are highly significant which is no surprise as the sample size of about 4,000 observations and the large number of variables almost guarantee for some significance. A first brief look at the results reveals that the R^2 is the same with 0.35 for all models and that sociodemographic influence is evident while the spatial influence is questionable.

In the original model all sociodemographic variables are highly significant. Reported estimates are odd-ratios hence an increase in household size by one unit increases the chance of having a residential Internet access by 1.23 times *ceteris paribus*. Imputed log income even has a stronger influence with an odd-ratio of 3.26. A small gender gap seems to exist in Austria because men have a 1.34 higher chance of being Internet users than women. The reference group for the education variables is compulsory school. As expected do all superior school degrees have a positive influence on Internet access compared to the lowest category. This influence is rising with the degree of education. Increasing age does have a negative influence on the depending variable because the reference group – namely the 16 to 24 years olds – already have a high penetration rate. Only the three oldest age groups differ significantly from the youngest especially with substantial lower probabilities. The 70 plus group is about 16 times[42] less likely to have residential Internet access than the 16 to 24 aged youngsters. The first model confirmed all assumed sociodemographic gaps.

For the measurement of the spatial dimension in the original basic model two dummies were created. The middle size town variable consists of all communities with a number of inhabitants between 10,000 and 50,000 and the metropolitan variable with all observations larger than 50,000. Small towns have a population smaller than 10,000 inhabitants and are the reference group. The results are mixed with the middle size town variable being insignificant and smaller than one which means that its influence on having home Internet access in comparison to small communities is *lower*. The same unexpected coefficient is found for the large city category which is significant at a 5 percent level. According to the model the inhabitants of cities with a population of over 50,000 people are 1.7 times less likely to have home Internet than rural communities members. The reviewed literature suggested the opposite effect. The continuous "number of inhabitants" variable has no influence on Internet access. Some state variables do have statistically significant influence. Vienna is the reference group hence coefficients smaller than one would be expected as the federal capital is a real city with excellent ICT infrastructure and a competitive market. Surprisingly Upper Austria has a significant odd-ratio of 1.6 while all other state variables have a coefficient smaller one. This finding suggests that inhabitants of Upper Austria are more likely to have residential Internet access than Viennese. Burgenland is known as the least developed part of Austria hence it is no surprise that the chance of having home Internet there is 3 times lower than in Vienna. Its coefficient is the only significant one at a level of 0.1 percent. Styria and Vorarlberg do have significant influence on the depending variable whereas all other not yet mentioned states have not. These results suggest existing regional differences which are difficult to interpret. A possible explanation for the significant Burgenland odd-ratio was already mentioned. But it is unclear why Vorarlberg and Styria influence the home Internet use and other states do not respectively what's the substantial regional difference between these two and the rest.

The suburban variable which includes all non-urban communities bordering large cities is insignificant hence these smaller towns do not seem to have an advantage from their metropolitan neighbourhood. The cost hypothesis is confirmed as the availability of an alternative ISP, which is the proxy for costs, increases the chance of having home Internet by 2.7 times.

---

[42] In case of odd-ratios smaller than one, $1/\exp(\beta)$ is computed to get a better idea of how many times it is less likely to have home Internet access.

Results of the original basic model suggest a gender, education, income, household size and age gap while there is only moderate evidence for the spatial dimension of the digital divide. Regional differences on state level are evident but rural-urban influences are not highly significant. Costs influence home Internet rates. The next step is to compare these results with the three other models based on the “number of inhabitants” urban-rural variables as listed in Table 22.

The quality of the other models measured with the help of the R^2 is the same with 0.35. The sociodemographic coefficients are almost the same in Model 2, 3 and 4. The already indicated gaps in education, gender, household size, age and income are confirmed in their significance and extent. The same states have significant coefficients than in the first model and even the surprising Upper Austria result is backed. The availability of an alternative ISP has statistical significant influence on the depending variable in all three models.

For the most interesting dimension of the digital divide - the spatial one - results are ambiguous. Keep in mind that the aim of these three other models is to check the results of model 1 for robustness. In model 2 the middle size town is insignificant whereas its metropolitan variable is significant even though its coefficient is smaller than 1. In model 3 and 4 the cut-off points for the rural-urban definitions are shifted again according to chapter 4.3 but not a single significant rural-urban odd-ratio could be found which casts serious doubts on the findings of model 1 and 2. The following section will take a look at the model with population density as the rural-urban definition and examine the crucial question whether there is a spatial gap or not in 2006.

5.2.1.2. Population density

Descriptive statistics in chapter 4.3 have shown that population density as a filter for urban-rural differences delivers the results which are suggested by the scientific literature. Urban communities had higher penetrations rates than their rural counterparts hence the same model is computed with the population density variables. In model 1 a middle size town is defined by a population density between 100 and 800 inhabitants/km^2. Communities with a density under 100 inhabitants/km^2 are considered to be rural and metropolitan communities have a density greater than 800.

Results for model 1 are very similar to those from the “number of inhabitants” models. The sociodemographic characteristics are all significant except for the age groups under 50. The odd-ratios are almost the same hence gender, household size, income, education and age have statistical significant influence on residential Internet access. The ratios of the rural-urban variables are both smaller than one which suggests that it is less likely that urban citizens are onliners in comparison to their rural colleagues. But both are insignificant at a 5 percent level. On state level there are again some significant differences. Upper Austria is the only state with an odd-ratio greater than one hence its population is 1.56 more likely to be online than Viennese. Like in the section before, Burgenland has a highly significant negative influence ($p < 0.000$) where as Styria and Vorarlberg have p-values of 0.022 respectively 0.034. Suburbs do not have a premium in Internet access and costs influence the dependent variable statistically significant in about the same extent a college degree does. The unexpected results of the population size models – namely urban odd-ratios smaller than one and an odd-ration for Upper Austria greater than one – are confirmed.

In model 2 and model 3 the cut-off points are changed for a mid-size town to having a population density between 100 and 500 respectively 50 and 500 inhabitants/km^2 as argued in chapter 4.3. Results of model 1 are backed in almost all cases. The sociodemographic influence on residential Internet access is evident and the rural-urban variables are insignificant. Small changes concerning regional differences at state level are observed as the p-value of Upper Austria drops from 0.033 in model 1 to 0.009 in model 2 respectively to 0.006 whereas the odd-ratio of Vorarlberg turns insignificant due to the changes. All other ratios and p-values stay as they are. The R^2 is exactly the same for all three models with 0.353.

>Insert Table 23 about here<

The population density models reinforced the doubts of the initial spatial gap findings because not a single odd-ratio was significant although population density was assumed to be a good filter for a possible spatial divide. Sociodemographic gaps are substantial and confirmed in every model. Regional differences do have influence on home Internet access although not all state ratios are relevant. Alternative ISPs seem to have a positive influence on the market price and increase the chance that the individual i has an Internet access.

As mentioned in chapter 4.3 EU-SILC provided an interesting measure of urbanisation consisting of a combination between population size and density. Full urbanised communities are supposed to have more than 50,000 inhabitants and a density of at least 500 inhabitants/km^2. Moderate urbanised towns have also more than 50,000 inhabitants but only a population density between 101 and 500 inhabitants/km^2. All other towns are considered to be rural. These definitions are the official Eurostat ones. The last model computed with the 2006 NASE data will include these urbanisation variables as a measure for a possible urban-rural gap.

5.2.1.3. Degree of Urbanisation

Population density and size of each observation is available for the 2006 and 2009 data therefore it is no problem to create dummy variables according to the Eurostat definitions. Only the rural-urban variables are different in this model. Results are presented in Table 24.

>Insert Table 24 about here<

The R^2 of the model is the same as for the other 2006 models. The sociodemographic odd-ratios are all significant and almost identical compared to the other models. Results for the spatial variables have changed. The surprising Upper Austria odd-ratio is backed as well as the statistical significant negative influence of Burgenland, Styria and Vorarlberg. A household living in Burgenland is 3 times less likely to have Internet access than a Viennese. Living in Vorarlberg reduces this chance by 1.9 times. The odd-ratio for Salzburg is with 0.651 the same as earlier but for the first time it is barely significant with a p-value of 0.05. Slight evidence for an urban-rural gap is found as the ratio for full urbanized cities has a p-value of 0.029. The odd-ratio is with 0.603 smaller than one which means that being an urban citizen reduces the chance of having residential Internet access by 1.7 times compared to a non-urban citizen. Model 1 and 2 of section 5.2.1.1 had the same unexpected ratio for urban communities. Moderate urbanized towns do not have an influence on the depending variable which is in line with the earlier results. Alternative ISPs increase the probability of being online by about the same extent as having a college degree does.

Conclusions for the 2006 NASE results

The results for the 2006 NASE data reveal a substantial sociodemographic gap while the spatial gap is doubtful. The gender, age, household size, education and income divide was significant in every model with almost the identical extent. It is very important to realise that a model just including these sociodemographic characteristics has a R^2 of about 0.33 hence almost all of the explanatory power of the depending variables is contributed by these variables while the spatial ones in fact contribute hardly anything. In addition, I performed an F-test at the end of each estimation with the H_o that the coefficients of the large and mid-size community variables equal zero. The H_o was never rejected. Considering these facts and the mixed evidence for a rural-urban gap, the spatial dimension of the digital divide in Austria is questionable. Regional influence on state level is backed and the surprising odd-ratio for Upper Austria is confirmed in all models whereas the findings for a rural-urban gap were rare. The mid-size town variable is not significant in a single model and the odd-ratio of the metropolitan variable indicates a reduced probability of having Internet when compared to rural towns which is the opposite of what was suggested by literature. Especially when the cut-off points are shifted, which was done as a check for robustness, the influence seems to disappear. Significant influence on the residential Internet access rate by the large city variable was only computed in three out of eight models.

Sociodemographic characteristics are the major determinants of home Internet access in 2006. Regional influence is present but there is certainly not enough evidence to state a rural-urban gap with the 2006 NASE data.

In the following sections these findings for 2006 are checked with different data collected for the EU-SILC and provided by Statistik Austria.

5.2.2 EU-SILC

According to its name the survey collected a lot of information about income and living conditions including residential Internet access, age, income, education, job position and geographical information. Due to concerns of data abuse the most precise information on the respondent's residence was the state. In addition there is a city variable which categorised each observation depending on the number of inhabitants of its community. Towns with a population under 10,000 inhabitants form one group and cities larger than 10,000 another. A third group consists of all cities larger than 100,000. Unfortunately it is not possible to create the exact same urban-rural variables as in the 2006 model with this data. Hence the reference group in the EU-SILC model will be communities with a population size under 10,000 and the mid-size towns are defined by a size between 10,000 and 100,000 inhabitants. Every observation above 100,000 inhabitants is categorised as large city. The ZIP codes are not available for each observation for obvious reasons thus it is not possible to compute the population density with the EU-SILC data to check the results of section 5.2.1.2.

Once again all correlations between the spatial variables are examined. Vienna is high correlated with the large community variables (≈ 0.75), which is expected and the way it should be. All other correlations are clearly under 0.5 respectively -0.5 hence it should be possible to estimate an urban-rural gap with the spatial variables.

Education, age, gender, household size, income and state will be part of the model with the same reference groups as in the 2006 model. One additional education category is available called "no compulsory school" which is included to the model. "Compulsory School" will stay the reference group to make a comparison with the 2006 odd-ratios possible. I expanded the EU-SILC model a little bit and added job position, degree of employment and household with children variables for two reasons: (1) this data is available and it is interesting to check whether these characteristics have influence on residential Internet access or not. (2) These variables, except for the household with children dummy variable, are part of the 2009 model and thanks to the EU-SILC model I can examine their influence over time. Job position and degree of employment data was not collected in the NASE 2006.

First, the EU-SILC model is computed with population size variables to compare these findings with those from the 2006 model to check if there are substantial differences which would question the representativeness of the NASE data on city level. Afterwards the results for the model with degree of urbanisation as rural-urban measure are presented. Table 25 contains both models.

>Insert Table 25 about here<

5.2.2.1 Population size

The model itself is highly significant but its R^2 is clearly lower with 0.2 than the 2006 R^2 (0.35) which indicates that the 2006 model does a better job in explaining the residential Internet access rates than the EU-SILC model with these variables. Household size has significant influence on the dependent variable with an odd-ratio of 1.5. The first interesting result is that there is no statistical significant difference in home Internet access between men and women. Hence the earlier stated gender gap is questioned. Donat (2008) conducted her own survey and came up with the same result for Austria but NASE data did not back her findings so far. Juveniles aged 16 to 24 form the reference group for age. Their Internet penetration rate is clearly above average hence all other age groups have an odd-ratio less than one which means they are less like to be online in comparison to young people. The only insignificant ratios are listed for people aged between 40 and 60. Respondents in their mid-twenties are 1.5 times less likely to have an Internet access whereas being an elderly person above the age of 70 reduces this chance by 6 times. All education variables have a strong influence on the dependent variable and as expected the more educated an individual i is the higher is the probability that home Internet access is present. A person with no compulsory school degree is 4 times less likely to be online whereas a university degree increases this probability by 9.6 times in comparison to a person with a compulsory school degree. Income and having children in a household are insignificant. One hypothesis of this thesis is the relationship between job positions and Internet use as presented in section 4.5.1. I chose the same reference group ("self-employed") as for the 2009 model to assure an over the years comparison. *Stata* only computed one significant job position odd-ratio for the EU-SILC model which is for the category "employee". Members of this group are 1.4 times less likely to use home Internet. All other job positions have an odd-ratio larger than one which means they are more likely to have residential Internet but none of them is significant. Actually this result was not expected because unemployed or retired persons had lower penetration rates than self-employed respondents according to the descriptive statistics. Even the degree of employment turns out to have no significant influence hence employment variables seem not to provide any explanatory power to the model.

Results for the spatial dimension of the digital divide are mixed again. For the first time five states have significant odd-ratios and some of them are even highly significant. Lower Austria, Carinthia and Tyrol influence the dependent variable although they never had p-values smaller than 0.05 in the 2006 NASE model. Almost all state odd-ratios are smaller than one which was expected. A person from Styria is 2.2 times less like to have residential Internet than a Viennese. Vorarlberg is the only state with a ratio greater than one. A brief look at the penetration rate calculated from the EU-SILC data for this state delivers an explanation. 64 percent of all households had an Internet access which is the highest rate nationwide. The assumed regional influence is backed by this model. The urban-rural variables defined by the number of inhabitants turn out to be insignificant and have once again a ratio smaller one. The most important fact concerning this model is that its reduced version without the spatial variables has a R^2 of 0.19 and that an F-test does not reject the H_o which states that the urban-rural coefficients are equal zero. Hence the explanatory power of the spatial variables is almost equal to zero thus residential Internet access seems to be a function of the sociodemographic characteristics.

Comparison with the 2006 NASE results

The aim of the EU-SILC model estimation is a comparison with the 2006 NASE model results to check for possible differences or confirmation of the 2006 findings. The sociodemographic influence on the access rate is confirmed. The extent and significance of the household size, age and education gap differs just a little bit. The age effect in the 2006 models appears to be stronger for the elderly people because their odd-ratios are smaller whereas in the EU-SILC model this is true for the younger age groups. In general the effect has the same direction hence one can confirm the age gap. The odd-ratio for household size is the same for both models. The differences in home Internet access can partly be explained by education variables. This effect is even stronger in the EU-SILC model than in the 2006 one. The income and gender gap of the NASE model is not confirmed. As noted earlier, Donat (2008) had the same interesting result for Austria. Reasons for a diminished income effect are difficult to find.

Regional effects are present in both models but with a different extent and significance. The surprising Upper Austria odd-ratio in the 2006 model is not confirmed but

Vorarlberg in the EU-SILC model had also a ratio greater than one. The highly significant Burgenland effect is gone in the EU-SILC model where as Styria is relevant in both models. The regional effect appears to be more developed in the EU-SILC model as Tyrol, Lower Austria and Carinthia also have a p-value smaller than 0.05. The weak urban-rural findings of model 1 and model 2 are not backed with the EU-SILC data because its city variables are not significant.

The EU-SILC model with population size confirmed the sociodemographic gaps in education, age and household size. The regional effects findings are backed as well whereas there is now new indication for an urban-rural divide. It has to be stressed that the spatial influence in all models so far was small and almost all explanatory power came from the sociodemographic characteristics. Until now residential Internet access is a function of social status with almost no spatial influence.

5.2.2.2 Degree of Urbanisation

The rural-urban variables in this model combine population density and size according to the Eurostat definitions which makes this attempt to measure a spatial gap even more interesting. The sociodemographic results are almost exactly the same in both EU-SILC models. Age, education and household size explain differences in access rates but not income, gender or job related variables. Relevant changes are observed for the spatial variables. Tyrol, Carinthia and Lower Austria which had significant odd-ratios for the first time ever are not relevant in this model which is an indication that the findings of the previous model were by chance not due to substantial differences. Styria remains significant and Vorarlberg even improves its p-value to 0.001. All other states are insignificant in comparison to Vienna. Results for the rural-urban variables are surprisingly different and for the first time the way they are expected to be. Moderate urbanised towns as well as full urbanised cities have an odd-ratio greater than one and p-values of 0.001 respectively 0.026. Inhabitants of mid-size towns are 1.26 times more likely to have home Internet access than rural community members and living in a large city increases this probability by 1.32 times. Be aware of the fact that the reduced model consisting only of sociodemographic variables has a R^2 of 0.19 hence the explanatory power of the spatial part is very low. However, an F-test performed at the end of the estimation rejects the H_o that the coefficients for the full and medium urbanised communities are equal zero.

Comparison with the 2006 NASE results

The comparison of the EU-SILC model with the 2006 degree of urbanisation model delivers to a large extent the same conclusions as before. The sociodemographic gaps in age, education and household size are confirmed while the income and gender gap is questionable. On state level the 2006 model had more significant odd-ratios than its EU-SILC counterpart. The degree of urbanisation model with the 2006 NASE data was one of the three models with a significant rural-urban variable and the EU-SILC model confirms this result although the 2006 odd-ratio for large cities is less than one while for the EU-SILC model is greater than one. Hence in one model living in a large city increases the probability and for the other not. Thus both models indicate a rural-urban gap but disagree about the direction of this effect which casts doubts on the existence of this gap again.

Conclusions for the 2006 results

Some effects are found in all models hence their existence is unquestionable. There is an age, education and household size gap in all models which is highly significant. The income and gender gap is only observed with the 2006 NASE data but not in the EU-SILC model. Considering the clearly lower R^2 of the EU-SILC model I put more weight on the NASE model which points towards a gender and income divide. The spatial determinants of Internet use are questionable. Regional influence in some extent seems to be given while there has to be serious doubts concerning the significance of the rural-urban variables. No mid-size town influence could be found and the large city odd-ratio was only significant in four out of ten models including the EU-SILC model with the degree of urbanity definitions which is the only one where the effect points toward the expected direction. Independent of whether a rural-urban influence can be stated or not, the spatial influence in general is subject of serious doubts. The reduced versions of all models containing only the sociodemographic characteristics as independent variables have almost the same R^2 as their full counterparts. Considering this fact and the results of the F-tests, it is more than obvious that the major determinants of residential Internet access are the sociodemographic characteristics of individual i independent of his or her residence.

In the next section the results of the 2009 NASE model are presented and then compared with the 2006 results to examine the developments of the sociodemographic gaps and possible spatial divides. The job related findings can be compared with the EU-SILC model as this data was not collected in the 2006 NASE.

5.2.3 2009 Logit Model

The model contains the well-known sociodemographic characteristics, different spatial variables depending on which definition is chosen and job related variables to check for a possible relationship between job position and Internet use.

The computed correlations between the spatial variables are as expected. Only Vienna is substantially correlated with the large (≈ 0.7) and mid-size community variables (≈ -0.6). All other federal states have correlations with the urban-rural variables clearly under 0.5 respectively -0.5. Hence these spatial variables can be included in the logit model.

5.2.3.1 Population size

The first analysis of the 2009 NASE data is done with the population size as urban-rural variables. In model 1 the initial cut-off points are used while model 2, 3 and 4 contain city variables from shifted cut-off points as argued in section 4.3. Results for all four models are presented in Table 26.

>Insert Table 26 about here<

The R^2 of all four models is the same with 0.31 and a bit lower than the 2006 R^2 but definitely better than the EU-SILC goodness-of-fit measure. All sociodemographic characteristics play a significant role in the various models. A change of one unit in household size, increases the probability of having Internet access by 1.5 times *ceteris paribus* which is the same extent as being male improves this probability in comparison to being female. The imputed income is significant but has an odd-ration of one. Each education category has substantial influence on residential Internet access rates which increases as the degree improves. Compulsory school is the reference group and all effects are relative to this category hence a completed apprenticeship boosts the home Internet probability of individual i by 1.7 times whereas a university degree has an effect of 4.3 times. Only two age groups differ significantly from the 16 to 24 aged people. A member of the 60 to 69 age group is 3.7 times less likely to be an onliner than the reference group. For the 70 plus generation the ratio is even worse with 12.8.

One subject of interest of this thesis is the relation between job and Internet use as argued in section 4.5.1. The 2009 model does not reveal any significant influence on the depending variable except for the Apprentice/Maternity variable. The insignificant odd-ratios are as expected with a ratio smaller than one for unemployed persons, framers and respondents who are keeping a household. Being employed or retired increases the probability in comparison to self-employed people. Having a completed apprenticeship or being on maternity leave makes a residential Internet access 3.1 times more likely. It is questionable if this significant odd-ratio is due to the influence of employment status or just an age effect as the average age of this category is with 23 years very low while the employee group has an average of 39 and the self-employed category 42.6. The sample average is even higher with 46 years. As only one of six job categories is significant, it is not possible to state a substantial influence of the job position on the dependent variable. The related degree of employment delivers equal insignificant results hence there seems to be no measurable relation between job related variables and Internet use.

No evidence for the spatial dimension of the digital divide is delivered by the 2009 model. While three years earlier at least some significant regional influence on state level is found, the 2009 results indicate absolutely no spatial impact. *Stata* computes only one significant odd-ratio on state level for Styria. All ratios are smaller than one for the first time and therefore as expected because Vienna is the reference group. The same is true for the rural-urban variables. Earlier results indicated that metropolitan residence lowers the probability of having residential Internet access which was not assumed by literature. The 2009 model computes odd ratios larger than one for both categories with the large city ratio being larger than the mid-size one but no significant influence at a 5 percent level is observed. The suburban variable measuring the influence of metropolitan neighbourhood does not contribute substantial explanatory power to the model as well. Hence while the spatial impact on Internet use in 2006 and its extent is subject of discussions and some arguments in favour of this assumption can be found, the situation in 2009 has changed. Absolutely no regional or urban-rural influence on the depending variable is found with population size as the filter for urban-rural variables which indicates that the only determinants of Internet use are the sociodemographic characteristics of individual i.

A shift of the cut-off points for the urban-rural variables as a check for robustness does not change the results in any way. The extent and significance of all variables stays the same. There is a substantial gap in home Internet access in household size, income, gender, age and education while job related and spatial variables deliver no explanatory power. The odd-ratios of the state and urban-rural variables are again as expected by literature but insignificant except for Styria. The R^2 of a reduced model only containing sociodemographic variables is with 0.3 almost the same as with the spatial variables and a statistical test whether the rural-urban coefficients are equal zero does not reject the H_o. Hence the earlier mentioned inexistence of the spatial divide is backed.

Comparison with the EU-SILC and 2006 NASE results

Due to three high quality survey samples it is possible to compare the 2009 results with the 2006 NASE and the EU-SILC results to examine possible developments and trends of the determinants of Internet use. Especially the sociodemographic odd-ratios in both years are subject of interest because results so far indicate them as the major factors of residential Internet access. The ratios within the different models of each sample are almost identical hence I will only compare the model 1 results of each section with each other.

The R^2 as the goodness-of-fit measure is the highest for the 2006 NASE models with around 0.35 which is slightly better than the 2009 R^2 with 0.31. The EU-SILC model has difficulties in explaining the depending variable and thus a substantial lower R^2 equal to 0.2.

Gaps in household size, income, gender, age and education are evident in both NASE samples. EU-SILC data questions the gender and income effect but the model itself has a substantial lower R^2 than the 2006 NASE model. The extent of the household size gap increased or remained constant from 2006 to 2009 depending on which results are compared. The 2006 NASE ratio was 1.2 while the EU-SILC one was 1.48. The odd-ratio in the 2009 model is with 1.46 a bit smaller than the EU-SILC ratio but larger than the NASE result. Hence this gap increased by a small portion or remained constant. The income effect decreased according to the comparison of the two NASE samples respectively did not exist in 2006 according to the EU-SILC. Independent of which 2006 sample is used the effect seems to diminish and is probably the weakest of all sociodemographic characteristics influences. The gender gap remained substantial with a small increase from 1.3 to 1.5 over three years. The EU-SILC results did not suggest such a gap. As argued before when discussing the income effect, the gender effect is not the major sociodemographic determinant of the model and its influence is weak in comparison with age and education which seem to be the most important influences on the depending variable. The impact of education on the probability of having residential Internet has slightly decreased but is still a major factor besides age. The EU-SILC ratios are a bit larger than the 2006 NASE ones but both are clearly larger than the 2009 results. Hence this gap is persistent but declining. The same is true for the age effect. The EU-SILC model delivered significant ratios for the younger and older age groups but their impact was not as big as in the 2006 NASE model. The inter-NASE comparison indicates that one age group – namely the 50 to 59 aged people – turned insignificant and the odd-ratios of the significant elderly groups are smaller than three years earlier.

The indicated minor spatial influence is gone in the 2009 model as well as there is no job related influence on the depending variable. Hence even the marginal regional divide of 2006 turns out to be not persistent while there is a substantial household size, age and education divide. The income and gender effect are questionable and their influence is definitely small when compared to the other three relevant sociodemographic factors.

5.2.3.2 Population density

A different approach to measure the rural-urban differences in home Internet access rates is the population density of the respondent's community. The cut-off points are shifted after the ratios for model 1 are computed to check for robustness of the results. The ratios for all three models are presented in Table 27.

>Insert Table 27 about here<

There are no significant changes between the population density and size model results. The sociodemographic gaps in household size, gender, income, age and education are confirmed and have the same extent as before. The ratios of the significant age and education groups are higher than those of the other sociodemographic characteristics hence these two are the major determinants. There are no results that back the assumed relation between job and Internet use. The apprentice/maternity variable again statistically influences the dependent variable but this effect is probably caused by age and not by employment status. As this is the only significant job variable the assumption that job position or degree of employment influences residential Internet access is rejected. The spatial variables are almost all insignificant. Styria and Carinthia are the only states with a p-value of smaller than 0.05. Their odd-ratio is smaller than one which means that people living in these states are 1.7 times less likely to have home Internet access than Viennese. The urban-rural variables have ratios smaller than one which is again not the expected result as rural villages are the reference group. However, not a single ratio is significant. The suburban variable also does not contribute explanatory power to the model.

A change in the definition of the rural-urban definitions has almost no effect of the outcome of the model. The ratios and their significance stay the same with only one exception. Carinthia turns insignificant which indicates that its initial significant result was by chance and is further evidence that there is no essential regional respectively spatial influence. The Styria result remains the same.

Comparison with the 2006 NASE results

Due to missing ZIP codes the population density for the EU-SILC observations could not be calculated hence only an inter-NASE comparison is possible. As the results for the population and the density model are the same the comparison with the 2006 models does not deliver new findings. The gap in household size and gender is persistent and increased a little bit over three years while the income effect almost diminished. Although these three variables are significant, their influence among the sociodemographic characteristics is small. Age and education are the major factors of the model and their effect diminished a bit over three years. The age group 50 to 59 has not a p-value smaller than 0.05 anymore and the ratios of the elderly people increased which means they are more likely to be online than in 2006. Having a university degree makes home Internet use 4.4 times more likely these days while three years earlier it boosted this probability by 6.5 times. The slight spatial influence in 2006 is completely gone in 2009 and the assumed rural-urban divide is rejected. The R^2 of a sociodemographic model is with 0.3 marginal smaller than with the spatial variables. In addition, F-tests back the no spatial influence assumption.

5.2.3.3 Degree of Urbanisation

The last 2009 model is computed with the Eurostat definitions of urbanity as a measure for the urban-rural divide.

>Insert Table 28 about here<

The different urban-rural measures do not change the results for the other variables. The gaps in access rates by household size, income, gender, age and education are confirmed in their significance and extent. Apprenticeship/maternity is again the only job related variable with an impact on the dependent variable hence there is no relationship between job and residential Internet access. The degree of employment as a different measure of this possible relation is also insignificant.

An interesting result for the spatial dimension is found. The moderate urbanized variable has a statistical significant influence on the residential Internet access rate with a p-value of 0.016. Its odd-ratio of 3.5 is larger than the insignificant ratio for full urbanized cities which means that living in a moderate urbanized town increases the probability of having home Internet more than living in a metropolitan area does. At least the full urban ratio is larger that one. The regional influence on state level is gone over three years as the 2009 model has no state variables with a significant ratio except for Styria.

Comparison with the EU-SILC and 2006 NASE results

The comparison with the EU-SILC and the 2006 NASE model indicates the same developments as mentioned earlier. The household size gap is significant in all models while the gender and income gap is questioned by the EU-SILC model. Independently of the question if these gaps exist or not, it is important to notice that the major sociodemographic determinants are age and education. The education gap is persistent but declining and the same is true for age.

The comparison of the spatial ratios delivers unexpected results. The EU-SILC model backed the rural-urban divide hypothesis with two significant odd-ratios larger than one and the full urbanized ration being a bit smaller than the moderate one. The same result was computed for the 2009 NASE model but the difference between the two ratios is clearly larger than for the EU-SILC model and the full urbanized city ratio does not contribute significant explanatory power to the model. The 2006 result for the rural-urban variables were smaller than one which is once again not the way things were expected to be. However, only the large city ratio was significant. While the EU-SILC model in 2006 came up with same result as the 2009 NASE model, the 2006 NASE result contradict these findings. In general the degree of urbanisation measure appears to distinguish more strictly between the three categories and therefore more significant results are computed. The combination of population size and density classifies the observations more accurate than the other two measures.

On state level both 2006 models had several significant ratios while in 2009 only Styria has a p-value smaller than 0.05. The striking finding is that Styria actually had relevant influence in every single of the 18 computed models. It is hard to argue that this is by chance. No other state was nearly as often significant as Styria. Its ratio is always between 0.5 and 0.65 hence its population appears to be approximately two times less likely to have residential Internet access than the Viennese one. A possible explanation is difficult to find. Neither its economic background, nor its geography indicates serious disadvantages for the deployment of Internet or restrictions of its inhabitants which disfavour the acquirement of an Internet access. However, the Styria finding is certainly not enough evidence to state a regional gap in 2009.

The EU-SILC model in 2006 and the NASE model in 2009 both rejected the hypothesis that there is an impact of employment status on home Internet access hence nothing changed over three years.

5.3 Conclusions

In all computed models the sociodemographic characteristics turned out to be major determinants of home Internet access. Among these variables, age and education have the most explanatory power while household size, gender and income have minor impact on the depending variable.[43] The EU-SILC results question the existence of a gender and income gap but the model's substantial lower R^2 has to be kept in mind as well. Donat (2008) also examined the Austrian situation and could not find a gender gap.

The evidence for a spatial dimension of the digital divide is ambiguous. In 2006 regional influence on state level appears to be present due to several significant results in every model whereas the rural-urban impact is subject of serious doubts. Depending on which measure is used the corresponding ratios have a p-value of smaller than 0.05 in just four out of nine cases. It is important to notice that these significant odd-ratios are almost all for the mid-size towns and smaller than one which was totally unexpected. In only one EU-SILC model the ratios are larger than one which implies that people living in metropolitan areas are more likely to have home Internet access than their rural counterparts. A special source of concern about the robustness of these findings is that once the cut-off points are shifted the significance is gone in most cases. Independently of the question whether the yet mentioned results are sufficient or insufficient to state a spatial influence on the depending variable, the R^2 of the reduced model containing only sociodemographic characteristics strongly indicates a solely marginal influence of the spatial variables and casts serious doubts on the existence of a rural-urban divide in general.

In 2009 absolutely no evidence for a spatial influence is found. Neither the state variables, nor the urban-rural variables turn significant independent of their definition. However, one remarkable finding is that Styria has a significant ratio between 0.5 and 0.65 in all models which is hardly caused by chance in consideration of the fact that no other state turned nearly as often significant as Styria. Especially in 2009 when no other spatial variable had an impact on Internet access, Styria again had a p-value smaller than 0.05. A possible explanation is hard to find as there are no obvious reasons why Styria should have systematically lower Internet access rates in comparison to Vienna.

The hypothesis that costs are a determinant of Internet use in 2006 is confirmed in all models whereas a metropolitan neighbourhood has no impact on the depending variable. Job related variables in the EU-SILC and the 2009 NASE model had no explanatory power

[43] In its digital competitiveness report the European Commission (2009) stated: "[…] age and education are the two most important factors influencing Internet take-up."

hence the assumed relation between employment status and residential Internet use is rejected.

The findings of the basic model are consistent with reality in the sense that these days Internet is available nationwide and it is a question of awareness whether one has Internet or not. Due to a decrease in prices in recent years income is definitely not a crucial factor in the decision process as well as household size and gender. Household size actually could be considered to be a proxy for kids[44] hence its minor influence is caused by the presence of kids who need Internet access for school duties or social networks. Gender differences in Internet use are hard to observe in real life and results indicate that their impact is marginal in comparison to age and education. EU-SILC data even questions the NASE gender gap findings. As age and education are two major factors that influence people's awareness concerning the important role of ICTs in today's world, the findings are plausible and reasonable. Both are also important determinants of income and are likely to affect levels of digital skill.

The evidence concerning a possible spatial divide also makes sense. While in 2006 differences on regional level had some marginal influence on residential Internet access, this influence is gone these days. Results concerning the urban-rural divide are ambiguous. For 2006, depending whether one is in favour or not of a spatial gap, pros and cons can be found. From a scientific point of view especially the robustness checks and the high R^2 of the reduced sociodemographic model point towards the inexistence of such a gap and cast serious doubts on the few significant findings. However, there is absolutely no indication of whatever spatial influence on the depending variable three years later.

All in all the findings of the basic model are in line with real life experience. Home Internet use is mainly a function of sociodemographic characteristics with age and education as the major factors.

The last part of my thesis deals with the determinants of the Internet access type and their development over time as 2006 and 2009 NASE data is available and contains this information.

---

[44] An alternative explanation for a household size larger than two could be other relatives e.g. parents or a shared flat community but the absolute majority of all observations with a household size larger than two probably has children.

6. Advanced Model

6.1 Idea

In the last part of my thesis, I will examine the possible determinants of the Internet access type which is even more interesting than to just take a look at the plain residential Internet access. It has been argued that the spatial gap in Internet use shifted to a gap in Broadband – Narrowband use (Whitacre, 2005). The unavailability of Broadband access for larger parts of the population was called a “third digital divide” by Smolenski (2000). A brief look at the data reveals a sudden drop in Narrowband use over three years from 32.6 percent in 2006 to 5.9 percent in 2009 (RTR, 2009). DSL is available in almost all places in Austria and TA has a nationwide competitive price. The mobile Broadband penetration rate increased to 27 percent and is an affordable alternative for most of the customers especially in more rural areas due to a fast deployment of the technology and good coverage rates. Austria is a forerunner country within the EU and worldwide. Internet access via Cabel TV (CATV) is available in all major cities and close-by communities. Hence it can be argued that no one is depending on a Narrowband access anymore and everybody who wants to get a Broadband access can get one. Actually the absolute majority of all Austrians can choose between different types of Broadband access. This choice is the subject of interest in this chapter.

The aim is to create a model to find the relevant determinants that influence the people’s choice of their Internet access type. The model will be evaluated with the data from two different years to check for possible changes over time.

Literature about this topic is rare and no empirical evidence can be found. Most of the papers deal with the determinants of the demand for Broadband access in comparison to Narrowband but not with the choice of Internet access type. The sociodemographic characteristics already mentioned in this thesis play a significant role in the decision between Broadband and Narrowband access. Households with higher income, a college education and multiple PCs prefer Broadband access according to Savage & Waldman (2005). They also found indications that younger people are more likely to use Broadband than elderly persons. Rappoport et al. (2002) confirmed these suggestions with a multinomial logit model when reporting that income and education level are major determinants of Broadband access. Moreover the United States Government Accountability Office (GAO) suggested that consumers with high incomes and college degrees are significantly more likely to adopt Broadband than the average of society (GAO, 2006). Whitacre (2005) published similar results. But not all findings from the residential Internet access case can be transformed to the Broadband access analysis. A more critical paper from Flamm & Chaudhuri (2007) suggests that the effects of gender and metropolitan location seem to be different for low – and high-speed services whereas the spatial gap is once again confirmed. NTIA (2004) stresses another interesting fact that among Internet users, those with Broadband connections at home are more likely to be daily users than those with dial-up service.

All of these findings concern the adoption of Broadband and the influences on the choice between Narrow- and Broadband access. But no paper was yet published in which the decision between different kinds of Broadband types and Narrowband was examined. The NASE data from 2006 and 2009 offer the unique chance to do so. The idea is to create a model similar to the one of chapter 5 with all available sociodemographic characteristics, the spatial and the survey specific variables e.g. job position or costs.

First, I will compute descriptive statistics for each access type to get an impression of the data and what to expect from the regression results. Second, the model is presented and finally the results will follow.

6.2 Descriptive statistics

The drastic changes in access type rates in recent years are an excellent example how fast developments in ICTs occur. While almost every third household (32.7%) in Austria used dial-up access to connect to the Internet in 2006, a diminishing part of all users (5.9%) did so three years later. ADSL was and is the most common access type with currently 44.8 percent market share followed by Cable Internet (29.1%). Mobile Internet had an incredible increase of 600 percent in three years to a share of 27 percent. Narrowband is almost history these days although it was the second most common access type in 2006. See Figure 1 in chapter 4 for an illustration of the changes in access types.

The following selected descriptive statistics should give the reader an impression of what to expect from the regression results. For each access type the corresponding average of the household's characteristics is calculated[45].

Table 15 contains the average age for both years. While the differences in 2006 for all four access types are small with mobile Broadband as an exception, the average DSL[46] user in 2009 is significantly older than all alternatives. Mobile Broadband user would be expected to be younger than DSL or Narrowband user because usually this access type is used in connection with a Laptop. In both years this is the case. As Narrowband is the "old" technology in 2009 it is surprising that the average age in this category is not higher than 44.1 years.

---

[45] Please note, that multiple answers for the Internet access type question were possible. In case a person did so his or her characteristics were included in both categories.
[46] The DSL category includes ADSL, XDSL and DSL user.

Table 15: Average age for Internet access types

Access Type	*2006*	*2009*
DSL	45.3	51.8
Narrowband	45.5	44.1
Mobile Broadband	42.3	39.9
CATV	44	39.5

Source: RTR survey data

A multiple-persons household is more likely to have multiple PCs with an Internet connection thus a Broadband access is need to deal with the increasing stream of data. Hence the three Broadband types should have a larger household size on average than a Narrowband household. The results in Table 16 are not as expected. The major Broadband access type DSL is larger than its alternatives but in both years Narrowband ranks second. An explanation for the average household size of CATV could be that it is only available in larger cities with an urban population where the average family size is smaller than in the countryside. Overall the picture for this sociodemographic characteristic is not as clear as expected.

Table 16: Average household size for Internet access types

Access Type	*2006*	*2009*
DSL	3.1	3.7
Narrowband	3	3.5
Mobile Broadband	2.9	3.3
CATV	2.8	3.2

Source: RTR survey data

Gender differences in Internet use have been examined earlier with the basic model and evidence is mixed so far. Thus it is interesting to take a look at the access type rates by gender to see if a possible gap is present. In 2006 only moderate gender differences in Narrowband access are found while the rest is balanced. Three years later men represent the majority for all types but CATV. As mentioned in chapter 4 the household was selected randomly via certain criteria but not the person living in this household. Hence these results have to be interpreted with great caution.

Table 17: Share of male users for Internet access types

Access type	*2006*	*2009*
DSL	51	55
Narrowband	46	54
Mobile Broadband	49	54
CATV	50	49

Source: RTR survey data

The results for the education categories reflect the general development in Internet use. Rates are increasing for all access types but Narrowband. The lower rates for compulsory school in 2009 are due to the newly introduced apprenticeship category which absorbs lots of users from the compulsory group as seen in the descriptive statistics in chapter 4.1. The differences within the samples are in general small and difficult to interpret. The comparison over the years reveals a shift from Narrowband to mobile Broadband and no real increase in CATV penetration. DSL is the by far most popular access type in all education groups.

Table 18: Share of educational groups in percent for Internet access types

Level of Education	DSL		Narrowband		Mobile Broadband		CATV	
	2006	*2009*	*2006*	*2009*	*2006*	*2009*	*2006*	*2009*
Compulsory School	40	37	32	9	5	26	24	28
Apprenticeship	NA	46	NA	7	NA	25	NA	22
Some Years of College	37	52	38	4	5	22	21	22
College Degree	32	47	36	6	5	20	27	27
University Degree	36	45	29	4	3	23	32	28

Source: RTR survey data

The access rates for job position are presented in Table 19 and deliver a similar picture. DSL is the most popular access type for almost all job positions except for unemployed persons who have a complete different usage rate distribution compared to other job groups with CATV as the most common access type and mobile Internet in second place. The farmer/household category indicates a strong spatial gap in access types because DSL which is provided by the national incumbent TA and is available in almost all places in Austria is the by far leading access type with 65 percent in this category. CATV and mobile Internet have significant lower rates probably due to bad coverage rates in the countryside. The 2009 survey data does not even contain a single farmer observation with a CATV connection. 28.5 percent of the apprentice/maternity group use a mobile Internet access which could be explained with the age effect as young people tend to adopt new technologies like mobile Internet faster than older persons.

Table 19: Share of job position in percent for Internet access types

Job Position	DSL	Narrowband	Mobile Internet	CATV
Self-employed	50	5	24	20
Civil servant	52	2	26	20
Leading employee	47	3	23	27
Employee	49	8	24	21
Farmer/Household	65	14	17	10
Old age pensioner	46	10	18	25
Unemployed	21	5	35	39
Apprentice/Maternity	38	5	28.5	30

Source: RTR survey data

A focus of this thesis is on a possible spatial dimension of the digital divide. Hence the spatial variables are included in the model. Table 20 contains descriptive statistics for these geographic categories. In both years, small villages with a population less than 10,000 inhabitants are leading in the DSL category as CATV and mobile Internet usually is not available in these areas hence DSL or Narrowband is the only option. Every second household in a large city with a population above 50,000 inhabitants has a CATV access these days. As Cable Internet is only available in Austria’s major cities the gap in access types between the large city variable and the other two is huge. Mobil Internet is uniformly distributed over all categories while DSL in metropolitan areas is not as popular as in small and middle size towns.

Table 20: Share of communities in percent for Internet access rates

Number of inhabitants	DSL		Narrowband		Mobile Internet		CATV	
	2006	*2009*	*2006*	*2009*	*2006*	*2009*	*2006*	*2009*
10,000	43	56.7	41.2	9.7	4.1	23.6	11.7	10
10,000 – 50,000	39.7	53.9	36.1	4.5	5.5	23.8	18.7	17.8
50,000	28	25.6	21.4	3.3	4	20.6	46.6	50.5

Source: RTR survey data

The descriptive results did not reveal a clear pattern of what to expect from the advanced model. DSL users seem to be on average older whereas mobile Internet users tend to be younger. Due to small differences between the groups, no patterns could be found for the household size and the education groups. Only the 2009 results indicate significant differences in gender Internet use. Unemployed people have an unusual usage pattern whereas all other job positions had a usage pattern as expected. The farmer/household category and the urban-rural variables indicate a strong spatial gap which can be assumed as CATV and mobile Internet are not available in rural areas.

The next section deals with the theoretical background and the structure of the advanced model.

6.3 Theoretical framework and model description

A multinomial logit model is used to examine the influences of the sociodemographic and other variables on the choice of the Internet access type. As mentioned earlier, there has been research on the determinants of the choice between Broadband and Narrowband access but none on which type of Internet access the individual or household has chosen. In my model Narrowband access is just one possible access type among DSL, mobile Broadband and CATV. Having no Internet is an option as well. Hence I will use the most common model so far which was the multinomial logit model (see Whitacre (2005); Rappoport et al. (2002) and Donat (2008)).

The difference between the logit model and the multinomial logit model is that the dependent variable is not dichotomous but has *m* outcomes where *m* is larger than two. Let's assume *m* = 3 and the outcomes are "Narrowband", "Broadband" and "No Internet". I will use an unordered multinomial logit model which means that the outcomes are coded with 1, 2 and 3 but these values are arbitrary and 1 < 2 < 3 does not imply that 1 is inferior and 3 the best. As Austria is a forerunner country in the mobile Broadband deployment, I want to include this category in the analysis therefore it is simply not possible to order the outcomes because neither DSL nor mobile Broadband is strictly better than the other. The same is true for DSL and CATV. Hence the unordered version of the multinomial logit model is chosen.

The coefficients $\beta^{(1)}$, $\beta^{(2)}$ and $\beta^{(3)}$ are computed via maximum-likelihood and corresponding to each category. The probabilities for each outcome are

$$\Pr(y=1) = \frac{e^{X\beta^{(1)}}}{e^{X\beta^{(1)}} + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

$$\Pr(y=2) = \frac{e^{X\beta^{(2)}}}{e^{X\beta^{(1)}} + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

$$\Pr(y=3) = \frac{e^{X\beta^{(3)}}}{e^{X\beta^{(1)}} + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

There is more than one solution to $\beta^{(1)}$, $\beta^{(2)}$ and $\beta^{(3)}$ thus the model is unidentified in the sense that these solutions lead to the same probabilities for y = 1, y = 2 and y = 3. To identify the model, one coefficient is set arbitrary to 0. In case we set $\beta^{(1)} = 0$ the reference group is y = 1 and the coefficients $\beta^{(2)}$ and $\beta^{(3)}$ measure the relative change to this group. If the relative change from group 2 to group 3 is the subject of interest, it is necessary to set $\beta^{(2)} = 0$ and compute the other two coefficients. Although the coefficients differ depending on which β is set to 0, the predicted probabilities for y = 1, 2 and 3 would be still the same.

Setting $\beta^{(1)} = 0$, the equations are

$$\Pr(y=1) = \frac{1}{1 + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

$$\Pr(y=2) = \frac{e^{X\beta^{(2)}}}{1 + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

$$\Pr(y=3) = \frac{e^{X\beta^{(3)}}}{1 + e^{X\beta^{(2)}} + e^{X\beta^{(3)}}}$$

The relative probability of y = 2 to the base category is

$$\frac{\Pr(y=2)}{\Pr(y=1)} = e^{X\beta^{(2)}}$$

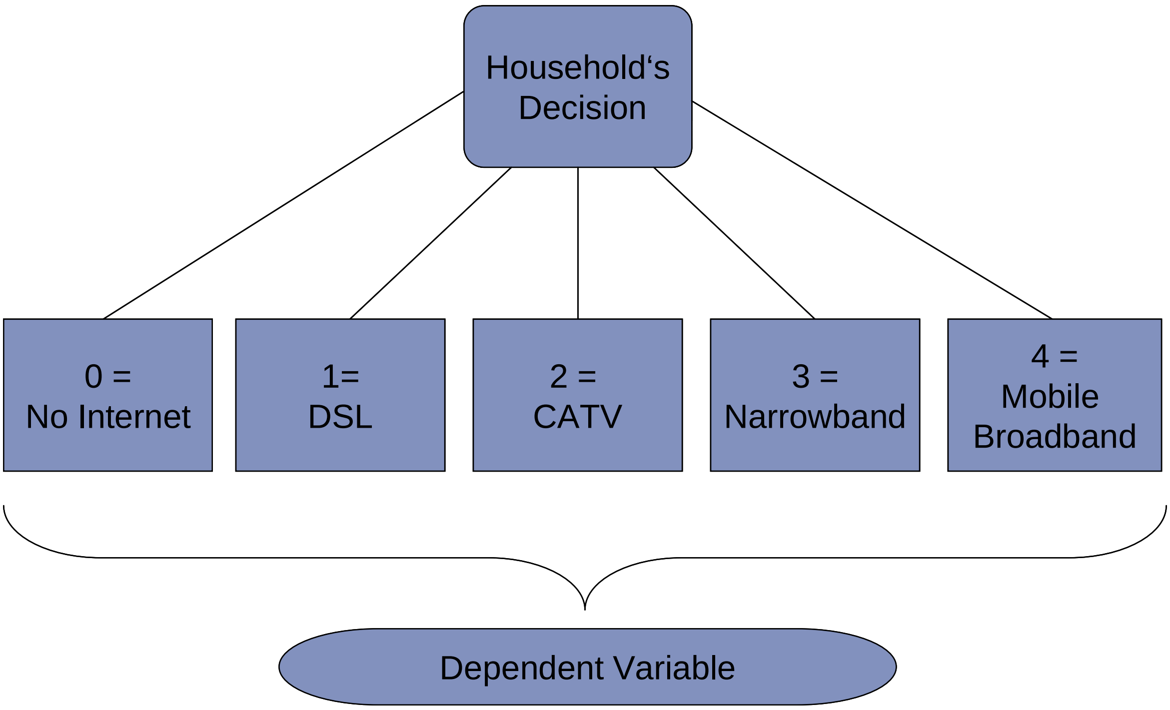
This ratio is called the relative risk and assume that X and $\beta_k^{(2)}$ are vectors equal to (x_1, x_2, …, x_k) and ($\beta_1^{(2)}$, $\beta_2^{(2)}$, …, $\beta_k^{(2)}$)'. Thus the ratio of the relative risk for a one-unit change in x_i is

$$\frac{e^{\beta_1^{(2)}x_1+...+\beta_i^{(2)}(x_i+1)+...+\beta_k^{(2)}x_k}}{e^{\beta_1^{(2)}x_1+...+\beta_i^{(2)}x_i+...+\beta_k^{(2)}x_k}} = e^{\beta_i^{(2)}}$$

The exponentiated value of a coefficient is the relative risk ratio for a one unit change in the corresponding variable. This risk has to been understood as the risk of the category relative to the base category.

Each respondent was asked which kind of Internet access he or she had in case residential Internet access existed. These answers form the dependent variable of the model as illustrated in Figure 6.

Figure 6: Dependent variable of the multinomial logit model



The models for both years differ a little bit due to differences in the available data in both years. In 2009 the job positions and the level of employment are available whereas this data was not collected in 2006. As I examine the household's decision between different types of access the availability of each access type has to be part of the model as well. Of course it is impossible for a household to choose CATV if it is not available in its community. In 2006 CATV operators covered about 50% of all Austrian households and no significant increase was observed the following years as all major cities are covered and markets are saturated. A CATV availability dummy variable is created with the help of the ZIP codes and added to the model. About 90 percent[47] of all households had the chance to get an ADSL connection in 2006 and today almost everyone can get ADSL. Hence only in 2006 an ADSL availability dummy is part of the model. Data for the mobile Broadband availability is not obtainable thus CATV will act as a proxy for mobile Broadband although their product characteristics are completely different. The reason for doing so is the assumption that mobile Broadband coverage was about the same in 2006 than for CATV – namely in all major Austrian cities. In 2009 mobile Broadband coverage is exceeding the CATV coverage by far therefore CATV is not considered to be a proxy for mobile Internet in the 2009 multinomial logit model.

The model for the choice of the Internet access type is

$$\Pr(y_i = j) = \frac{e^{\beta'_j x_i}}{\sum_{k=0}^{4} e^{\beta'_k x_i}} \text{ for j = 0,1,…,4 and i = 1,2,…,N}$$

The index *i* runs over all individuals *N* while *j* is for the five access types. *β* is the vector with all coefficients and *X* the matrix of all observed characteristics.

The reference group for the independent variables are the same as in chapter 5.1. All coefficients will measure the probabilities relative to “No Internet” which is the base category in both models. The effects of the other access types will be compared relatively to “No Internet” and between each other to find possible different effects for the same characteristics.

The *mlogit* command in *Stata* computes the coefficient *β* for every variable of each access type. As I would need to calculate $Pr(y_i = j)$ by hand, I use the *mfx* command to get the marginal effects for each access type which are easier to interpret and computed by *Stata*. The *mfx* command has to be run separately for each outcome e.g. Narrowband, DSL etc. after the *mlogit* procedure. Due to a larger number of independent variables and a complex computation method, the runtime of this procedure is quite long.

The *mfx* results report the marginal effect of the independent variable which is the rate of change in probability for the outcome *j* given a unit change of the variable *ceteris paribus*. Most of the model’s variables are dummies hence the coefficient refers to a discrete change of the dichotomous variable from 0 to 1. The sign, magnitude and significance of the change rate are subject of interest. The interpretation is not as straight forward as in the basic model. The outcome reference group is “No Internet” hence the results for the other four outcomes reflect the impact of the independent variables on the probability of individual *i*

---

[47] http://www.telekomaustria.com/presse/news/2006/0526-pk-neuestruktur.php (29th May, 2009)

to change its access type from “No Internet” to the corresponding type. As almost all influences are measured via several dummies these impacts are also relative to the dummies’ reference group. For example, the coefficient for “men” in the DSL category is the increase (or decrease - depending on the sign) in probability for being male instead of female to change from “No Internet” to DSL. In the following section all results will be interpreted in this manner.

EU-SILC data does not contain which Internet access type the respondent has chosen thus the advanced model is solely estimated with the NASE data. Only one model is computed for each year for two reasons: (1) In chapter 5, differences between the various models in both years turned out to be marginal and irrelevant as the major findings were the same for all models. (2) Furthermore it is more reasonable to discuss one model in detail and examine developments over time as these multinomial models are substantial larger than the already discussed logit models.

As the measure for a possible urban-rural divide, the degree of urbanity variables are chosen because they combine the population size and density approach and represent the official Eurostat definition. Results from chapter 5 indicate that the degree of urbanity definition classifies more accurate than its alternatives. Moreover it is a more intuitive approach.

Both models will contain sociodemographic, spatial and job-related variables but no product characteristics. This is a possible shortcoming of this analysis as these days consumers tend to demand product packages e.g. a triple play offer which includes Broadband, TV and fixed line telephone service instead of just a plain Internet access. This aspect is not modelled as well as the substantial differences between the four access types in performance, quality and costs. Depending on the results of my model, future research could include these product characteristics in its analysis.

In the next section the estimation results are presented and interpreted for both years.

6.4 Estimation results

6.4.1 2006 Multinomial Logit Results

All available sociodemographic, state, suburban and degree of urbanity variables as a measure for a possible spatial divide are part of the model. In addition dummies which indicate the availability of CATV and DSL are required because not every household in Austria had access to Cable or DSL Internet in 2006. It is easy to classify each observation when respondent's ZIP code is known.

No constant is included because *Stata* reported a problem when computing the marginal effects due to the constant. Results are presented in Table 29.

>Insert Table 29 about here<

Please note that especially the determinants of mobile Broadband would be of interest because Austria is one of the leading countries in mobile Broadband penetration but the 2006 NASE sample only contains 75 observations while the full sample size is 4,020. Unfortunately there are just not enough observations to find real significant variables that influence the probabilities of switching to mobile Broadband.

The model itself is highly significant which is not unexpected due to the large number of observations and explanatory variables. The reported pseudo R^2 is McFadden's goodness-of-fit measure and equals 0.377. The earlier found sociodemographic influence from the basic model appears to be present as well in the multinomial logit model. Household size has a positive but marginal influence on the probability to turn from offline to online. Its coefficient is significant for all four alternatives even though its magnitude is small in comparison to the other sociodemographic characteristics. The evidence for an already disputed gender gap is ambiguous again. For two out of four possible alternatives being male increases the chance of having residential Internet access relative to being female. The extent of the impact is limited. Interestingly the imputed income has a negative sign which means that it actually decreases the chances to have any form of Internet access although one would expect income to have a positive influence. However, the magnitude of its coefficient suggests that income is not a major determinant of the Internet access type. In contrary, education has a statistically significant impact on the choice probabilities of all types except for mobile Broadband. Its positive influence increases as the degree improves e.g. having a university degree has almost the largest absolute magnitude for each alternative. Keep in mind that compulsory school is the reference group hence the computed rates of change are relative to this group. An inter-alternative comparison reveals that education has the same influence on CATV than the three yet mentioned characteristics while for DSL and Narrowband the difference is substantial. The age impact is structured similar to the basic models. The differences between the age groups under 50 are insignificant which means they are inexistent. The three oldest groups decrease the online probability relative to the youngsters and the negative coefficient increases with age. The 70 plus group even reduces the probability of changing to mobile Broadband hence its impact has to be more than obvious otherwise a p-value of under 0.001 would not be possible.

The sociodemographic variables seem to be once again major determinates in the choice between different access types. Similar to the basic models, education and age have a greater impact than household size, income and gender. DSL and CATV have the most significant coefficients while almost no substantial impacts where found for mobile Broadband due to a small number of observations for this category. Interestingly the differences between the variables' coefficients for CATV and DSL are significantly. The sociodemographic influences on the probabilities for CATV are marginal in their magnitude in comparison with DSL. Narrowband appears to be distributed more equally within the groups as fewer significant differences are computed.

The spatial impact is examined in the advanced model as well. The degree of urbanity is used as the definition for the urban-rural variables. As usual, the results are mixed. There is no rural-urban impact for DSL whereas the full urbanized variable makes a change to CATV more likely relative to non-urbanized towns. Moderate urbanized communities reduce this probability. Please note that the model contains DSL and CATV availability variables hence this rural-urban influence is not a result of a better supply in metropolitan areas or unavailability in rural towns as the availability variable covers this effect. Living in a large city also makes a change from "No Internet" to Narrowband less likely in comparison to rural areas. Almost no evidence for a state impact on the outcome probabilities is found. Burgenland inhabitants are less likely to have DSL, CATV and Narrowband relative to Viennese which makes sense as this state has a low Internet penetration rate in general. Two other states have impact on the outcome: Upper Austria has a positive coefficient for CATV relative to Viennese and Vorarlberg two negative ones for DSL and mobile Internet. As there is no real mobile Internet availability variable just the CATV availability as proxy, the negative coefficient for mobile Broadband could point towards bad coverage with mobile services in the second smallest of all nine federal states. The suburb variable has no impact at all.

The spatial dimension appears to be almost inexistent. Only CATV is influenced a bit and Burgenland seems to be an outlier but the scattered significant coefficients are not enough to state a spatial gap in any form.

Cable and DSL availability have an impact on CATV and DSL as one would expect. Interestingly the possibility to get CATV makes a change from "No Internet" to DSL more likely but the coefficient is small compared to the other significant coefficients for DSL. Within the Cable category the availability influence is the most powerful influence by far and highly significant whereas the possibility to have DSL is just one significant explanatory variable among others for the DSL change probability. The likelihood of individual i to switch to CATV decreases a bit when DSL is available. Narrowband and mobile Internet are not influenced by these two variables according to the model.

The sociodemographic variables contribute most of the explanatory power to the model while the spatial influence is marginal. Age and education are the major determinants on the choice of Internet access type. The probability to get online decreases as age increases. Uneducated people are less frequent online than their educated counterparts. Household size, income and gender only have minor impact. However, the gender effect is questionable. Furthermore, the availability influences the probabilities for DSL and CATV. The inter-type differences are substantial. While the sociodemographic influence on DSL and Narrowband is large, CATV in general appears to be explained mainly by its availability as this coefficient exceeds the others by far. Hence the main motivation for people to switch from "No Internet" to Cable Internet is the chance to do so. Unfortunately there are not enough observations to get significant results for mobile Broadband.

The presentation of the 2009 results and a comparison of both models are following in the next section.

6.4.2 2009 Multinomial Logit Results

Two changes are made in the 2009 model compared to 2006: Job related variables are added and the DSL availability variable is gone because DSL is available nationwide these days and no employment data was collected in 2006. Table 30 contains the computed results.

>Insert Table 30 about here<

Please note that due to an increasing Broadband penetration over three years Narrowband now has the former mobile Broadband problem. Solely 105 respondents had Narrowband access hence once again it is difficult to get significant results for this access type as the sample size is 3,001. The number of observations for mobile Broadband has increased to 444.

The model itself is highly significant but its R^2 of 0.3 is worse than the 2006 score. The household size influence is marginal in its magnitude. Interestingly an increasing household size decreases the probability of having mobile Broadband. Hence singles or households without children tend to have mobile Internet which is intuitively because this access type unlikely satisfies the needs of a household with multiple PCs and a corresponding down- and upload stream not to mention technical restrictions. Gender and income appear to have no impact on the choice of access type. Only the DSL probability differs significantly between the sexes. Significant results for education are rare as well. CATV, Narrowband and mobile Broadband have not a single significant coefficient solely the switching probability for DSL is higher for better educated persons than for respondents with a compulsory school degree. All yet mentioned sociodemographic characteristics had minor impact in comparison with age. The reference group is formed by the 16 to 24 years old who have an over average penetration rate therefore all computed coefficients for the other age groups have a negative sign. The old pattern, that increasing age reduces the online probability respectively the probability to switch to one of the four access types, is confirmed. The 70 plus group is significantly less likely to have a DSL, CATV, Narrowband or mobile Broadband access than the youngsters. The inter-access comparison indicates that the magnitude of this effect is the same for all access types relative to the category's significant coefficients.

Sociodemographic influence in the 2009 multinomial model is not as dominant as it has been in all yet estimated models. Household size, gender, income and education have almost no influence on CATV, Narrowband and mobile Broadband. Only the probability to change from “No Internet” to DSL is influenced by those characteristics. Age contributes the major sociodemographic explanatory power to the model.

A large number of spatial variables are significant in comparison to other models. Although a full urbanized environment reduces the likelihood of having DSL and increases the chance to switch to CATV relative to a non-urbanized community, the evidence for a rural spatial impact is moderate. A moderate urbanized residence solely decreases the Narrowband switching probability hence only three out of eight rural-urban coefficients have a p-value of smaller than 0.05. State influence appears to be present especially for CATV because seven states have a significant negative impact relative to Vienna on the chance to switch from “No Internet” to Cable Internet. Upper Austria has a negative impact on all access types except for mobile Broadband. In addition Styria, Carinthia and Vorarlberg have a negative influence on Narrowband, DSL and mobile Broadband. Living in a suburb has only a marginal negative impact on the mobile Broadband switching probability.

Evidence for job related effects on the probabilities to change the access type is mixed. Mobile Internet and DSL appear to be unaffected by the employment status of the respondent while Narrowband and CATV exhibit significant results. Interestingly all coefficients are negative which means that these job groups are less likely to switch to from “No Internet” to CATV or Narrowband in comparison with self-employed persons. Farmers and respondents who keep the household are less likely to change their access type to CATV or Narrowband as well as civil servants. Being unemployed makes a change to DSL or Narrowband less probable. The level of employment has no effect on the depending variable. Finally, the availability of CATV has the expected negative sign for DSL and a positive one for Cable Internet.

The 2009 multinomial model delivered substantial different results compared to earlier estimated models. The sociodemographic influence was not predominant as household size, gender, income and education had only minor impact while age was the major explanatory variable. Results did not indicate a rural-urban urban gap for any access type but state influence for CATV and Narrowband appears to be present. Some job-related coefficients were significant but the overall evidence is only moderate.

The inter-access comparison reveals essential differences. DSL is mostly a function of the sociodemographic characteristics whereas the spatial and job related variables have almost no explanatory power. On the contrary, only age has impact on the switching probability of CATV while regional and job related variables are significant which could be motivated by a worse coverage in some states and in the countryside. The highly significant farmer/household variable is an indication for this assumption as CATV is only available in metropolitan areas thus farmers can not get such an access. A possible explanation is hard to find because availability issues should be covered with the availability variable. But not the whole effect is probably measured for two reasons: (1) ZIP codes do not match exactly the areas for which availability data is feasible and (2) the ZIP code problem mentioned in chapter 4.3 arrives thus it could be possible that some of the significant state coefficients are due to a worse CATV coverage than Vienna. An alternative explanation could be that maybe more consumers in Vienna are aware of the fact that CATV is an alternative to DSL while on the countryside there is a lack of information. However, the 2009 results indicate regional differences. Evidence for mobile Broadband is sparsely again. Increasing household size and age reduce the likelihood for a change to mobile Internet. Vorarlberg and suburbs have a negative coefficient which could point towards availability problems as the model does not include a mobile Internet availability variable.

Comparison with the 2006 multinomial logit model

The sociodemographic influence seems to decrease over three years. The minor impact of household size, gender and income disappears and the education effect diminishes. Hence no gender gap appears to be present. Age is the only persistent major explanatory sociodemographic variable for all categories. The marginal spatial effect from 2006 is reinforced as quite a few states have a p-value smaller than 0.05 in 2009. It is questionable if these findings indicate real regional differences or just measure availability issues as the CATV availability variable is not 100 percent accurate. Neither the 2006 nor the 2009 model deliver enough evidence to report a rural-urban influence on the depending variable. An over the year comparison for the job-related influence is not possible as the 2006 NASE data does not include this information.

The inter-access comparison indicates that the sociodemographic influence on DSL remained over the years while it diminished for Narrowband. In 2006, CATV mainly depended on its availability whereas it is a function of age, spatial and job related variables these days. Although a variable in both models covers availability issues, especially the significant state coefficients could point towards structural obstacles to get CATV. Unfortunately, evidence for the determinants of the probability to switch to mobile Broadband is sparsely in both years.

6.5 Conclusions

The interpretation of the multinomial logit results is not as straight forward as it was for the basic model. The main finding is that the sociodemographic influence on access types is diminishing over the years, which is consistent with our real life experience. These days the Internet is an essential tool in our everyday life and the choice of access is independent of income, gender or education. The revealed age effect on the choice was also found in the basic model. The negative impact of increasing age on the probability to switch from “No Internet” to any Internet access type is caused by a substantial lower average Internet access rate of elderly persons rather than on certain access type preferences. Both models suggested that sociodemographic determinants play a minor role in comparison to the basic model. The spatial findings seem to be driven by availability issues rather than by real regional differences and the inexistence of a rural-urban influence is also confirmed by our day-to-day experience.

The switching probability of DSL and Narrowband seem to be more affected by sociodemographic influences than the rest. A change to Broadband appears to be basically a function of its availability and spatial influences in both models. Unfortunately, the evidence for the determinants of mobile Broadband is sparsely. The 2009 results suggest that smaller households and younger consumers prefer mobile access which makes sense. These access-related findings are hard to check for their plausibility because no papers concerning the determinants of this choice have been published and in real life this decision process is usually driven by other factors e.g. product characteristics, personal preferences.

All in all, the general picture created by the multinomial logit results is not as clear as with the basic model because the advanced analysis includes more factors and unicausal explanatory structures are inappropriate. A model with additional variables containing product characteristics should be used in future research to get an in-depth analysis of the determinants.

7. Conclusions

Internet access in every household is an essential thing for Austria's way from an industrial to a knowledge based society. Hence it is important to have an insight about the main determinants of residential Internet access. The aim of this thesis was to examine which factors influence the household's decision whether and which Internet access is acquired. The existence of various dimensions of the digital divide was tested with the help of statistical models and survey data from two different years to observe possible developments. A focus of this thesis is dedicated to the so called spatial gap.

Descriptive statistics of the available data from RTR and EU-SILC surveys indicate substantial differences in residential Internet access rates by gender, income, household size, education and age but only moderate evidence for a spatial gap. Various logit models for each survey sample are computed with different urban-rural definitions to check for robustness. In both years the sociodemographic characteristics turned out to be the major determinants of Internet use and among these education and age delivered the most explanatory power to the model while household size, gender and income only had minor impact. The presence of a gender and income gap is questionable as there are no significant findings with the EU-SILC data. Sociodemographic influence tends to diminish over the years.

Evidence for an urban-rural gap in residential Internet access is rare in 2006. Regional influence on state level appears to be present while the metropolitan and mid-size town variables are seldom significant independently of the used urban-rural measure. Three years later, absolutely no spatial impact on the dependent variable could be found. Internet service costs are negatively correlated with home Internet access whereas the respondent's job position and a suburban environment had no effect on the penetration rates at all.

These findings are in line with real life experience and backed by recent studies which also report that age and education are the main sociodemographic determinants (European Commission, 2009) and question the gender gap (Donat, 2008). Household's residence appears not to have impact on its decision to acquire home Internet access hence there is no rural penalty which disadvantage Internet diffusion in the countryside.

The choice of Internet access type was examined via a multinomial logit model. Sociodemographic influence plays a minor role compared to the basic model as only age had significant and substantial effects on all access types. Spatial variables did explain some of the variation of the depending variable but this effect seems to be driven by availability issues rather than by real spatial differences although availability variables were part of the model. In 2006, DSL and Narrowband are mainly influenced by sociodemographic characteristics while the choice for CATV was motivated by its availability. Three years later, CATV was a function of spatial and job related variables which is again an indication for availability issues. Sociodemographic influence diminished in general. Unfortunately there are not enough mobile Broadband observations to get significant results for this category and to observe possible developments of its determinants. A model containing product characteristics is recommended for future research on this topic to get an in-depth analysis.

8. References

Agarwal, Ritu; Animesh, Animesh & Prasad, Kislaya (2005): "Social Interactions and the 'Digital Divide': Explaining Regional Variations in Internet Use." in Robert H. Smith School Research Paper No. RHS 06-024. Available at SSRN: http://ssrn.com/abstract=796090 (19th August, 2009)

Angrist, Joshua & Lang, Kevin (2002): "How Important are Classroom Peer Effects? Evidence from Boston's Metco Program" in NBER Working Paper 9263

Beilock, Richard & Dimitrova, Daniela (2003): "An exploratory model of inter-country Internet diffusion" in Telecommunications Policy 27(3-4): 237 – 252

Bimber, Bruce (2000): "Measuring the Gender Gap on the Internet" in Social Science Quarterly 81(3)

Brandtweiner, Roman & Donat, Elisabeth (2007): "The Digital Divide – Any Reasons for Enthusiasm? The Case of Austria" from http://ecom.fov.uni-mb.si/proceedings.nsf/0/4cb4c828f0887f21c12572ee0075b93c/$FILE/Paper42.pdf (9th July, 2009)

Chaudhuri, Anindya; Flamm, Kenneth & Horrigan, John (2005): "An Analysis of the Determinants of Internet Access" in Telecommunications Policy 29(9-10): 731 – 755

Chen, Wenhong & Wellman, Barry (2003): "Charting and Bridging Digital Divides: Comparing Socio-economic, Gender, Life Stage, and Rural Urban Internet Access and Use in Eight Countries" from http://www.amd.com/us-en/assets/content_type/DownloadableAssets/FINAL_REPORT_CHARTING_DIGI_DIVIDES.pdf (28th May, 2009)

Chen, Wenhong & Wellman, Barry (2004): "The Global Digital Divide – Within and Between Countries" in IT&Society, Volume 1, Issue 7: 39-45

Donat, Elisabeth (2008): "Determinants of Internet Usage. An In-Depth Analysis of the Digital Divide" in Verlag Kovac, Hamburg

Drabenstott, Mark (2001): "New Policies for a New Rural America." International Regional Science Review 24(1):3-15

Edwards, Paul (1995): "From 'Impact' to Social Process." in Handbook of Science and Technology Studies: 257 – 285

European Commission (2008): "14th Report on the Implementation of the Telecommunications Regulatory Package" from http://ec.europa.eu/information_society/policy/ecomm/library/communications_reports/annualreports/14th/index_en.htm (6th July, 2009)

European Commission (2009): "Europe's Digital Competitiveness Report - Volume 1: i2010 — Annual Information Society Report 2009" from http://ec.europa.eu/information_society/eeurope/i2010/docs/annual_report/2009/sec_2009_1060_vol_1.pdf (11th August, 2009)

Flamm, Kenneth & Chaudhuri, Anindya (2007): "An analysis of the determinants of Broadband access" in Telecommunications Policy: 31(6-7): 312 – 326

Fong, Eric; Wellman, Barry; Kew, Melissa & Wilkes, Rima (2001): "Correlates of the Digital Divide: Individual, Household and Spatial Variation" from http://www.chass.utoronto.ca/~wellman/publications/digitaldivide/digital-report2d.PDF (28th May, 2009)

Forestier, Grace & Kenney (2002): "Can information and communication technologies be pro-poor?" in Telecommunications Policy 26(11): 623-646

Future Foundation (2004): "The Digital Divide in 2025" from http://www.btplc.com/societyandenvironment/pdf/digitaldivide2025.pdf (28th May, 2009)

General Accounting Office (2001): "Characteristics and Choices of Internet Users" from http://www.gao.gov/new.items/d01345.pdf (28th May, 2009)

General Accounting Office (2006): "Telecommunications: Broadband Deployment Is Extensive throughout the United States, but It Is Difficult to Assess the Extent of

Deployment Gaps in Rural Areas" from http://www.gao.gov/new.items/d06426.pdf (21st July, 2009)

Goolsbee, Austan & Klenow, Peter (2002): "Evidence on Learning and Network Externalities in the Diffusion of Home Computers" in The Journal of Law and Economics 45: 317 – 343

Grubesic, Tony (2002): "Spatial dimensions of Internet activity" in Telecommunications Policy 26(7-8): 363-387

Hindman, Douglas (2000): "The rural-urban digital divide" in Journalism and Mass Communication Quarterly 77(3): 549 - 560

Kraut, Robert; Scherlis, William; Mukhopadhyay, Tridas & Manning, Jane (1996): "The HomeNet Field Trial of Residential Internet Services." in Communications of the ACM 39(12): 55 – 64

Krings, Bettina & Riehm, Ulrich (2006): "Internet for all? The Discussion on the "digital divide" revisited." in Rehberg, Karl-Siegbert (Hg.): Soziale Ungleichheit, kulturelle Unterschiede (2006): pp. 3052-3061.

Lenhart, Amanda et al. (2003): "The Ever-Shifting Internet Population – A new look at Internet access and the digital divide" from The Pew Internet & American Life Project, Washington DC

Madden, Gary & Coble-Neal, Grant (2003): "Internet use in rural and remote Western Australia" in Telecommunications Policy 27(3-4): 253 – 266

Mahler, Alwin & Rogers, Everett (1999): "The Diffusion of Interactive Communication Innovations and the Critical Mass: the Adoption of Telecommunications Services by German Banks." in Telecommunications Policy 23(10-11): 719-740

Malecki, Ed (2003): "Digital Development in rural areas: potentials and pitfalls." in Journal of Rural Studies 19: 201-214

Mills, Bradford & Whitacre, Brian (2003): "Understanding the Non-Metropolitan-Metropolitan Digital Divide" in Growth and Change 34(2): 219 – 243

Niles, Sarah & Hanson, Susan (2003): "A New Era of Accessibility." in Urban and Regional Information Systems Association Journal 15(1)

NTIA (1999): „Falling through the net: Defining the Digital Divide" from http://www.ntia.doc.gov/ntiahome/fttn99/ (28th May, 2009)

NTIA (2000): "Falling through the net: Towards digital inclusion" from http://search.ntia.doc.gov/pdf/fttn00.pdf (28th May, 2009)

NTIA (2002): "A Nation Online: How Americans are expanding their Use of the Internet" from http://www.ntia.doc.gov/ntiahome/dn/anationonline2.pdf (28th May, 2009)

NTIA (2004): "A Nation Online: Entering the Broadband Age" from http://www.ntia.doc.gov/reports/anol/NationOnlineBroadband04.pdf (28th May, 2009)

OECD (2001): "Understanding the Digital Divide" from http://www.oecd.org/dataoecd/38/57/1888451.pdf (28th May, 2009)

Ono, Hiroshi & Zavodny, Madeline (2002): "Gender and the Internet" from http://swopec.hhs.se/hastef/papers/hastef0495.pdf (8th June, 2009)

Rappoport, Paul; Kridel, Donald; Taylor, Lester & Alleman, James (2002): "Residential Demand for Access to the Internet" from http://www.colorado.edu/engineering/alleman/print_files/Forecasting_the_Demand_for_Internet_Services.PDF (21st July, 2009)

RTR (2006): "Abgrenzung des Marktes für breitbandigen Zugang auf Vorleistungsebene" from http://www.rtr.at/de/komp/KonsultationBBMarkt2007/Untersuchung_Breitbandmarkt.pdf (19th June, 2009)

RTR (2009): "Der österreichische Breitbandmarkt aus Sicht der Nachfrager im Jahre 2009" from http://www.rtr.at/de/komp/BerichtNASE2009 (29th May, 2009)

Savage, Scott & Waldman, Donald (2005): "Broadband Internet access, awareness, and use: Analysis of United States household data" in Telecommunications Policy: 29(8): 615 – 633

Selhofer, Hannes & Hüsing, Tobias (2002): “The digital divide index – a measure of social inequalities in the adoption of ICT” from http://www.sibis-eu.org/files/Huesing_Selhofer_DDIX_2002.pdf (28th May, 2009)

Selhofer, Hannes & Hüsing, Tobias (2004): “DIDIX: A Digital Divide Index for measuring Inequality in IT Diffusion” in IT&Society, Volume 1, Issue 7: 21 - 38

Shashaani, Lily (1997): “Gender Differences in Computer Attitudes and Use among College Students” in Journal of Educational Computing Research 16(1): 37 - 51

Smolenski, Mark (2000): “The Digital Divide and American Society: A Report on the Digital Divide and its Social and Economic Implications for Our Nation and Its Citizens” from Gartner Group, October.

UCLA Center for Communication Policy (2003): “The UCLA Internet Report: Surveying the Digital Future Year Three” from http://www.digitalcenter.org/pdf/InternetReportYearThree.pdf (9th June, 2009)

Van Dijk, Jan (2005): “The Deepening Divide: Inequality in the Information Society“ Sage Publications, Thousand Oaks, California

Vlosky, Richard & Poku, Kofi (2002): “Bridging the Rural/Urban Digital Divide” from http://www.lfpdc.lsu.edu/publications/working_papers/wp53.pdf (29th May, 2009)

Warren, Martyn (2007): “The digital vicious cycle: Links between social disadvantage and digital exclusion in rural areas” in Telecommunications Policy 31(6-7): 374 – 388

Wellman, Barry; Haase, Anabel; Witte, James & Hampton, Keith (2001): “Does the Internet Increase, Decrease, or Supplement Social Capital? Social Networks, Participation, and Community Commitment.“ in American Behavioral Scientist 45(3): 437 – 456

Whitacre, Brian E. (2005): “Bridging the Rural – Urban Digital Divide in Residential Internet Access” from http://scholar.lib.vt.edu/theses/available/etd-09262005-123744/unrestricted/Whitacre.pdf (2nd June, 2009)

Willis, Sue & Tranter, Bruce (2002): “Beyond the Digital Divide: Socio-economic Dimensions of Internet Diffusion in Australia” in Conference on the Digital Divide: Technology and Politics in the Information Age, 2002

Wunnava, Phanindra & Leiter, Daniel (2009): “Determinants of Intercountry Internet Diffusion Rates“ in American Journal of Economics and Sociology 68(2)

9. Appendix

Table 21: OLS estimation results

Variables	2006		2009	
	Coef.	Se	Coef.	Se
Age	0.040***	(0.00)	4.326	(6.81)
Age squared	-0.000***	(0.00)	-0.010	(0.07)
Compulsory School	Reference group			
Apprenticeship			182.093**	(61.08)
Some Years of College	0.163***	(0.03)	367.919***	(69.63)
College Degree	0.303***	(0.03)	528.894***	(65.83)
University Degree	0.508***	(0.03)	802.856***	(71.00)
Women	Reference group			
Men	0.164***	(0.02)	160.476***	(38.75)
Household size	0.126***	(0.01)	245.875***	(13.91)
Marginal Employed	Reference group			
Full time			337.846***	(87.36)
Part time			50.270	(93.41)
Self-employed	Reference group			
Civilservant			44.163	(115.06)
Leading Employee			221.821*	(111.86)
Retired			-238.409**	(90.29)
Unemployed			-641.383***	(153.19)
Employee			-156.506	(100.40)
Farmer/Household			-442.286***	(105.30)
Apprentice/Maternity			-214.514*	(97.13)
Constant	6.177***	(0.10)	825.328***	(179.94)
R^2	0.4615		0.2943	
Adjusted R^2	0.4594		0.2891	
F-Value	211.11		56.62	
p-Value	0.0000		0.0000	
Number of Observations	1,732		2,326	

Source: RTR survey data

Note: "***" significance at 0.1%; "**" significance at 1%; "*" significance at 5%; *Huber-White standard errors* in parentheses and "R^2" refers to McFadden's. "F-value" is the F-statistic of the model and "p-value" the corresponding observed probability.

Table 22: Estimation results for the 2006 logit model with different population size variables

Variables	(1)		(2)		(3)		(4)	
	Coef.	Se	Coef.	Se	Coef.	Se	Coef.	Se
Household size	1.230***	(0.05)	1.230***	(0.05)	1.233***	(0.05)	1.229***	(0.05)
Imputed Log Income	3.260***	(0.57)	3.252***	(0.57)	3.251***	(0.57)	3.260***	(0.57)
Women	Reference group							
Men	1.336**	(0.13)	1.337**	(0.13)	1.337**	(0.13)	1.334**	(0.13)
Compulsory School	Reference group							
Some Years of College	1.924***	(0.23)	1.930***	(0.23)	1.926***	(0.23)	1.920***	(0.23)
College Degree	3.108***	(0.41)	3.107***	(0.41)	3.092***	(0.41)	3.108***	(0.41)
University Degree	6.591***	(1.32)	6.631***	(1.33)	6.569***	(1.32)	6.534***	(1.31)
Age 16 – 24	Reference group							
Age 25 – 29	0.771	(0.23)	0.766	(0.23)	0.764	(0.23)	0.767	(0.23)
Age 30 – 39	0.909	(0.23)	0.908	(0.24)	0.901	(0.23)	0.908	(0.23)
Age 40 – 49	0.779	(0.19)	0.772	(0.19)	0.774	(0.19)	0.778	(0.19)
Age 50 – 59	0.459**	(0.11)	0.456**	(0.11)	0.457**	(0.11)	0.458**	(0.11)
Age 60 – 69	0.266***	(0.07)	0.264***	(0.07)	0.265***	(0.07)	0.266***	(0.07)
Age 70 plus	0.063***	(0.02)	0.062***	(0.02)	0.062***	(0.02)	0.063***	(0.02)
Under 10,000 inhabitants	Reference group							
Between 10,000 & 50,000 inhabitants	0.936	(0.12)						
More than 50,000 inhabitants	0.594*	(0.15)	0.518*	(0.16)	0.665	(0.16)		
Under 5,000 inhabitants			Reference group					
Between 5,000 & 50,000 inhabitants			0.828	(0.16)				
Under 15,000 inhabitants					Reference group			
Between 15,000 and 50,000 inhabitants					1.062	(0.14)		
Under 10,000 inhabitants							Reference group	
Between 10,000 & 40,000 inhabitants							0.927	(0.12)
More than 40,000 inhabitants							0.638	(0.16)
Number of inhabitants	1.000	(0.00)	1.000	(0.00)	1.000	(0.00)	1.000	(0.00)
Vienna	Reference group							
Lower Austria	0.944	(0.19)	0.942	(0.19)	0.942	(0.19)	0.969	(0.20)
Upper Austria	1.589*	(0.30)	1.593*	(0.30)	1.561*	(0.29)	1.628**	(0.31)
Salzburg	0.660	(0.14)	0.657	(0.14)	0.669	(0.15)	0.677	(0.15)
Burgenland	0.328***	(0.10)	0.323***	(0.10)	0.342***	(0.10)	0.336***	(0.10)
Styria	0.643*	(0.12)	0.645*	(0.12)	0.631*	(0.12)	0.662*	(0.13)
Carinthia	0.735	(0.17)	0.739	(0.18)	0.744	(0.18)	0.765	(0.18)
Tyrol	0.732	(0.17)	0.728	(0.16)	0.737	(0.17)	0.749	(0.17)
Vorarlberg	0.568*	(0.15)	0.562*	(0.15)	0.560*	(0.15)	0.581*	(0.16)
Suburbs	0.925	(0.14)	0.918	(0.14)	0.892	(0.14)	0.938	(0.15)
Alternative ISP available	2.681***	(0.35)	2.681***	(0.34)	2.598***	(0.33)	2.686***	(0.35)
Constant	0.000***	(0.00)	0.000***	(0.00)	0.000***	(0.00)	0.000***	(0.00)
Pseudo R^2	0.354		0.354		0.354		0.353	
Number of Observations	3,719		3,719		3,719		3,719	

Source: RTR survey data

Note: "***" significance at 0.1%; "**" significance at 1%; "*" significance at 5%; *Huber-White standard errors* in parentheses and "Pseudo R^2" refers to McFadden's pseudo R-square. Model (1): With fewer than 10,000 inhabitants as reference group and 10,000 – 50,000 & over 50,000 inhabitants variables. Model (2): With fewer than 5,000 inhabitants as reference group and 5,000 – 50,000 & over 50,000 inhabitants variables. Model (3): With fewer than 15,000 inhabitants as reference group and 15,000 – 50,000 & over 50,000 inhabitants variables. Model (4): With fewer than 10,000 inhabitants as reference group and 10,000 – 40,000 & over 40,000 inhabitants variables.

Table 23: Estimation results for the 2006 logit model with different population density variables

Variables	(1)		(2)		(3)	
	Coef.	Se	Coef.	Se	Coef.	Se
Household size	1.227***	(0.05)	1.229***	(0.05)	1.226***	(0.05)
Imputed Log Income	3.259***	(0.57)	3.263***	(0.57)	3.268***	(0.58)
Women	Reference group					
Men	1.338**	(0.13)	1.337**	(0.13)	1.341**	(0.13)
Compulsory School	Reference group					
Some Years of College	1.928***	(0.23)	1.927***	(0.23)	1.924***	(0.23)
College Degree	3.087***	(0.41)	3.100***	(0.41)	3.118***	(0.41)
University Degree	6.475***	(1.30)	6.484***	(1.30)	6.524***	(1.30)
Age 16 – 24	Reference group					
Age 25 – 29	0.759	(0.23)	0.761	(0.23)	0.764	(0.23)
Age 30 – 39	0.901	(0.23)	0.909	(0.23)	0.911	(0.23)
Age 40 – 49	0.771	(0.19)	0.776	(0.19)	0.778	(0.19)
Age 50 – 59	0.456**	(0.11)	0.459**	(0.11)	0.461**	(0.12)
Age 60 – 69	0.263***	(0.07)	0.267***	(0.07)	0.267***	(0.07)
Age 70 plus	0.062***	(0.02)	0.063***	(0.02)	0.063***	(0.02)
Pop Density under 100	Reference group					
Pop Density between 100 & 800	0.964	(0.11)				
Pop Density more than 800	0.721	(0.16)				
Pop Density under 100			Reference group			
Pop Density between 100 & 500			0.981	(0.12)		
Pop Density more than 500			0.769	(0.15)	0.673	(0.15)
Pop Density under 50					Reference group	
Pop Density between 50 & 500					0.854	(0.13)
Number of inhabitants	1.000	(0.00)	1.000	(0.00)	1.000	(0.00)
Vienna	Reference group					
Lower Austria	0.958	(0.20)	1.027	(0.21)	1.029	(0.21)
Upper Austria	1.527*	(0.30)	1.637**	(0.31)	1.681**	(0.32)
Salzburg	0.686	(0.15)	0.716	(0.15)	0.704	(0.15)
Burgenland	0.338***	(0.10)	0.356***	(0.10)	0.363***	(0.11)
Styria	0.631*	(0.13)	0.669*	(0.13)	0.672*	(0.13)
Carinthia	0.741	(0.18)	0.787	(0.19)	0.775	(0.18)
Tyrol	0.739	(0.17)	0.773	(0.17)	0.764	(0.17)
Vorarlberg	0.550*	(0.15)	0.608	(0.16)	0.619	(0.16)
Suburbs	0.919	(0.14)	0.922	(0.14)	0.949	(0.15)
Alternative ISP available	2.599***	(0.33)	2.629***	(0.34)	2.666***	(0.34)
Constant	0.000***	(0.00)	0.000***	(0.00)	0.000***	(0.00)
Pseudo R^2	0.353		0.353		0.353	
Number of Observations	3,719		3,719		3,719	

Source: RTR survey data

Note: "***" significance at 0.1%; "**" significance at 1%; "*" significance at 5%; *Huber-White standard errors* in parentheses and "Pseudo R^2" refers to McFadden's pseudo R-square. Model (1): With a population density of less than 100 inhabitants/km^2 as reference group and 100 - 800 & over 800 inhabitants/km^2 variables. Model (2): With a population density of less than 100 inhabitants/km^2 as reference group and 100 – 500 & over 500 inhabitants/km^2 variables. Model (3): With a population density of less than 50 inhabitants/km^2 as reference group and 50 – 500 & over 500 inhabitants/km^2 variables.

Table 24: Estimation results for the 2006 logit model with degree of urbanity variables

Variables	(1)	
	Coef.	Se
Household size	1.230***	(0.05)
Imputed Log Income	3.253***	(0.57)
Women	Reference group	
Men	1.339**	(0.13)
Compulsory School	Reference group	
Some Years of College	1.928***	(0.23)
College Degree	3.112***	(0.41)
University Degree	6.634***	(1.34)
Age 16 – 24	Reference group	
Age 25 – 29	0.767	(0.23)
Age 30 – 39	0.900	(0.23)
Age 40 – 49	0.773	(0.19)
Age 50 – 59	0.455**	(0.11)
Age 60 – 69	0.264***	(0.07)
Age 70 plus	0.062***	(0.02)
Non-Urbanized	Reference group	
Full Urbanized	0.603*	(0.14)
Moderate Urbanized	0.840	(0.43)
Number of inhabitants	1.000	(0.00)
Vienna	Reference group	
Lower Austria	0.925	(0.19)
Upper Austria	1.538*	(0.29)
Salzburg	0.651*	(0.14)
Burgenland	0.330***	(0.10)
Styria	0.626*	(0.12)
Carinthia	0.730	(0.17)
Tyrol	0.722	(0.16)
Vorarlberg	0.527*	(0.15)
Suburbs	0.890	(0.14)
Alternative ISP available	2.637***	(0.33)
Constant	0.000***	(0.00)
Pseudo R^2	0.354	
Number of Observations	3,719	

Source: RTR survey data

Note: "***" significance at 0.1%; "**" significance at 1%; "*" significance at 5%; *Huber-White standard errors* in parentheses and "Pseudo R^2" refers to McFadden's pseudo R-square.

Table 25: Estimation results for the EU-SILC logit model with degree of urbanity and population size variables

Variables	(1)		(2)	
	Coef.	Se	Coef.	Se
Household size	1.482***	(0.04)	1.515***	(0.05)
Women	Reference group			
Men	0.956	(0.06)	0.954	(0.06)
Age 16 – 24	Reference group			
Age 25 – 29	0.647**	(0.10)	0.650**	(0.10)
Age 30 – 39	0.642***	(0.08)	0.641***	(0.08)
Age 40 – 49	0.867	(0.10)	0.877	(0.11)
Age 50 – 59	0.860	(0.11)	0.872	(0.11)
Age 60 – 69	0.391***	(0.05)	0.393***	(0.05)
Age 70 plus	0.162***	(0.02)	0.166***	(0.03)
Vienna	Reference group			
Lower Austria	0.640**	(0.11)	0.765	(0.11)
Upper Austria	0.856	(0.13)	0.983	(0.13)
Salzburg	0.745	(0.14)	0.863	(0.15)
Burgenland	0.798	(0.17)	1.037	(0.21)
Styria	0.447***	(0.07)	0.544***	(0.08)
Carinthia	0.626*	(0.12)	0.764	(0.13)
Tyrol	0.688*	(0.12)	0.850	(0.13)
Vorarlberg	1.759**	(0.38)	1.870***	(0.34)
Compulsory School	Reference group			
No compulsory School	0.250**	(0.11)	0.249**	(0.11)
Apprenticeship	1.732***	(0.15)	1.729***	(0.15)
Some Years of College	3.164***	(0.38)	3.095***	(0.38)
College Degree	5.409***	(0.60)	5.230***	(0.58)
University Degree	9.651***	(1.49)	9.216***	(1.42)
Income	1.000	(0.00)	1.000	(0.00)
Household with kids	1.081	(0.08)	1.069	(0.08)
Self-employed	Reference group			
Employee	0.705**	(0.10)	0.697**	(0.09)
Unemployed	1.117	(0.68)	1.074	(0.67)
Keeping Household	1.512	(0.86)	1.513	(0.88)
Retired	1.857	(1.09)	1.850	(1.10)
Student/Pupil	1.696	(1.01)	1.709	(1.03)
Maternity Leave	2.416	(1.54)	2.436	(1.57)
For other reasons non-employed	1.320	(0.87)	1.299	(0.86)
Non-employed	Reference group			
Part Time	2.685	(1.55)	2.676	(1.57)
Full Time	2.464	(1.40)	2.492	(1.44)
Less than 10,000 inhabitants	Reference group			
Between 10,000 & 100,000 inhabitants	0.889	(0.08)		
More than 100,000 inhabitants	0.878	(0.11)		
Non-Urbanized			Reference group	
Full Urbanized			1.257*	(0.13)
Moderate Urbanized			1.322**	(0.11)
Constant	0.208*	(0.13)	0.140**	(0.09)
Pseudo R^2	0.202		0.204	
Number of Observations	5,997		5,997	

Source: EU-SILC survey data

Table 25 (cont'd):

Note: "***" significance at 0.1%; "**" significance at 1%; "*" significance at 5%; *Huber-White standard errors* in parentheses and "Pseudo R^2" refers to McFadden's pseudo R-square. Model (1): With fewer than 10,000 inhabitants as reference group and 10,000 – 100,000 & more than 100,000 inhabitants variables. Model (2): With non-urbanized communities as reference group and moderate & full urbanized cities variables.

Table 26: Estimation results for the 2009 logit model with different population size variables

Variables	(1)		(2)		(3)		(4)	
	Coef.	Se	Coef.	Se	Coef.	Se	Coef.	Se
Household size	1.465***	(0.08)	1.465***	(0.08)	1.474***	(0.08)	1.466***	(0.08)
Imputed Income	1.001***	(0.00)	1.001***	(0.00)	1.001***	(0.00)	1.001***	(0.00)
Women	Reference group							
Men	1.507***	(0.18)	1.507***	(0.18)	1.517***	(0.18)	1.507***	(0.18)
Compulsory School	Reference group							
Apprenticeship	1.686**	(0.28)	1.687**	(0.28)	1.676**	(0.28)	1.681**	(0.28)
Some Years of College	2.429***	(0.47)	2.430***	(0.47)	2.408***	(0.47)	2.428***	(0.47)
College Degree	3.849***	(0.77)	3.851***	(0.77)	3.803***	(0.76)	3.845***	(0.77)
University Degree	4.312***	(0.99)	4.314***	(0.99)	4.245***	(0.98)	4.326***	(1.00)
Age 16 – 24	Reference group							
Age 25 – 29	1.208	(0.39)	1.208	(0.39)	1.226	(0.40)	1.211	(0.39)
Age 30 – 39	1.056	(0.28)	1.055	(0.28)	1.077	(0.29)	1.054	(0.28)
Age 40 – 49	0.826	(0.22)	0.826	(0.22)	0.837	(0.22)	0.824	(0.22)
Age 50 – 59	0.666	(0.18)	0.666	(0.18)	0.682	(0.18)	0.666	(0.18)
Age 60 – 69	0.265***	(0.09)	0.265***	(0.09)	0.268***	(0.09)	0.264***	(0.09)
Age 70 plus	0.078***	(0.03)	0.078***	(0.03)	0.080***	(0.03)	0.078***	(0.03)
Less than 10,000 inhabitants	Reference group							
Btw. 10,000 & 50,000 inh.	1.003	(0.14)						
More than 50,000 inh.	1.174	(0.32)	1.141	(0.37)	1.353	(0.36)		
Less than 5,000 inh.			Reference group					
Btw. 5,000 & 50,000 inh.			0.975	(0.21)				
Less than 15,000 inh.					Reference group			
Btw. 15,000 and 50,000 inh.					1.211	(0.18)		
Less than 10,000 inh.							Reference group	
Btw. 10,000 & 40,000 inh.							0.999	(0.14)
More than 40,000 inh.							1.120	(0.29)
Number of inhabitants	1.000	(0.00)	1.000	(0.00)	1.000	(0.00)	1.000	(0.00)
Vienna	Reference group							
Lower Austria	0.912	(0.23)	0.911	(0.23)	0.903	(0.23)	0.896	(0.22)
Burgenland	0.986	(0.36)	0.981	(0.36)	1.029	(0.38)	0.967	(0.35)
Styria	0.632*	(0.15)	0.634*	(0.15)	0.620*	(0.14)	0.619*	(0.14)
Carinthia	0.634	(0.18)	0.634	(0.17)	0.638	(0.18)	0.614	(0.17)
Upper Austria	0.841	(0.18)	0.842	(0.18)	0.828	(0.18)	0.829	(0.18)
Salzburg	0.960	(0.28)	0.959	(0.28)	0.970	(0.28)	0.945	(0.27)
Tyrol	0.794	(0.21)	0.793	(0.21)	0.795	(0.21)	0.780	(0.20)
Vorarlberg	0.666	(0.21)	0.667	(0.21)	0.620	(0.20)	0.655	(0.21)
Suburbs	1.254	(0.21)	1.257	(0.21)	1.201	(0.20)	1.251	(0.21)
Self-employed	Reference group							
Civilservant	1.025	(0.41)	1.027	(0.41)	1.012	(0.40)	1.022	(0.41)
Leading Employee	1.431	(0.57)	1.433	(0.57)	1.440	(0.57)	1.429	(0.57)
Employee	1.376	(0.48)	1.378	(0.48)	1.377	(0.48)	1.373	(0.48)
Farmer/Household	0.717	(0.23)	0.716	(0.23)	0.711	(0.23)	0.716	(0.23)
Unemployed	0.873	(0.36)	0.874	(0.36)	0.867	(0.35)	0.873	(0.36)
Apprentice/Maternity	3.123**	(1.17)	3.128**	(1.18)	3.121**	(1.17)	3.126**	(1.18)
Retired	1.356	(0.39)	1.358	(0.39)	1.347	(0.38)	1.357	(0.39)
Non-Employed	Reference group							
Full Time	0.898	(0.28)	0.898	(0.28)	0.896	(0.28)	0.901	(0.28)
Part Time	1.196	(0.40)	1.195	(0.40)	1.190	(0.40)	1.199	(0.41)
Constant	0.172***	(0.07)	0.176***	(0.08)	0.163***	(0.07)	0.175***	(0.08)

Table 26 (cont'd):

Pseudo R^2	0.312	0.312	0.313	0.312
Number of Observations	2,923	2,923	2,923	2,923

Source: RTR survey data

Note: "***" significance at 0.1%; "**" significance at 1%; "*" significance at 5%; *Huber-White standard errors* in parentheses and "Pseudo R^2" refers to McFadden's pseudo R-square. Model (1): With fewer than 10,000 inhabitants as reference group and 10,000 – 50,000 & over 50,000 inhabitants variables. Model (2): With fewer than 5,000 inhabitants as reference group and 5,000 – 50,000 & over 50,000 inhabitants variables. Model (3): With fewer than 15,000 inhabitants as reference group and 15,000 – 50,000 & over 50,000 inhabitants variables. Model (4): With fewer than 10,000 inhabitants as reference group and 10,000 – 40,000 & over 40,000 inhabitants variables.

Table 27: Estimation results for the 2009 logit model with different population density variables

Variables	(1)		(2)		(3)	
	Coef.	Se	Coef.	Se	Coef.	Se
Household size	1.457***	(0.08)	1.459***	(0.08)	1.464***	(0.08)
Imputed Income	1.001***	(0.00)	1.001***	(0.00)	1.001***	(0.00)
Women	Reference group					
Men	1.504***	(0.18)	1.502***	(0.18)	1.508***	(0.18)
Compulsory School	Reference group					
Apprenticeship	1.685**	(0.28)	1.690**	(0.28)	1.687**	(0.28)
Some Years of College	2.439***	(0.48)	2.442***	(0.48)	2.439***	(0.48)
College Degree	3.927***	(0.79)	3.920***	(0.79)	3.920***	(0.79)
University Degree	4.445***	(1.02)	4.425***	(1.02)	4.408***	(1.01)
Age 16 – 24	Reference group					
Age 25 – 29	1.223	(0.39)	1.218	(0.39)	1.214	(0.39)
Age 30 – 39	1.051	(0.28)	1.049	(0.28)	1.048	(0.28)
Age 40 – 49	0.825	(0.22)	0.825	(0.22)	0.818	(0.21)
Age 50 – 59	0.660	(0.17)	0.661	(0.17)	0.657	(0.17)
Age 60 – 69	0.266***	(0.09)	0.266***	(0.09)	0.265***	(0.09)
Age 70 plus	0.078***	(0.03)	0.078***	(0.03)	0.078***	(0.03)
Pop Density less than 100	Reference group					
Pop Density btw. 100 & 800	0.934	(0.13)				
Pop Density more than 800	0.808	(0.20)				
Pop Density less than 100			Reference group			
Pop Density btw. 100 & 500			0.931	(0.13)		
Pop Density more than 500			0.873	(0.19)	0.992	(0.25)
Pop Density less than 50					Reference group	
Pop Density btw. 50 & 500					1.108	(0.19)
Number of inhabitants	1.000	(0.00)	1.000	(0.00)	1.000	(0.00)
Vienna	Reference group					
Lower Austria	0.809	(0.20)	0.845	(0.21)	0.848	(0.21)
Burgenland	0.882	(0.32)	0.919	(0.33)	0.892	(0.32)
Styria	0.577*	(0.13)	0.597*	(0.13)	0.591*	(0.13)
Carinthia	0.572*	(0.16)	0.597	(0.16)	0.605	(0.16)
Upper Austria	0.783	(0.17)	0.809	(0.17)	0.791	(0.17)
Salzburg	0.869	(0.25)	0.891	(0.25)	0.897	(0.26)
Tyrol	0.702	(0.19)	0.727	(0.19)	0.737	(0.19)
Vorarlberg	0.598	(0.20)	0.643	(0.20)	0.631	(0.20)
Suburbs	1.246	(0.22)	1.259	(0.22)	1.208	(0.20)
Self-employed	Reference group					
Civilservant	1.011	(0.40)	1.010	(0.40)	1.007	(0.40)
Leading Employee	1.412	(0.56)	1.411	(0.56)	1.410	(0.56)
Employee	1.357	(0.47)	1.352	(0.47)	1.347	(0.47)
Farmer/Household	0.710	(0.23)	0.711	(0.23)	0.715	(0.23)
Unemployed	0.875	(0.36)	0.872	(0.36)	0.871	(0.36)
Apprentice/Maternity	3.152**	(1.19)	3.135**	(1.18)	3.107**	(1.17)
Retired	1.356	(0.39)	1.352	(0.39)	1.346	(0.38)
Non-Employed	Reference group					
Full Time	0.909	(0.28)	0.912	(0.28)	0.912	(0.28)
Part Time	1.212	(0.41)	1.209	(0.41)	1.213	(0.41)
Constant	0.196***	(0.08)	0.190***	(0.08)	0.173***	(0.08)
Pseudo R^2	0.313		0.312		0.312	
Number of Observations	2,923		2,923		2,923	

Source: RTR survey data

Table 27 (cont'd):

Note: "***" significance at 0.1%; "**" significance at 1%; "*" significance at 5%; *Huber-White standard errors* in parentheses and "Pseudo R^2" refers to McFadden's pseudo R-square. Model (1): With a population density of less than 100 inhabitants/km^2 as reference group and 100 - 800 & over 800 inhabitants/km^2 variables. Model (2): With a population density of less than 100 inhabitants/km^2 as reference group and 100 – 500 & over 500 inhabitants/km^2 variables. Model (3): With a population density of less than 50 inhabitants/km^2 as reference group and 50 – 500 & over 500 inhabitants/km^2 variables.

Table 28: Estimation results for the 2009 logit model with different degree of urbanity variables

Variables	(1)	
	Coef.	Se
Household size	1.463***	(0.08)
Imputed Income	1.001***	(0.00)
Women	Reference group	
Men	1.499***	(0.17)
Compulsory School	Reference group	
Apprenticeship	1.686**	(0.28)
Some Years of College	2.429***	(0.48)
College Degree	3.875***	(0.78)
University Degree	4.327***	(1.00)
Age 16 – 24	Reference group	
Age 25 – 29	1.221	(0.39)
Age 30 – 39	1.055	(0.28)
Age 40 – 49	0.823	(0.22)
Age 50 – 59	0.664	(0.18)
Age 60 – 69	0.264***	(0.09)
Age 70 plus	0.077***	(0.03)
Non-Urbanized	Reference group	
Full Urbanized	1.045	(0.27)
Moderate Urbanized	3.503*	(1.82)
Number of inhabitants	1.000	(0.00)
Vienna	Reference group	
Lower Austria	0.865	(0.22)
Burgenland	0.928	(0.34)
Styria	0.604*	(0.14)
Carinthia	0.605	(0.17)
Upper Austria	0.792	(0.17)
Salzburg	0.918	(0.27)
Tyrol	0.758	(0.20)
Vorarlberg	0.561	(0.19)
Suburbs	1.204	(0.20)
Self-employed	Reference group	
Civilservant	1.036	(0.41)
Leading Employee	1.447	(0.57)
Employee	1.374	(0.48)
Farmer/Household	0.717	(0.23)
Unemployed	0.886	(0.36)
Apprentice/Maternity	3.115**	(1.17)
Retired	1.373	(0.39)
Non-Employed	Reference group	
Full Time	0.900	(0.28)
Part Time	1.194	(0.40)
Constant	0.182***	(0.08)
Pseudo R^2	0.313	
Number of Observations	2,923	

Source: RTR survey data

Note: "***" significance at 0.1%; "**" significance at 1%; "*" significance at 5%; *Huber-White standard errors* in parentheses and "Pseudo R^2" refers to McFadden's pseudo R-square.

Table 29: Estimation results for the 2006 multinomial logit model

Reference group: No Internet

Variables	DSL		CATV		Narrowband		Mobile BB	
	Coef.	Se	Coef.	Se	Coef.	Se	Coef.	Se
Household size	0.037***	(0.006)	0.007***	(0.002)	0.027***	(0.008)	0.001*	(0.000)
Women	Reference group							
Men (d)	0.069***	(0.015)	0.009*	(0.004)	0.035	(0.020)	0.002	(0.001)
Imputed Log Income	-0.043***	(0.009)	-0.017***	(0.003)	-0.019	(0.013)	-0.003***	(0.001)
Compulsory School	Reference group							
Some Years of College (d)	0.055*	(0.021)	0.002	(0.005)	0.150***	(0.032)	0.002	(0.002)
College Degree (d)	0.110***	(0.025)	0.011*	(0.006)	0.225***	(0.035)	0.003	(0.002)
University Degree (d)	0.236***	(0.038)	0.032**	(0.010)	0.260***	(0.047)	0.002	(0.002)
Age 16 – 24	Reference group							
Age 25 – 29 (d)	-0.066	(0.034)	0.004	(0.011)	-0.065	(0.048)	0.005	(0.005)
Age 30 – 39 (d)	-0.020	(0.034)	-0.013*	(0.006)	0.037	(0.057)	-0.000	(0.003)
Age 40 – 49 (d)	-0.000	(0.035)	0.002	(0.008)	-0.013	(0.049)	0.001	(0.003)
Age 50 – 59 (d)	-0.069*	(0.029)	-0.014*	(0.006)	-0.055	(0.045)	-0.003	(0.002)
Age 60 – 69 (d)	-0.107***	(0.026)	-0.028***	(0.005)	-0.102**	(0.037)	-0.003	(0.002)
Age 70 plus (d)	-0.263***	(0.016)	-0.047***	(0.005)	-0.171***	(0.028)	-0.005***	(0.001)
Non-Urbanized	Reference group							
Full Urbanized (d)	-0.059	(0.034)	0.033**	(0.012)	-0.141**	(0.043)	0.002	(0.003)
Moderate Urbanized (d)	-0.006	(0.066)	-0.020*	(0.010)	0.077	(0.105)	0.013	(0.019)
Vienna	Reference group							
Lower Austria (d)	-0.011	(0.033)	-0.006	(0.007)	-0.008	(0.049)	0.002	(0.003)
Upper Austria (d)	0.037	(0.035)	0.029*	(0.011)	0.010	(0.051)	0.002	(0.003)
Salzburg (d)	-0.060	(0.031)	0.001	(0.008)	-0.042	(0.047)	0.001	(0.003)
Burgenland (d)	-0.115***	(0.025)	-0.014*	(0.007)	-0.105**	(0.038)	-0.000	(0.003)
Styria (d)	-0.044	(0.030)	-0.007	(0.006)	-0.072	(0.038)	0.004	(0.004)
Carinthia (d)	0.022	(0.042)	0.003	(0.014)	-0.052	(0.047)	0.007	(0.007)
Tyrol (d)	0.006	(0.038)	0.014	(0.011)	-0.072	(0.041)	0.002	(0.004)
Vorarlberg (d)	-0.080*	(0.031)	-0.003	(0.010)	-0.070	(0.045)	-0.020***	(0.004)
Suburbs (d)	-0.004	(0.022)	0.009	(0.008)	-0.047	(0.027)	0.001	(0.002)
Cable available (d)	0.035*	(0.016)	0.157***	(0.015)	-0.016	(0.023)	0.002	(0.001)
DSL available (d)	0.114***	(0.019)	-0.088*	(0.044)	0.015	(0.061)	0.002	(0.002)
Pseudo R^2	0.3774							
Number of Observations	3,719							

Source: RTR survey data

Note: "***" significance at 0.1%; "**" significance at 1%; "*" significance at 5%; *Huber-White standard errors* in parentheses and "Pseudo R^2" refers to McFadden's pseudo R-square. "(d)" indicates the effect of a discrete change of the dummy variable from 0 to 1.

Table 30: Estimation results for the 2009 multinomial logit model

Reference group: No Internet

Variables	DSL		CATV		Narrowband		Mobile BB	
	Coef.	Se	Coef.	Se	Coef.	Se	Coef.	Se
Household size	0.066***	(0.009)	0.009	(0.006)	0.004***	(0.001)	-0.016*	(0.007)
Women	Reference group							
Men (d)	0.066**	(0.022)	-0.017	(0.016)	-0.003	(0.004)	0.017	(0.016)
Imputed Income	0.000***	(0.000)	0.000	(0.000)	-0.000	(0.000)	0.000***	(0.000)
Compulsory School	Reference group							
Apprenticeship (d)	0.050	(0.035)	-0.020	(0.025)	-0.006	(0.005)	0.041	(0.029)
Some Years of College (d)	0.139**	(0.043)	-0.009	(0.028)	-0.008	(0.004)	0.015	(0.033)
College Degree (d)	0.212***	(0.041)	-0.019	(0.025)	-0.001	(0.006)	0.027	(0.032)
University Degree (d)	0.165***	(0.047)	0.003	(0.030)	-0.005	(0.006)	0.074	(0.039)
Age 16 – 24	Reference group							
Age 25 – 29 (d)	-0.152***	(0.044)	0.005	(0.040)	-0.009	(0.005)	-0.050	(0.028)
Age 30 – 39 (d)	-0.149***	(0.041)	-0.010	(0.032)	-0.013**	(0.004)	-0.095***	(0.021)
Age 40 – 49 (d)	-0.151***	(0.041)	-0.050	(0.028)	-0.015***	(0.004)	-0.103***	(0.021)
Age 50 – 59 (d)	-0.142***	(0.043)	-0.085***	(0.025)	-0.008	(0.005)	-0.141***	(0.017)
Age 60 – 69 (d)	-0.242***	(0.038)	-0.133***	(0.025)	-0.018***	(0.004)	-0.150***	(0.021)
Age 70 plus (d)	-0.294***	(0.026)	-0.195***	(0.014)	-0.017***	(0.003)	-0.177***	(0.017)
Non-Urbanized	Reference group							
Full Urbanized (d)	-0.167***	(0.049)	0.125**	(0.042)	0.010	(0.010)	-0.048	(0.033)
Moderate Urbanized (d)	0.177	(0.115)	0.069	(0.102)	-0.023***	(0.003)	0.010	(0.074)
Vienna	Reference group							
Lower Austria (d)	-0.016	(0.047)	-0.119***	(0.021)	0.008	(0.010)	-0.047	(0.027)
Burgenland (d)	-0.067	(0.063)	-0.075*	(0.036)	-0.002	(0.009)	-0.007	(0.046)
Styria (d)	-0.079	(0.042)	-0.105***	(0.017)	-0.010*	(0.005)	-0.033	(0.026)
Carinthia (d)	-0.126**	(0.043)	-0.133***	(0.017)	-0.000	(0.007)	0.006	(0.038)
Upper Austria (d)	-0.084*	(0.039)	-0.058**	(0.021)	-0.010*	(0.005)	-0.015	(0.027)
Salzburg (d)	-0.031	(0.057)	-0.017	(0.030)	-0.004	(0.010)	-0.052	(0.030)
Tyrol (d)	-0.080	(0.045)	-0.114***	(0.018)	-0.006	(0.007)	0.027	(0.036)
Vorarlberg (d)	-0.100	(0.060)	-0.073*	(0.030)	0.020	(0.021)	-0.117***	(0.020)
Suburbs (d)	0.047	(0.033)	0.044	(0.034)	-0.004	(0.006)	-0.045*	(0.020)
Self-employed	Reference group							
Civilservant (d)	-0.099	(0.054)	-0.083**	(0.028)	-0.016***	(0.004)	0.086	(0.062)
Leading Employee (d)	-0.073	(0.055)	-0.033	(0.039)	-0.011	(0.006)	0.069	(0.055)
Employee (d)	-0.093	(0.051)	-0.075*	(0.034)	-0.006	(0.007)	0.080	(0.044)
Farmer/Household (d)	-0.090	(0.053)	-0.112***	(0.027)	-0.010*	(0.005)	-0.022	(0.044)
Unemployed (d)	-0.206***	(0.058)	-0.000	(0.060)	-0.031***	(0.004)	-0.031	(0.051)
Apprentice/Maternity (d)	-0.061	(0.047)	0.011	(0.035)	-0.011*	(0.004)	0.031	(0.042)
Retired (d)	-0.061	(0.051)	0.007	(0.046)	-0.002	(0.008)	-0.029	(0.049)
Non-Employed	Reference group							
Full Time (d)	0.002	(0.049)	0.039	(0.035)	-0.000	(0.007)	-0.055	(0.031)
Part Time (d)	-0.001	(0.051)	0.036	(0.040)	-0.004	(0.007)	0.011	(0.038)
Cable available (d)	-0.080**	(0.029)	0.110***	(0.024)	-0.006	(0.006)	-0.025	(0.021)
Pseudo R2	0.3001							
Number of Observations	2,923							

Source: RTR survey data

Note: “***” significance at 0.1%; “**” significance at 1%; “*” significance at 5%; *Huber-White standard errors* in parentheses and “Pseudo R^2” refers to McFadden’s pseudo R-square. “(d)” indicates the effect of a discrete change of the dummy variable from 0 to 1.

Abstract

The Internet is an important thing for many people these days: Writing emails, Internet surfing, online banking, online games or online communities are parts of our daily life. From a technical point of view everyone in Austria could have an Internet access but not everyone has. A brief look at the data reveals that there are significant differences in Internet penetration rates among households depending on the filter used e.g. age, income, gender, etc. This thesis provides an in-depth analysis of these various gaps and their possible determinants. In addition, an advanced model examines the relevant factors of the Internet access type decision. A special focus is dedicated to the urban-rural divide in penetration rates.

In the first part of the thesis, theories and dimension of the so called digital divide are introduced to give the reader an idea which gaps and determinants can be expected and have been found in the past. An alternative explanatory approach is the peer group theory which could only be presented in the theoretical part but not be tested with a statistical model due to data unavailability. The applied part starts with the discussion of the descriptive statistics of data sets from the years 2007 and 2009 provided by the Austrian Regulatory Authority for Broadcasting and Telecommunications (RTR) and Statistik Austria. As data from two different years is available, developments of possible gaps and their determinants can be examined. A logit model is created with the sociodemographic and geographical characteristics of each respondent as the explanatory variables and the dichotomous residential Internet access variable on the left side. Several models are estimated to check for robustness as there are different possible measures and cut-off points for a spatial gap. In a second step, an unordered multinomial logit model is estimated to find the determinants of the Internet access type e.g. DSL, Cable Internet (CATV), mobile Broadband, Narrowband.

Results from the descriptive statistics indicate substantial sociodemographic gaps in residential Internet access rates. Male, well educated and younger persons and multiple-person households are more likely to be online than others. Evidence for an urban-rural gap isn’t as conclusive as expected. Only moderate differences can be found depending on the spatial measure used e.g. population size, population density, degree of urbanity. In both years the estimated models confirm the assumed educational, age and household size gaps. The gender and income divide is doubtful because the only significant results were computed with the RTR data but not with the Statistik Austria data. Age and education are the major determinants of residential Internet access while the findings for an urban-rural divide are sparse. For 2006 depending on the measure and cut-off point, some significant odd-ratios are estimated but the overall explanatory contribution of the spatial variables to the models is

marginal. Absolutely no spatial influence on the dependent variable could be found with the 2009 data.

The findings from the multinomial logit model confirm once again the strong sociodemographic influence. In 2006 a change from “No Internet” to DSL or Narrowband is mostly effected by age and education while a switch to CATV is driven by its own availability. Unfortunately there are not enough observations for mobile Broadband to get significant results. The sociodemographic impact diminished over three years, as only age is a major determinant in the 2009 model. The computed significant federal state variables for CATV point towards availability issues rather than real regional influence.

Most of the logit and multinomial logit results are in line with recent published papers and real life experience. A more advanced multinomial logit model including product characteristics could be subject of future research.

Zusammenfassung

Das Internet nimmt heutzutage bei vielen Menschen einen sehr hohen Stellenwert ein: Emails schreiben, Internet surfen, online Banking, online Spiele oder online Communities sind Teil unseres täglichen Lebens. Theoretisch kann jeder österreichische Haushalt über einen Internetanschluss verfügen. Erhebungen jedoch zeigen, dass dies nicht die Realität ist. Ein kurzer Blick in die Statistiken offenbart, dass es signifikante Unterschiede zwischen den Internetpenetrationsraten und den soziodemograpischen Merkmalen (Alter, Einkommen, Geschlecht, etc.) gibt. Diese Diplomarbeit untersucht den Einfluss soziodemografischer Merkmale auf die Internetpenetrationsrate privater Haushalte. Zusätzlich wird versucht die relevanten Einflüsse auf die Wahl der Internetzugangsart zu schätzen. Ein weiterer Schwerpunkt wurde auf das Stadt-Land Gefälle bei den privaten Internetzugängen gelegt.

Im ersten Teil der Arbeit werden Theorien und die verschiedenen Dimension des sogenannten „digital divide“ vorgestellt um dem Leser eine Vorstellung davon zu geben welche Kluft und relevanten Faktoren erwartet werden können und welche davon in der Vergangenheit bei vorherigen Untersuchungen gefunden wurden. Ein alternativer Erklärungsansatz ist die „peer group“ (dt. Bezugsgruppe) These, welche nur im Theorieteil der Arbeit vorgestellt wird und nicht mittels statistischem Model getestet werden kann aufgrund nicht vorhandener Daten. Der angewandte Teil der Arbeit beginnt mit einer Diskussion der deskriptiven Auswertungen der Datensätze aus den Jahren 2006 und 2009, welche von der Rundfunk und Telekom Regulierungs-GmbH (RTR) und der Statistik Austria zur Verfügung gestellt wurden. Da Daten aus zwei verschiedenen Jahren vorhanden sind können Entwicklungen einer möglichen Kluft und deren bestimmenden Faktoren über die Zeit untersucht werden. Ein logit Model mit den soziodemographischen und geographischen Eigenschaften der Respondenten als erklärende Variablen und dem privaten Internetzugang als abhängige Dummy Variable wird erstellt. Mehrere Modelle werden geschätzt um die Ergebnisse auf ihre Robustheit zu überprüfen, weil es verschiedene Maße und Definitionen für ein Stadt-Land Gefälle gibt. Weiters wird ein ungeordnetes multinomiales logit Modell berechnet um die signifikanten Einflüsse auf die Wahl der Internetzugangsart (DSL, Kabel, mobiles Breitband, Schmalband) zu finden.

Die deskriptiven Ergebnisse deuten auf substantielle soziodemographische Lücken in der Verteilung der privaten Internetanschlüsse hin. Männliche, gut ausgebildete und jüngere Personen als auch Mehr-Personen Haushalte sind mit einer höheren Wahrscheinlichkeit online als andere. Hinweise auf ein Stadt-Land Gefälle sind nicht so eindeutig wie erwartet. Es werden nur geringe Differenzen gefunden abhängig davon welches Maß genommen wird z.B.: Einwohneranzahl, Einwohnerdichte, Urbanisierungsgrad. In beiden Jahren bestätigen die geschätzten Modelle den vermuteten „digital divide“ nach Bildung, Alter und

Haushaltsgröße. Geschlechter- und Einkommensunterschiede in der Internetnutzung sind fraglich da diese nur mit den RTR Daten gefunden wurden jedoch nicht mit den Daten der Statistik Austria. Alter und Bildung sind die bestimmenden Faktoren der privaten Internetnutzung während der Wohnort nur sehr wenig Einfluss darauf hat. Für das Jahr 2006 wurden abhängig vom gewählten Maß und Definition einige signifikante odd-ratios berechnet jedoch war im Allgemeinen der Beitrag der geographischen Variablen zur Erklärung der abhängigen Variable ein geringer. Für das Jahr 2009 konnte überhaupt kein relevanter Einfluss dieser Variablen gefunden werden.

Die Ergebnisse des multinomialen logit Modells bestätigen den starken Einfluss der soziodemographischen Variablen. Im Jahr 2006 wurde die Anschaffung eines DSL oder Schmalband Anschlusses hauptsächlich von Alter und Bildung des Respondenten beeinflusst während ein Wechsel zu Kabelinternet davon abhing ob dieses verfügbar war. Unglücklicherweise gibt es für 2006 nicht genügend Beobachtungen in der Kategorie „mobiles Breitband“ um signifikante Ergebnisse zu erhalten. Der soziodemographische Einfluss auf die abhängige Variable nimmt über einen Zeitraum von drei Jahren ab, da nur noch Alter ein bestimmender Faktor im Jahr 2009 ist. Die signifikanten Bundesländer odd-ratios für Kabelinternet deuten eher auf mangelnde Verfügbarkeit in diesen Bundesländern hin als auf echten regionalen Einfluss.

Fast alle Ergebnisse des logit als auch des multinomialen logit Models werden durch kürzlich erschienene Publikationen und Erfahrungen des alltäglichen Lebens gestützt. Ein vertiefendes multinomiales logit Modell welches Produktcharakteristika enthält könnte Gegenstand zukünftiger Forschung sein.

Curriculum Vitae Philipp Sandner

Citizenship:	Austrian
Born:	1985 in Neunkirchen, NÖ

Education

June 2003	High School Diploma (Matura) in BORG Wr. Neustadt
2004 – 2009	Studies in Economics and Statistics at the University of Vienna
September 2007 – February 2008	Studies in Economics at the Universidad de Alicante, Spain
March 2009	Bachelor degree in Statistics

Employment

2003 – 2004	Civilian Service at the Red Cross Wr. Neustadt
2009 – until now	Economist at RTR GesmbH

Languages

German:	Mother Tongue
English:	Very good knowledge
Spanish:	Basic skills