



MASTERARBEIT | MASTER'S THESIS

Titel | Title

Evaluating the Effectiveness of Targeted R&D Subsidies
A Staggered Difference-in-Differences Analysis of the “KMU-innovativ” Program in the German Automotive Industry

verfasst von | submitted by

Dominik Jonas Pfeifer B.Eng.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | UA 066 915
Degree programme code as it appears on
the student record sheet:

Studienrichtung lt. Studienblatt | Degree Masterstudium Betriebswirtschaft
programme as it appears on the student
record sheet:

Betreut von | Supervisor: Univ.-Prof. Dr. Stefan Wagner

Abstract

This thesis asks whether receiving KMU-innovativ funding leads German automotive SMEs to file more patents. The study builds a matched panel of 387 firms, 96 of which received KMU-innovativ grants and 291 that serve as controls, observed annually from 2004 to 2024. Patent data are collected from three databases (PATSTAT, LENS.ORG, and DPMA) and combined using the maximum count across sources to minimize undercounting. The causal effect is estimated using a staggered Difference-in-Differences design with a Poisson Pseudo-Maximum Likelihood estimator, which handles both the different entry dates of treated firms and the fact that over 92% of firm-year patent counts are zero.

The main result is a null. The pooled model finds no significant treatment effect on patent output, neither during the active treatment phase (IRR = 0.887, $p = 0.665$) nor in the post-treatment period (IRR = 1.270, $p = 0.545$). The event study confirms that treated and control firms followed parallel patent trajectories before treatment. The only individually significant coefficient is at $k = 0$, the year firms received the grant, where treated firms filed roughly 44% fewer patents than controls (IRR = 0.562, $p = 0.029$). This is consistent with the coordination costs of entering multi-partner R&D consortia, though the result rests on a single coefficient and should be treated as suggestive rather than definitive.

A key part of the analysis asks why the null result arises. A logistic regression shows that 86 of the 96 treated firms had fewer than 150 employees at baseline, which is the threshold at which patenting probability first crosses 50% in this sample. Most program recipients simply operate in a size range where patent filing is not something firms do regularly, regardless of how innovative they are. This suggests that the null result is better understood as a mismatch between the evaluation metric and the population being evaluated, rather than as evidence that the program had no effect on innovation.

Keywords: R&D subsidies, patent output, SME innovation, Difference-in-Differences, KMU-innovativ

Abstract (Deutsch)

Diese Masterarbeit untersucht, ob die Förderung durch das Programm KMU-innovativ dazu führt, dass deutsche KMU der Automobilbranche mehr Patente anmelden. Grundlage ist ein gematchtes Panel von 387 Unternehmen (96 geförderte, 291 Kontrollunternehmen), die jährlich von 2004 bis 2024 beobachtet werden. Die Patentdaten stammen aus drei Datenbanken (PATSTAT, LENS.ORG und DPMA) und werden nach dem Maximum-Prinzip zusammengeführt, um Untererfassungen zu vermeiden. Der kausale Effekt wird mit einem gestaffelten Difference-in-Differences-Ansatz und einem Poisson-Pseudo-Maximum-Likelihood-Schätzer (PPML) geschätzt, der sowohl die zeitlich versetzten Förderzeitpunkte als auch den hohen Nullanteil von über 92 % in den jährlichen Patentzahlen berücksichtigt.

Das zentrale Ergebnis ist ein Nullbefund. Das gepoolte Modell zeigt weder in der aktiven Förderphase ($IRR = 0,887$; $p = 0,665$) noch in der Nachförderphase ($IRR = 1,270$; $p = 0,545$) einen signifikanten Behandlungseffekt auf die Patentaktivität. Die Event-Study-Spezifikation bestätigt parallele Patentverläufe vor der Förderung. Der einzige individuell signifikante Koeffizient tritt bei $k = 0$ auf, dem Jahr des Förderbeginns: Geförderte Unternehmen meldeten rund 44 % weniger Patente an als vergleichbare Kontrollunternehmen ($IRR = 0,562$; $p = 0,029$). Dieses Ergebnis ist vereinbar mit den Koordinationskosten, die beim Eintritt in F&E-Konsortien mit mehreren Partnern entstehen, beruht jedoch auf einem einzelnen Koeffizienten und ist daher als Hinweis und nicht als gesicherter Befund einzuordnen.

Ein wesentlicher Teil der Analyse widmet sich der Frage, warum der Nullbefund zustande kommt. Eine logistische Regression zeigt, dass 86 der 96 geförderten Unternehmen zu Studienbeginn weniger als 150 Beschäftigte hatten. Dies entspricht dem Schwellenwert, ab dem die Patentierungswahrscheinlichkeit in dieser Stichprobe erstmals 50 %

erreicht. Die Mehrheit der Programmteilnehmer bewegt sich also in einem Größenbereich, in dem Patentanmeldungen unabhängig von der tatsächlichen Innovationstätigkeit keine regelmäßige unternehmerische Praxis darstellen. Der Nullbefund ist daher eher als Diskrepanz zwischen der gewählten Evaluationsmetrik und der untersuchten Firmenpopulation zu verstehen als Beleg dafür, dass das Programm keine Wirkung auf die Innovation entfaltet hat.

Schlüsselwörter: F&E-Förderung, Patentoutput, KMU-Innovation, Difference-in-Differences, KMU-innovativ

Table of contents

Abstract.....	I
Abstract (Deutsch).....	III
List of Abbreviations.....	VII
List of Tables.....	VIII
List of Figures.....	IX
1. Introduction	- 1 -
2. Literature Review	- 6 -
2.1. Market Failure, SME Financing, and the Rationale for Public R&D Support.....	- 6 -
2.2. R&D Subsidy Effectiveness and the Output Measurement Problem	- 8 -
2.2.1. What the Empirical Literature Has Established	- 8 -
2.2.2. Patent Counts as an SME Outcome Metric	- 10 -
2.3. KMU-innovativ, the Automotive SME Sector, and the Research Gap	- 12 -
2.3.1. The KMU-innovativ Program	- 12 -
2.3.2. The German Automotive SME Sector.....	- 14 -
2.3.3. Research Gaps and Hypothesis Development	- 15 -
3. Data and Descriptive Statistics	- 19 -
3.1. Project Selection	- 19 -
3.2. Definition of Treated and Control Groups	- 20 -
3.2.1. The Treated Group.....	- 20 -
3.2.2. Identification of Potential Controls	- 21 -
3.2.3. Matching Strategy	- 21 -
3.3. Data Sources	- 21 -
3.3.1. Project and Corporate Data.....	- 21 -
3.3.2. Patent Data Retrieval and Triangulation	- 22 -
3.3.3. Geospatial and Regional Data	- 23 -
3.4. Variable Construction.....	- 23 -
3.4.1. Core and Control Variables	- 23 -
3.4.2. Time and Treatment Phase Indicators	- 24 -
3.4.3. Derived Variables	- 24 -
3.5. Descriptive Statistics	- 24 -
3.5.1. Sample Composition and Balance	- 25 -
3.5.2. Distributional Properties and PPML Sample Restriction	- 27 -

3.5.3.	Patenting Propensity and Firm Size	- 28 -
3.6.	Summary and Methodological Outlook	- 30 -
4.	Empirical Strategy and Results	- 31 -
4.1.	Empirical Strategy	- 31 -
4.1.1.	Difference-in-Differences Framework	- 31 -
4.1.2.	Staggered Treatment Timing and the PPML Estimator	- 32 -
4.1.3.	Event Time and Observation Window	- 34 -
4.1.4.	Parallel Trends Test and Robustness Checks	- 35 -
4.2.	Empirical Results.....	- 37 -
4.2.1.	Parallel Trends Validation	- 37 -
4.2.2.	Main PPML Estimates.....	- 40 -
4.2.3.	Patenting Consistency Analysis	- 42 -
4.2.4.	Robustness Checks	- 45 -
4.2.5.	Firm Size Threshold	- 48 -
5.	Discussion.....	- 50 -
5.1.	Hypothesis Evaluation.....	- 50 -
5.2.	The Treatment-Onset Effect and a Possible Collaboration Paradox	- 53 -
5.3.	Patenting Behavior Among Treated Firms.....	- 55 -
5.4.	The Metric Mismatch: Synthesizing the Evidence.....	- 56 -
5.5.	Limitations.....	- 58 -
6.	Conclusion.....	- 62 -
	References	- 65 -
	Appendix	- 71 -
	Appendix A:	- 71 -
	Appendix B:.....	- 73 -
	Appendix C:.....	- 74 -
	Appendix D:	- 75 -
	Appendix E:.....	- 76 -
	AI Tools Disclaimer	- 77 -

List of Abbreviations

Abbreviation	Full Term
ANOVA	Analysis of Variance
BMBF	Bundesministerium für Bildung und Forschung (Federal Ministry of Education and Research)
BMFTR	Bundesministerium für Forschung, Technologie und Raumfahrt (Federal Ministry of Research, Technology and Space)
BKG	Bundesamt für Kartographie und Geodäsie (Federal Agency for Cartography and Geodesy)
CASE	Connected, Autonomous, Shared, Electric
CI	Confidence Interval
DiD	Difference-in-Differences
DPMA	Deutsches Patent- und Markenamt (German Patent and Trade Mark Office)
EPO	European Patent Office
EU	European Union
IP	Intellectual Property
IRR	Incidence Rate Ratio
KMU	Kleine und mittlere Unternehmen (Small and Medium-Sized Enterprises)
NACE	Nomenclature statistique des activités économiques dans la Communauté européenne (Statistical Classification of Economic Activities in the European Community)
OECD	Organisation for Economic Co-operation and Development
OEM	Original Equipment Manufacturer
PATSTAT	EPO Worldwide Patent Statistical Database
PPML	Poisson Pseudo-Maximum Likelihood
R&D	Research and Development
SME	Small and Medium-Sized Enterprise
VDA	Verband der Automobilindustrie (German Association of the Automotive Industry)
ZEW	Zentrum für Europäische Wirtschaftsforschung (Centre for European Economic Research)

List of Tables

Table	Title	Page
Table 1	Balance Statistics of the Matched Sample	26
Table 2	Distribution of Annual Patent Applications	27
Table 3	Predicted Patenting Probability by Firm Size	28
Table 4	Event Study Coefficients	37
Table 5	Main PPML Results	40
Table 6	Patenting Consistency Before and After Treatment	43
Table 7	Placebo Test Results	45
Table A1	Pairwise Correlations: Full Sample (N = 387)	70
Table A2	Pairwise Correlations: Treated Group (N = 96)	70
Table A3	Pairwise Correlations: Control Group (N = 291)	70
Table B1	Balance Statistics of the Patenting Subsample	72
Table B2	Geographic Distribution and Patent Output	73
Table C	Patenting Consistency by Treatment Status	73
Table D1	Characterization of the Treated Sample by Collaboration Type	74
Table D2	Partner Count and Patent Output: Key Correlations (Treated Group, N = 96)	74

List of Figures

Figure	Title	Page
Figure 1	Covariate Balance: Standardized Mean Differences	26
Figure 2	Dynamic Treatment Effects on Patent Output	38
Figure 3	Distribution of Placebo Post-Treatment Coefficients (500 Permutations)	46
Figure B	Geospatial Distribution of the Matched Sample	72
Figure E	Distribution of Treatment Cohorts	75

1. Introduction

The German automotive industry is at a turning point. Decades of dominance are being tested by a structural transition toward electric vehicles, software-defined mobility, and intensifying competition from manufacturers in China and the United States (Krzywdzinski et al., 2023). Thousands of small and medium-sized enterprises are affected by this and form the backbone of the German automotive supply chain. These are firms with between ten and 249 employees that manufacture components, develop processes, and carry a respectful part of the applied technological innovation that underpins the sector's global competitiveness.

Whether these firms can successfully navigate the transition depends in large part on their ability to invest in research and development (Cohen & Levinthal, 1990). But small firms face well-documented barriers to R&D investment. The knowledge generated by innovation tends to spill over to competitors, reducing the private return below the social optimum and leading to systematic underinvestment (Arrow, 1962). Small firms additionally face financing constraints that large corporations do not: without the collateral or internal capital to fund high-risk research projects, they are particularly vulnerable to credit rationing in private capital markets (Hall & Lerner, 2009). Public R&D subsidies are designed to address this, by reducing the cost of R&D investment while providing the upfront capital that resource-constrained firms could not otherwise access.

Germany's KMU-innovativ program is an instrument through which support reaches the SME sector. Established in 2007 by the Federal Ministry of Education and Research and operational since 2008, the program funds collaborative applied R&D projects between SMEs and university or research institute partners, with project durations of typically two to three years and grants that can cover up to 60% of eligible project costs. In the automotive

technology field alone, the program has funded 96 distinct projects between 2008 and 2019, involving firms spread across 14 German federal states.

Despite the program's scale and duration, rigorous causal evaluation of its effects on firm-level innovation output remains limited. The broader literature on R&D subsidy evaluation has established positive effects on R&D expenditure, so-called input additionality, with reasonable consistency. Whether that funding actually translates into measurable innovation outcomes like patents, new products, or productivity gains is a different question. The evidence on this so-called output additionality is much thinner and far less conclusive (Czarnitzki & Licht, 2006; Hottenrott & Lopes-Bento, 2014). For programs targeting small firms through collaborative structures, evidence is especially sparse. The only published system evaluation of the program, conducted by Aschhoff et al. (2012) for the BMBF, confirms that KMU-innovativ generates substantial input additionality in the form of increased R&D expenditure, but does not apply a causal estimation framework to patent output. Most existing German R&D subsidy evaluations, moreover, are cross-sectoral or cover broad high-technology samples, and their findings may not transfer to the automotive supply chain, where innovation cycles are tightly synchronized with vehicle generation timelines and supplier contracts increasingly mandate formal IP ownership. No existing evaluation isolates the automotive sector or formally examines whether patent counts are an appropriate outcome metric for the SME population that the program serves. This thesis addresses both gaps directly.

The central research question is: does receiving KMU-innovativ funding cause automotive SMEs to file more patents? Patents are used as the primary outcome measure because they are the most widely available, objective, and comparable indicator of innovative output at the firm level, and because they are the standard metric in the R&D subsidy evaluation literature (Griliches, 1990). At the same time, the literature documents that patent propensity

is substantially lower among SMEs than among large firms, driven by fixed costs, legal complexity, and the portfolio effect whereby smaller firms simply produce fewer innovations of patentable quality in any given period (Holgersson, 2012; Athreye et al., 2021). Whether patent counts are a reliable outcome metric for the specific firm population served by KMU-innovativ is therefore an empirical question that this thesis addresses alongside the primary treatment effect.

The analysis constructs a matched panel of 387 firms, 96 KMU-innovativ recipients and 291 control firms, matched on four-digit NACE industry code, baseline year, and employee count observed annually from 2004 to 2024, yielding 8,127 firm-year observations. Patent data are triangulated across three databases (PATSTAT, LENS.ORG, and DPMA) to minimize undercounting, with particular attention to German national filings that the SME literature identifies as the preferred filing route for small firms (Holgersson, 2012). To estimate the causal effect, the study uses a Difference-in-Differences design with a Poisson Pseudo-Maximum Likelihood (PPML) estimator. This approach handles two features of the data that standard linear models cannot: firms entered the program at different points in time, and over 92% of firm-year patent counts are zero. The PPML estimator keeps all those zeros in the estimation rather than dropping them, and the resulting coefficients can be read directly as incidence rate ratios (Santos Silva & Tenreyro, 2006).

The answer is no, at least not in any way detectable in the available patent data. The pooled PPML estimates are statistically insignificant for both the active treatment phase and the post-treatment phase, and the event-study specification confirms the absence of differential pre-treatment trends. But this null result is itself informative, and the bulk of the analytical contribution of this thesis lies in understanding why it arises. The descriptive analysis reveals that the program's firm population is structurally unlikely to generate patent output regardless of their innovation activity: of the 96 treated firms, 86 had fewer than 150

employees at baseline, placing the majority below an empirically estimated threshold at which patenting propensity first crosses 50% among R&D-active automotive SMEs. The event-study further documents a potential collaboration paradox at treatment onset: treated firms filed approximately 44% fewer patents in the year they received the grant, consistent with the IP coordination costs that arise when firms enter multi-partner R&D consortia. The patenting consistency analysis confirms that the null result holds whether patent activity is measured by volume or by frequency.

These findings add up to something bigger than the evaluation of a single program. The most direct contribution is that no previous study has estimated the causal effect of KMU-innovativ on patent output within a specific technology sector. Existing evaluations either pool across technology fields or stop at input additionality. Beyond the program-level result, the analysis shows that the majority of KMU-innovativ recipients operate below the firm size at which patent filing becomes statistically regular, which raises a more general question about whether patent-based evaluation is appropriate for programs that target this kind of firm. There is also a suggestive finding on the collaborative structure of the program itself: treated firms filed fewer patents in the year they entered the consortium, which points to IP coordination costs that are specific to multi-partner R&D grants and would not show up in evaluations of single-firm subsidies. The implication is that evaluating collaborative SME R&D programs through patent counts is likely to understate their impact, and that alternative approaches, including behavioral additionality measures and survey-based innovation indicators, would be better matched to the firms these programs actually serve.

The thesis is structured as follows. Chapter 2 reviews the theoretical foundations of public innovation policy and the empirical literature on R&D subsidy evaluation, and derives the three hypotheses that guide the analysis. Chapter 3 describes the data, the sample construction, and the descriptive statistics that characterize the firm population. Chapter 4 sets

out the econometric methodology, including the Difference-in-Differences framework, the PPML estimator, and the parallel trends test, and reports the results: the parallel trends validation, the main PPML estimates, the patenting consistency analysis, the robustness checks, and the firm size threshold analysis. Chapter 5 discusses the findings in light of the literature and draws out the policy and methodological implications. Chapter 6 concludes.

2. Literature Review

This chapter reviews the theoretical and empirical literature that underpins the research question of this thesis. It is organized in three sections. Section 2.1 establishes the theoretical rationale for public R&D support by examining the market failure arguments that justify government intervention and the mechanisms through which grants like KMU-innovativ are designed to correct them. Section 2.2 examines the empirical evidence on what such programs achieve, identifies the key debates and limitations. Section 2.3 places these general findings in the specific context of the KMU-innovativ program and the German automotive SME sector, and derives the hypotheses that guide the empirical analysis.

2.1. Market Failure, SME Financing, and the Rationale for Public R&D Support

The standard economic justification for public intervention in private R&D markets rests on the concept of market failure. Arrow (1962) pointed to three reasons why private markets underinvest in R&D: the results are hard to keep private (inappropriability), the outcomes are uncertain, and the costs are lumpy and hard to divide (indivisibility). The knowledge that comes out of R&D essentially behaves like a public good. Once it exists, it tends to spill over to competitors who did not bear the cost of producing it. The firm that does the research cannot capture all the benefits, because competitors can free-ride on the knowledge it creates. This lowers the incentive to invest, and the economy ends up with less R&D than would be socially optimal. That gap is the market failure that public subsidies are supposed to fix.

For SMEs, the problem goes further. On top of the general market failure, small firms face a financing gap that large firms do not. They usually lack the internal capital to fund risky R&D on their own, and they lack the collateral to borrow for it, which makes them especially vulnerable to credit rationing (Hall & Lerner, 2009). Beck and Demirguc-Kunt

(2006) show that this pattern holds across more than 80 countries: small firms consistently face bigger financing constraints than large firms, and these constraints directly limit their ability to grow and invest in innovation. The problem is not simply that small firms have less money. It is that information asymmetries between borrowers and lenders are larger for small firms, which makes private creditors unwilling to finance projects whose risk cannot be fully evaluated from the outside (Hall & Lerner, 2009).

Public innovation policy instruments are generally designed to intervene at one of two points in the innovation process. Input-oriented instruments like direct grants or R&D tax credits work by lowering the cost of doing R&D. The idea is straightforward: if the government puts in one euro, private R&D spending should go up by at least one euro as well, so that total R&D in the economy ends up higher than what the market would have produced without the subsidy (David, Hall, & Toole, 2000). Output-oriented instruments, by contrast, operate on the results of innovation rather than the inputs, for example through patent boxes that reduce the tax burden on income derived from intellectual property (OECD/Eurostat, 2018).

KMU-innovativ is an input-oriented grant program. But the program's stated goal goes beyond just increasing R&D spending. It aims for output additionality: new products, new processes, or new intellectual property that would not have been created without the funding (Aschhoff et al., 2012; BMBF, 2024). For the SME population it targets, the input-side design matters for a practical reason. Smaller firms frequently lack the upfront capital to initiate high-risk collaborative R&D projects at all, regardless of their willingness to do so. The grant removes the capital barrier. Whether that removal actually translates into measurable innovation output is a separate empirical question, and one that depends critically on the characteristics of the firms receiving the subsidy. This is the central question the empirical analysis of this thesis addresses.

2.2. R&D Subsidy Effectiveness and the Output Measurement Problem

2.2.1. What the Empirical Literature Has Established

The empirical literature on the effectiveness of public R&D subsidies is extensive, but the results are more mixed than the theoretical case for intervention might suggest. The most consistent finding concerns input additionality. When researchers compare the R&D spending of subsidized firms to unsubsidized controls, they generally find that the subsidies do increase spending. Czarnitzki and Licht (2006) document this for German manufacturing firms using a matching estimator, and Hottenrott and Lopes-Bento (2014) confirm it specifically for SME-targeted collaborative programs, including in the German context. The official system evaluation of KMU-innovativ, conducted by Aschhoff et al. (2012) for the BMBF, finds that the program increased annual R&D expenditure of participating firms by approximately 30%, with grants averaging around 110,000 euros per firm per year. This evidence of input additionality is well-established for the program.

Output additionality, whether funding translates into measurably more patents, new products, or productivity gains, is considerably harder to establish. Where positive effects on patent output are found, they tend to be concentrated among larger firms and programs with longer observation windows. A complementary strand of the literature documents a certification effect, whereby receiving a competitive public grant signals firm quality to external investors and thereby stimulates private investment beyond the grant itself (Wu, 2017). This effect operates independently of whether the grant produces direct innovation output.

A persistent methodological challenge across this literature is selection bias. Programs that use competitive application processes tend to attract firms with already-high innovative potential, making it difficult to separate genuine program effects from the continuation of pre-existing trajectories. This is the picking-winners problem. Almus and Czarnitzki (2003),

in their evaluation of East German R&D subsidies, find positive effects even after controlling for selection, but acknowledge that residual selection bias cannot be fully eliminated without randomization. In the context of KMU-innovativ, this issue is directly relevant: firms that apply for and win competitive grants are systematically different from firms that do not apply, and any evaluation that does not account for this risks attributing pre-existing firm characteristics to program effects.

A further consideration is that null results in this literature are more common than the share of published studies might suggest. Zuniga-Vicente et al. (2012), in a review of 77 empirical studies, find that results differ substantially depending on which firms are studied, over what time period, and how the subsidy and outcome variables are measured. This heterogeneity means that a null result is not necessarily evidence that a program has no effect. It may instead reflect a mismatch between the firm population and the outcome metric, an observation window too short to capture the time between grant receipt and measurable output, or a sample too small to detect an effect of the true size. As Section 2.2.2 documents, all three of these conditions are plausibly present in a patent-based evaluation of a small-firm program like KMU-innovativ.

A final dimension of program effectiveness that standard output metrics fail to capture is behavioral additionality. Gök and Edler (2012) define this as persistent changes in how firms organize their innovation activities as a consequence of participating in a funded program. These changes are distinct from input or output additionality. They include shifts in the research paths firms pursue, new collaborative relationships with universities or industry partners, and improvements in innovation management capabilities (Clarysse et al., 2009). Patent counts do not capture such changes. A program that successfully shifts how a firm conducts R&D without that shift producing a measurable patent within the observation window will appear ineffective in a patent-based evaluation. This problem is compounded for

small firms, where the propensity to patent is structurally lower than in large firms regardless of the innovation's value (Arundel & Kabla, 1998). Furthermore, engaging in collaborative research networks, while highly beneficial for behavioral learning, has actually been shown to have a negative impact on measurable output additionality like patents (Clarysse et al., 2009). A null result on patent output therefore need not indicate that the program had no effect. It may instead reflect the limits of the outcome metric rather than the limits of the program.

2.2.2. Patent Counts as an SME Outcome Metric

Most evaluations of SME innovation programs use patent counts as the primary outcome metric because they are publicly available, legally standardized, and precisely dated (Griliches, 1990). This is methodologically convenient, but it rests on an assumption that is systematically violated for small firms: that the population being evaluated has a meaningful propensity to patent its innovations.

SMEs patent at much lower rates than large firms, and this is not because they innovate less. Holgersson (2012) shows that small firms patent less even when they are equally innovative, simply because the fixed costs of patenting hit them harder. Legal fees, administrative burden, and the cost of enforcing patents all weigh more heavily on a firm with 30 employees than on one with 3,000. Arundel and Kabla (1998) make a similar point at the industry level: even in sectors with high patent propensity like chemicals, only about 36% of innovations get patented. In fabricated metals, the number drops to 8%. What this means in practice is that patent data only capture a fraction of all innovation, and that fraction gets smaller as the firm gets smaller. A small firm can be just as innovative as a large one and still show up with zero patents. The gap at the firm level is driven by portfolio volume rather than by reluctance to patent individual inventions. Both mechanisms predict that patent counts will systematically understate the innovation output of small firms relative to large

ones, not only because cost barriers sometimes deter small firms from protecting specific innovations, but primarily because they generate fewer innovations in any given observation window and are therefore statistically less likely to produce one of patentable quality within a fixed evaluation period.

These structural patterns have direct consequences for program evaluation. For a given level of innovation output, smaller firms have a lower probability of converting that output into a filed patent (Brouwer & Kleinknecht, 1999). This means that patent-based evaluation is structurally biased toward null or weak results for small firm populations, regardless of whether a genuine treatment effect exists. A program that successfully stimulates innovation in SMEs may leave no trace in patent databases if recipients protect their innovations through trade secrets, confidentiality agreements, or lead-time advantages instead of formal filings, or if the innovations produced do not meet the patentability threshold. Hughes & Mina (2010) documents precisely this pattern in a study of UK SMEs, finding that the majority of innovative small firms relied on non-patent appropriation mechanisms as their primary form of IP protection.

How large is this gap in practice? Athreye, Fassio, and Roper (2021) find that only 0.6 to 1.5% of SMEs published a patent in any given year, depending on the size definition used. For large firms, the figure is around 4%. For R&D-intensive small firms specifically, Hayter, Link and Scott (2018) estimate the elasticity of patenting with respect to R&D investment at between 0.41 and 0.55, meaning that even firms actively engaged in R&D convert less than half of a proportionate increase in R&D investment into additional patent filings. These figures imply that patent counts are a reliable innovation output metric primarily above a firm size threshold that may sit at or beyond the upper boundary of the SME definition itself. Identifying where that threshold lies for German automotive SMEs, and establishing

whether the majority of KMU-innovativ recipients fall below it, is one of the empirical contributions of this thesis.

A further complication arises when DiD methods are applied to patent data. Patent counts are non-negative integers with a large mass of observations at zero. Standard linear estimators applied to this kind of count outcome with many zeros produce inconsistent estimates under heteroscedasticity (Santos Silva & Tenreyro, 2006). The conventional workaround of applying a logarithmic transformation is unavailable when zeros are present, and the alternative of estimating the log of patents plus one produces results that are not directly interpretable and depend arbitrarily on the size of the constant added (Bellemare & Wichman, 2020). Santos Silva and Tenreyro (2006) recommend Poisson Pseudo-Maximum Likelihood (PPML) as the appropriate estimator for count outcomes with zero-inflation. PPML retains all zero-count observations as genuine data points, produces consistent estimates even when the conditional variance is misspecified, and generates coefficients that are directly interpretable as incidence rate ratios after exponentiation.

2.3. KMU-innovativ, the Automotive SME Sector, and the Research Gap

2.3.1. The KMU-innovativ Program

KMU-innovativ is run by the Federal Ministry of Education and Research (BMBF, 2024) and falls under the Zukunft der Wertschöpfung (future of value creation) initiative. The program started in 2007 as part of the Hightech-Strategie of the Federal Government (Aschhoff et al., 2012), the program funds high-risk, application-oriented R&D projects conducted by SMEs, either as standalone projects or in collaboration with universities, Fraunhofer Institutes, or other firms (Aschhoff et al., 2012; BMBF, 2024). Project durations are typically two to three years (Aschhoff et al., 2012), and grants are provided as non-repayable subsidies covering up to 50 percent of eligible costs for industrial research and up to 25 percent for experimental development, with rates elevated by 10 to 20 percentage

points for small and medium enterprises (BMBF, 2024; Commission Regulation (EU), 2014). The program targets firms meeting the EU SME definition of fewer than 250 employees and annual turnover below 50 million euros (Aschhoff et al., 2012).

Firms go through a two-stage competitive selection process (Aschhoff et al., 2012). They first submit a project outline, and if it passes, they are invited to submit a full application (Aschhoff et al., 2012; BMBF, 2024). This process creates non-random treatment assignment: firms that successfully navigate the selection procedure are likely to differ systematically from non-applicants in their pre-existing innovation capacity, managerial capability, and familiarity with formal R&D processes (Aschhoff et al., 2012). This is the institutional basis for the selection bias concern discussed in Section 2.2.1, and it is one reason why a matched difference-in-differences design with control groups is preferable to a simple before-after comparison for evaluating the program.

A mandatory incentive effect requirement, referred to as the *Anreizeffekt*, stipulates that the grant must induce R&D activity that would not have occurred in its absence, requiring applications to be submitted before work begins (Commission Regulation (EU), 2014). This legal requirement provides the institutional foundation for the additionality assumption that any causal evaluation of the program must rely upon. The only published system evaluation of *KMU-innovativ*, conducted by Aschhoff et al. (2012), confirms that the program generates substantial input additionality in the form of increased R&D expenditure. That evaluation does include patent data, tracking applications via the EPO's PATSTAT database and reporting that 14% of funded projects had already generated patent filings at the time of the survey. However, it does not apply a causal estimation framework to patent output in the way that this thesis does, leaving open the question of whether a formally identified treatment effect on patent filings exists.

2.3.2. The German Automotive SME Sector

The German automotive industry is structured as a tiered supplier hierarchy. At the top sit the large OEMs: Volkswagen, BMW, Mercedes-Benz. Below them, thousands of SMEs work as Tier 2 and Tier 3 suppliers. Many of these firms are what Simon (1996) calls Hidden Champions, companies that lead in global niche markets but operate with far fewer resources than their OEM customers. The VDA puts the number of workers affected by the powertrain transition at around 300,000 across 900 companies. SMEs making traditional combustion components face the most direct hit, of which SMEs form a substantial part (Verband der Automobilindustrie, 2024).

The sector is currently undergoing a structural transformation driven by the shift toward electric vehicles, software-defined mobility, and intensifying international competition, particularly from Chinese manufacturers (Krzywdzinski et al., 2023). This transition is pushing the innovation frontier away from the mechanical engineering domains where German SME suppliers have historically been competitive, toward software, electronics, and systems integration under what the industry calls the CASE paradigm: Connected, Autonomous, Shared, and Electric. For an SME supplier, this shift requires a level of R&D expenditure and organizational capability that frequently exceeds internal absorptive capacity as defined by Cohen and Levinthal (1990). Firms that cannot keep pace risk displacement from supplier networks as OEMs consolidate around technologically capable partners.

For automotive SMEs, patents play a complicated role. OEMs increasingly require formal IP protection from their suppliers, which pushes SMEs toward patenting. But patenting is expensive and legally complex, and most small firms do not have dedicated IP staff to handle it consistently (Holgersson, 2013).

Hughes and Mina (2010) document that the majority of innovative small firms in a comparable manufacturing context relied on non-patent appropriation mechanisms such as trade

secrecy and lead-time advantages as their primary form of IP protection. This sector-specific evidence reinforces the general finding from Section 2.2.2 that patent counts are an unreliable innovation metric for this population.

The quantitative evidence on SME patenting rates in this context is consistent with those expectations. Between 0.6 and 1.5 percent of SMEs had patents published in a given year in the broader literature (Athreye, Fassio, & Roper, 2021), and the elasticity of patenting with respect to R&D investment for small R&D-intensive firms is estimated at between 0.41 and 0.55 (Link & Scott, 2018). The empirical analysis in Chapter 3 tests where the patenting propensity threshold lies for German automotive SMEs specifically, generating an estimate of the minimum firm size at which patent-based evaluation becomes a statistically reliable tool.

2.3.3. Research Gaps and Hypothesis Development

The review of the literature in Sections 2.1 through 2.3.2 identifies three gaps that motivate this thesis.

The first gap is sectoral granularity. Most existing German R&D subsidy evaluations are cross-sectoral or cover broad high-technology samples. Findings from those studies may not transfer to the automotive supply chain, where innovation cycles are tightly synchronized with vehicle generation timelines and supplier contracts increasingly mandate formal IP ownership. No existing evaluation of KMU-innovativ isolates the automotive sector, leaving open whether program effects in this context differ from those found in more flexible industries.

The second gap is the absence of a formal test of metric appropriateness. As documented in Section 2.2.2, patent propensity increases sharply with firm size, and within the size range served by KMU-innovativ the majority of recipients are likely to fall below the threshold at which patenting becomes a reliable regular output. No existing evaluation of the program

has formally quantified this mismatch or examined whether patent-count models are statistically appropriate for the firm population to which they are applied. This thesis directly addresses that question.

The third gap is methodological. Standard difference-in-differences evaluations assume simultaneous treatment onset, which does not reflect the structure of a program running continuous competitive rounds since 2007. Callaway and Sant'Anna (2021) and Goodman-Bacon (2021) have shown that applying standard two-way fixed effects estimators in staggered settings can produce biased estimates when treatment effects are heterogeneous over time. This thesis applies a PPML estimator with an event-time specification that addresses this problem, representing a methodological improvement over prior evaluations of the program.

From these three gaps and the theoretical framework developed above, three hypotheses are derived.

H1 (Output Additionality): Drawing on Arrow's (1962) market failure framework and the input-output additionality literature reviewed in Section 2.2.1, KMU-innovativ grants are intended to reduce the marginal cost of R&D and thereby induce innovation output that would not otherwise occur. The Aschhoff et al. (2012) evaluation confirms input additionality for the program but does not estimate a causal effect on patent output. Given the structural patenting constraints documented in Section 2.2.2, any such effect is expected to be detectable primarily among firms with sufficient organizational scale to convert R&D activity into formal patent filings.

Hypothesis 1: KMU-innovativ funding is associated with an increase in annual patent output among treated firms relative to matched controls in the post-treatment period.

H2 (Selection Effect): Because the application process is competitive, the firms that get funded are not random (Aschhoff et al., 2012; BMBF, 2024). Following the picking-winners literature (Almus & Czarnitzki, 2003), treated firms are expected to exhibit higher pre-

treatment patenting activity than controls, reflecting the fact that firms already engaged with formal IP processes are more likely to apply for and win competitive R&D grants.

Hypothesis 2: Funded firms exhibit a higher pre-treatment patenting propensity than matched control firms, consistent with positive selection into the program.

H3 (Metric Mismatch): Athreye, Fassio, and Roper (2021) show that SMEs patent less than large firms mainly because they produce fewer innovations overall, not because they choose not to patent the ones they have. Add to that the fixed costs of filing (legal fees, prosecution, enforcement) and the fact that only 0.6 to 1.5% of SMEs publish a patent in any given year, and the picture is clear: turning R&D activity into formal patent filings takes a certain organizational scale. What nobody has established for the German automotive context is where exactly that threshold sits, and whether most KMU-innovativ recipients fall below it. If they do, then patent counts are simply the wrong metric for evaluating this program, no matter what it actually does for firm-level innovation.

Hypothesis 3: The majority of KMU-innovativ recipients in the automotive sector fall below the firm size threshold at which patenting becomes a statistically reliable measure of innovation output.

Beyond the three hypotheses, the collaborative structure of KMU-innovativ raises a question that this thesis can explore descriptively but cannot answer causally. The program requires multi-partner project structures, which provide access to knowledge spillovers and have been shown to translate into market novelties, particularly for internationally collaborating SMEs (Hottenrott & Lopes-Bento, 2014). But the behavioral additionality literature points to a tension: the same collaborative networks that help firms learn and innovate commercially can actually reduce their measurable output in the form of patents (Clarysse et al., 2009). When multiple firms with differing IP interests collaborate on a project, agreeing on patent ownership is complex, and the path of least resistance is often to protect results

through trade secrecy rather than formal filings. If this mechanism operates in the KMU-innovativ context, one would expect a negative association between partner count and patent output, independent of firm size. The estimation strategy in this thesis is designed to identify the average treatment effect of program participation, not to isolate the mediating role of consortium structure, so this relationship is examined descriptively in Appendix D rather than tested causally.

3. Data and Descriptive Statistics

To test these hypotheses, the study needed a dataset that covers a long enough time period and contains reliable measures of innovation output. Early on, alternative measures like product announcements in industry trade journals were considered. But the data was too patchy and inconsistent over the 21-year observation window, so patent counts were used instead. While certain industry bodies, such as the VDA (Verband der Automobilindustrie, 2021), provide historical innovation records, gaps in digital archives posed a substantial risk of measurement bias. Furthermore, the qualitative nature of industry magazine mentions makes them difficult to quantify reliably, as the volume of reported innovations often remains static regardless of actual R&D intensity. Consequently, this study adopts patent filing data as the primary proxy for innovation performance. Patents provide a standardized, legally verified, and time-stamped metric that is consistently available across the entire observation period (Griliches, 1990).

3.1. Project Selection

The primary data source for the treated group is the project database provided by the Federal Ministry of Education and Research (Bundesministerium für Forschung, Technologie und Raumfahrt [BMFTR], 2026).¹ This database catalogs initiatives under the umbrella program "Zukunft der Wertschöpfung," which seeks to develop innovative recommendations for the design of products, services, and industrial processes. One component of this initiative is the KMU-innovativ funding program, which provides direct R&D subsidies to small and medium-sized enterprises.

¹ The ministry was renamed from Bundesministerium für Bildung und Forschung (BMBF) to BMFTR in 2025. References to BMBF (2024) elsewhere in this thesis refer to directives published under the earlier name.

To identify the treated companies, a two-stage filtering process was implemented. First, the database was restricted to projects explicitly labeled under the KMU-innovativ program. Second, these projects were filtered to isolate those specifically relevant to the automotive industry. Two distinct search strategies were considered: categorizing every funded firm based on its primary industry affiliation, or categorizing firms based on the specific technological focus of the funded project. The first approach was rejected to avoid measurement noise, as firms that occasionally interact with the automotive sector might receive funding for projects entirely unrelated to vehicle technology. Consequently, the second approach was adopted, utilizing a keyword-based filtration strategy. Project titles and descriptions were searched for automotive-related keywords. This turned up 82 projects initially. Borderline cases were then reviewed by hand, bringing the final count to 96 automotive R&D projects.

The BMFTR database provided the following critical parameters for each project: the company name of the primary funding recipient, identification of consortia members including other SMEs or research institutions, the funding period spanning from 2008 to 2019, and the project title indicating the specific technological domain of the research.

3.2. Definition of Treated and Control Groups

3.2.1. The Treated Group

The 96 firms identified through the project selection procedure constitute the treated group. Each firm is assigned a treatment start year corresponding to the calendar year in which its KMU-innovativ project began. This year defines the event time origin ($t = 0$) for that firm, around which the pre- and post-treatment observation windows are constructed. The staggered nature of these start dates across the 2008 to 2019 window enables the Difference-in-Differences identification strategy described in Chapter 4.

3.2.2. Identification of Potential Controls

Potential control firms were pulled from the Orbis database (Moody's, 2025), which has longitudinal firm-level data on employee counts and industry codes. The initial pool was restricted to firms satisfying three criteria. First, the firm had to be registered in Germany, since KMU-innovativ only funds domestic firms. Second, it had to operate in NACE codes 24 through 29, which cover fabricated metals, machinery, motor vehicles, and transport equipment. This way the control firms face the same industry shocks as the treated firms. Third, the firm must satisfy the SME size threshold throughout the observation period. This screening procedure produced an initial pool of 50,592 potential control firms.

3.2.3. Matching Strategy

Control firms were selected in two steps to make them as comparable as possible to the treated firms. In the first stage, exact matching was applied on NACE four-digit industry code and baseline year, ensuring that each treated firm is compared only to control firms operating in the same sub-sector and observed over the same calendar years. In the second stage, nearest neighbor matching on baseline employee count was applied. For each treated company, three control firms were selected based on the closest available employee count, with a strict tolerance margin allowing a maximum size discrepancy of six percent. In instances where more than three identical matches were available, three firms were selected at random to avoid selection bias. The final sample has 387 firms: 96 treated and 291 controls. The panel runs from 2004 to 2024, giving 8,127 firm-year observations.

3.3. Data Sources

3.3.1. Project and Corporate Data

Project-level information was drawn directly from the BMFTR project database, providing treatment timing, partner count, and the university collaboration indicator for each treated firm. Corporate data including employee counts, industry classification, and

geographic location were sourced from Orbis (Moody's, 2025), which consolidates financial and structural firm data from national commercial registries across Europe. The integration of these two sources required a standardization of firm identifiers and geographic denominations. This was performed manually for the treated group, cross-referencing firm names, addresses, and registration numbers across both databases. For the control group, Orbis data were used directly.

3.3.2. Patent Data Retrieval and Triangulation

The primary dependent variable, annual patent filings per firm, was constructed through a multi-stage triangulation procedure across three platforms. PATSTAT (European Patent Office, 2025) served as the initial reference source, providing comprehensive coverage of European and international patent applications. LENS.ORG (LENS, 2025) was employed as a complementary source to enhance name-matching precision between funded firms and patent applicants, particularly for firms operating under multiple legal entities or trading names. Given the documented tendency of German SMEs to prioritize national patent protection over European or international applications (Holgersson, 2012; Athreye et al., 2021), all firms in the sample were additionally queried manually through the DPMA database (Deutsches Patent- und Markenamt, 2025). Where discrepancies arose between sources, the maximum count rule was applied: the higher filing count across all sources was recorded for each firm-year, under the assumption that a lower count likely indicates a failure of search parameters to capture all filings rather than an overcount. Prior to estimation, missing values in the annual patent variable were replaced with zero to complete the balanced panel structure, since the absence of a patent filing in a given year constitutes a genuine zero observation rather than missing data. This procedure resulted in a longitudinal panel of 2,656 filed patents across 644 non-zero data points within the 2004 to 2024 observation period.

3.3.3. Geospatial and Regional Data

Firm location data at the community level was extracted from Orbis and matched with the VG250 community-level administrative boundary files provided by the Federal Agency for Cartography and Geodesy (Moody's, 2025; Bundesamt für Kartographie und Geodäsie, 2026). The federal state identifier derived from this matching procedure was retained as a geographic control variable. As reported in Appendix B, geographic location at the state level shows no significant relationship with patent output in this sample, and the variable functions as an absorptive control in the regression specifications rather than carrying an independent narrative role.

3.4. Variable Construction

The resulting dataset is a balanced panel where each firm is represented by one observation per year over the 21-year period (2004 to 2024).

3.4.1. Core and Control Variables

The treatment status variable is a binary indicator equal to one for the 96 KMU-innovativ recipients and zero for the 291 matched control firms. The primary outcome variable records the annual count of patent filings per firm. Baseline employee count captures firm size at the year preceding treatment for treated firms and at the corresponding year for matched control firms. This variable serves both as the matching criterion and as a control variable in the regression specifications. The competition index is a distance-weighted measure of firm density in each firm's geographic neighborhood, constructed as the sum of inverse distances to all other firms in the sample. While this variable is retained as a control in the regression specifications to absorb regional variation, its construction conflates geographic density with the number of nearby firms in a way that limits direct interpretation. It is therefore not assigned an independent descriptive role in the analysis.

3.4.2. Time and Treatment Phase Indicators

The staggered structure of the KMU-innovativ program requires a relative time scale to synchronize analysis across firms with different treatment start years. The event time variable measures the distance in years from each firm's individual treatment start, with $t = 0$ corresponding to the first year of funding receipt. The treatment phase indicator bins event time into three analytically distinct periods: pre-treatment ($t < 0$), active treatment ($t = 0$ to $t = 2$, corresponding to the typical two- to three-year KMU-innovativ project duration), and post-treatment ($t \geq 3$). This categorical variable is the primary treatment indicator in the PPML specifications reported in Chapter 4.

3.4.3. Derived Variables

The binary variable `ever_patent` is defined as equal to one if a firm's cumulative patent count across all observed years is greater than zero, and zero otherwise. This variable identifies the 132 firms that filed at least one patent at any point in the observation window and defines the restricted sample used in the PPML estimation. The patenting consistency variables record the proportion of years within the five-year pre-treatment window ($t = -5$ to $t = -1$) and post-treatment window ($t = 0$ to $t = 4$) respectively in which a firm filed at least one patent. These variables are constructed for the full sample of 387 firms and used in the supplementary consistency analysis reported in Chapter 4. The university collaboration indicator is a binary variable equal to one if the KMU-innovativ project consortium included at least one university or publicly funded research institution, and zero for industry-only consortia. This variable exists only for the 96 treated firms and is used in the characterisation of the treated sample reported in Appendix D.

3.5. Descriptive Statistics

This section presents the descriptive analysis of the matched dataset. It establishes the quality of the matching procedure, examines the distributional properties of the outcome

variable, and quantifies the relationship between firm size and patenting propensity. Supplementary descriptive analyses, including the correlation structure, patenting subsample characterization, patenting consistency, consortium structure, and geographic distribution, are reported in Appendices A through D.

3.5.1. Sample Composition and Balance

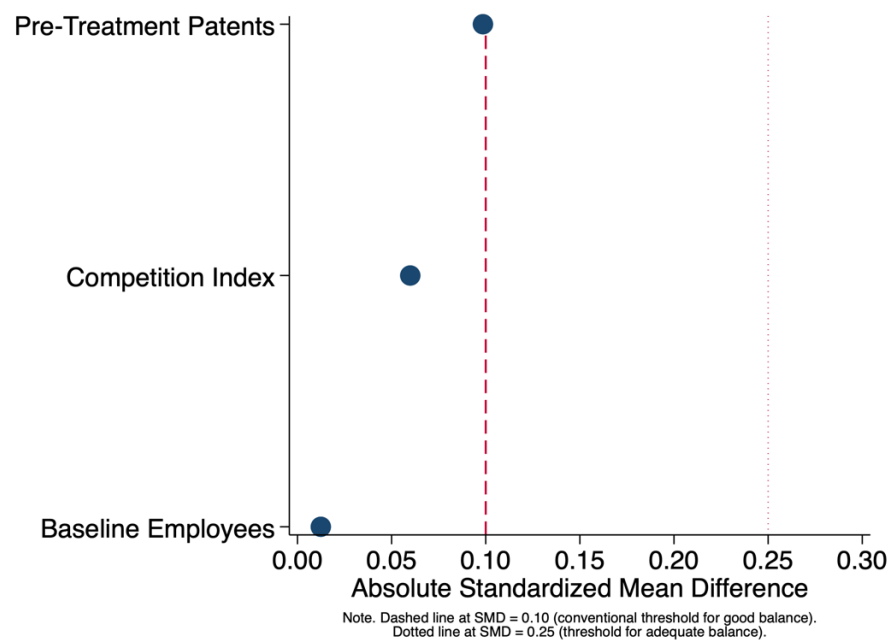
The final dataset comprises a balanced panel of 387 uniquely matched SMEs, 96 treated and 291 control firms, observed annually from 2004 to 2024. Before examining treatment effects, it is essential to verify that the matching procedure produced groups that are structurally and historically comparable.

Table 1 presents balance statistics for the matched sample. To assess matching quality, standardized mean differences (SMDs) are reported alongside conventional t-tests, following standard practice in the matching literature. Firm size is well balanced: treated firms average 66.2 employees compared to 67.0 for controls, with an SMD of -0.012 and a t-test p-value of 0.916. The competition index is slightly higher for control firms (20.4 versus 19.8, $SMD = -0.060, p = 0.611$), a difference that is far from statistical significance and will be absorbed by firm and year fixed effects in all specifications. Pre-treatment patent volume, the cumulative patent count over the four years prior to treatment, is 1.93 for treated firms and 1.27 for controls ($SMD = 0.099, p = 0.415$). Figure 2 presents these SMDs as a Love plot. All three variables fall at or below the conventional 0.10 threshold. It should be noted, however, that the 0.10 rule of thumb was developed for continuous, approximately normal covariates. Pre-treatment patent counts are zero-inflated integers with high variance, and the pooled standard deviation used in the SMD denominator is inflated by that variance, which compresses the SMD. The balance on this variable is adequate but should not be overinterpreted.

Table 1*Balance Statistics of the Matched Sample*

Variable	Treated (N=96)	Control (N=291)	Diff.	SMD	p-value
Firm Characteristics					
Baseline Employees	66.2	67.0	-0.76	-0.012	0.916
Competition Index	19.8	20.4	-0.58	-0.060	0.611
Innovation History					
Pre-Trend Patents (t-4 to t-1)	1.93	1.27	-0.66	0.099	0.415

Note. Pre-treatment patents reflect the cumulative count over the four years prior to treatment. SMD is the standardized mean difference, computed as the difference in group means divided by the pooled standard deviation. SMDs below 0.10 indicate good balance. p-values are from two-sample t-tests.

Figure 1*Covariate Balance: Standardized Mean Differences*

Note. Each point shows the absolute standardized mean difference between treated and control group means. The dashed line marks SMD = 0.10 (conventional threshold for good balance). The dotted line marks SMD = 0.25 (threshold for adequate balance). All three matching variables fall below the 0.10 threshold.

The pre-treatment patent balance is particularly important because it establishes innovation-trajectory equivalence between the two groups before any program effect could plausibly occur. The formal parallel trends test is presented in Chapter 4.

3.5.2. Distributional Properties and PPML Sample Restriction

The dependent variable exhibits severe zero-inflation. As shown in Table 2, 92.1% of all 8,127 firm-year observations record zero patent filings. Among the 7.9% of observations with positive counts, intensity is low: single filings account for 3.3% of all observations, counts between two and five for a further 3.1%, and counts above five for only 1.6%. This distributional structure is consistent with the broader SME patenting literature, which documents that occasional and irregular patent activity is the norm rather than the exception for small firms (Brouwer & Kleinknecht, 1999).

Table 2

Distribution of Annual Patent Applications

Annual Patent Count	Frequency	Percent	Cumulative
0	7,483	92.1%	92.1%
1	266	3.3%	95.4%
2-5	248	3.1%	98.5%
> 5	130	1.6%	100.0%
Total	8,127	100.0%	

This distributional structure motivates the use of the Poisson Pseudo-Maximum Likelihood estimator (Santos Silva & Tenreiro, 2006), which provides consistent estimates for

count data with severe zero-inflation without requiring the log transformation that would be undefined for zero-valued observations.

A further distributional feature creates a technical constraint for the PPML estimator. Firms that record zero patent filings across all observed years generate a complete separation problem within the PPML fixed effects framework: their firm-level fixed effects converge to negative infinity as the algorithm attempts to fit a zero outcome across all periods. The PPML estimation is therefore restricted to firms that filed at least one patent at any point in the 21-year observation window. This restriction is a technical requirement of the estimator rather than a substantive sample selection choice.

This restriction reduces the PPML estimation sample to 132 unique firms, 45 treated and 87 control, contributing 2,772 firm-year observations. The shift from the 3:1 matching ratio in the full sample to approximately 2:1 in the patenting subsample reflects the higher patenting propensity of treated firms, which is consistent with positive selection into the KMU-innovativ competitive grant process. Balance statistics for the patenting subsample, confirming that treated and control patenting firms remain well-matched on all key dimensions, are reported in Appendix B.

3.5.3. Patenting Propensity and Firm Size

To quantify the relationship between firm size and the probability of ever patenting, a logistic regression of `ever_patent` on baseline employee count was estimated across all 387 firms. The model is highly significant overall (LR chi-squared = 25.90, $p < 0.001$) with a pseudo R-squared of 0.052, indicating that firm size explains a meaningful but partial share of the variation in patenting participation. The coefficient on baseline employees is positive and significant (0.009, $p < 0.001$), confirming a monotonically increasing relationship between firm size and patenting probability across the full size range of the sample.

Table 3 reports the predicted patenting probabilities at selected firm sizes derived from the marginal effects of the logit model, holding all other covariates at their sample means.

Table 3

Predicted Patenting Probability by Firm Size

Firm Size (Employees)	Predicted Probability	95% CI
10	23.2%	[17.7%, 28.7%]
50	30.2%	[25.3%, 35.0%]
100	40.3%	[34.5%, 46.1%]
150	51.4%	[42.5%, 60.3%]
200	62.3%	[50.4%, 74.3%]
249	72.0%	[58.4%, 85.5%]

Note. Probabilities reflect patenting propensity within this R&D-active matched sample and are elevated relative to the general SME population. Derived from logistic regression of ever_patent on baseline employee count across 387 firms.

Two interpretive caveats are important. First, the absolute probabilities are elevated relative to what would be found in a random sample of German automotive SMEs, because both treated and control firms are R&D active by construction. Second, the patenting propensity measure captures whether a firm ever files a patent across the full 21-year observation window, which is a lower bar than annual patenting activity. The gradient from 23% at 10 employees to 72% at 249 employees nonetheless clearly illustrates the structural disadvantage faced by the smaller firms in the sample when evaluated against a patent-based outcome metric. Of the 96 treated firms, 86 had fewer than 150 employees at baseline, placing the majority below the estimated 50% patenting probability threshold. This finding directly supports Hypothesis 3, which predicts that the majority of KMU-innovativ recipients

fall below the firm size at which patenting becomes a statistically reliable measure of innovation output.

3.6. Summary and Methodological Outlook

The descriptive analysis presented in this chapter establishes several properties of the dataset that are directly relevant to the interpretation of the causal results in Chapter 4. The matching procedure produced a well-balanced sample on all pre-treatment dimensions that matter for identification: firm size, pre-treatment patent volume, and geographic distribution are statistically equivalent between treated and control groups, and the balance extends into the restricted patenting subsample of 132 firms (Appendix B).

The extreme zero-inflation of the patent count variable, with 92.1% of firm-year observations recording zero filings, motivates the PPML estimation strategy, while the complete separation problem for firms with zero patents across all periods motivates the restriction to the 132 firms with positive lifetime patent counts. This restriction is a technical necessity rather than a substantive sample selection, but its implications for statistical power are significant: the causal analysis ultimately rests on the behavior of 45 treated firms.

The logistic regression shows that patenting propensity increases steeply with firm size, and that the majority of treated firms fall below the 150-employee threshold at which patenting probability first crosses 50%. Supplementary analyses in the appendices confirm that treated firms are positively selected on pre-treatment patenting consistency (Appendix C), that the negative correlation between consortium partner count and patent output is independent of firm size (Appendix D), and that geographic location does not confound the treatment effect (Appendix B). Chapter 4 presents the empirical strategy and the causal results.

4. Empirical Strategy and Results

This chapter presents the econometric framework used to evaluate the causal effect of KMU-innovativ funding on patent output, followed by the empirical results. Section 4.1 describes the identification strategy, including the Difference-in-Differences framework, the adaptations required for staggered treatment timing and count data, and the design of the parallel trends test and robustness checks. Section 4.2 reports the results in five stages: the parallel trends validation, the main PPML estimates, the patenting consistency analysis, the robustness checks, and the firm size threshold analysis.

4.1. Empirical Strategy

4.1.1. Difference-in-Differences Framework

The core problem in evaluating any public subsidy is that the counterfactual outcome for funded firms cannot be observed (Rubin, 1974; Holland, 1986). A simple before-and-after comparison would confuse the effect of the grant with general industry trends, while a simple comparison between funded and unfunded firms would confuse it with the pre-existing differences that led firms to apply for and win competitive grants (Czarnitzki & Licht, 2006; Almus & Czarnitzki, 2003).

The Difference-in-Differences (DiD) method addresses both problems simultaneously (Angrist & Pischke, 2009). Rather than asking how much the treated firms changed, it asks how much more they changed compared to the control firms over the same period. If both groups would have followed the same trajectory without the program, then any extra change in the treated group can be attributed to the funding. This is done by including two sets of fixed effects in the model. Firm fixed effects remove any time-invariant differences between firms, such as their location, size history, or founding circumstances. Year fixed effects

absorb any shock affecting all firms in a given year, such as the 2008 financial crisis or a general downturn in the automotive sector. The basic equation takes the form:

$$Y_{it} = \alpha_i + \delta_t + \beta(Treated_i \times Post_t) + \varepsilon_{it}$$

Here Y_{it} is the number of patents filed by firm i in year t , α_i is the firm fixed effect, δ_t is the year fixed effect, and β is the coefficient of interest. If the parallel trends assumption holds, β gives an unbiased estimate of the causal impact of the KMU-innovativ funding (Angrist & Pischke, 2009).

4.1.2. Staggered Treatment Timing and the PPML Estimator

Two features of the data require adaptations to the standard DiD framework.

First, firms in this sample received KMU-innovativ funding at different points between 2008 and 2019. A standard DiD model assumes that all firms get treated at the same time. Here they do not. Goodman-Bacon (2021) shows that this creates a problem: the regression can end up comparing later-treated firms to already-treated firms, which biases the estimates. Callaway and Sant'Anna (2021) offer solutions for large samples. This study handles the issue by converting calendar time into event time, measured relative to each firm's own treatment start. That way, every treated firm is compared to control firms at the same point in the event timeline. The Callaway and Sant'Anna (2021) group-time ATT estimator was considered as an alternative, but the treatment cohort structure in this sample makes it impractical. As Appendix E shows, the 96 treated firms are spread across 12 entry cohorts between 2008 and 2019, with most cohorts containing fewer than 10 firms. After restricting to the patenting subsample, several cohorts shrink to two or three firms. Estimating cohort-specific treatment effects with groups this small would produce estimates dominated by noise, and the resulting confidence intervals would be uninformative. The event-time

specification used here pools across cohorts, which sacrifices the ability to detect treatment effect heterogeneity across entry waves but provides enough statistical power to estimate an average effect.

Second, the patent count variable is a non-negative integer with 92.1% of observations at zero. Standard linear estimators produce inconsistent estimates for count outcomes with this degree of zero-inflation under heteroscedasticity (Santos Silva & Tenreyro, 2006). The conventional logarithmic transformation is undefined for zeros, and the alternative of estimating the log of patents plus one produces results that are not directly interpretable and depend arbitrarily on the size of the constant added (Bellemare & Wichman, 2020).

This study therefore uses the Poisson Pseudo-Maximum Likelihood (PPML) estimator proposed by Santos Silva and Tenreyro (2006). PPML models the expected patent count as an exponential function of the explanatory variables, guaranteeing positive predicted values and handling zero observations without transformation:

$$E[Y_{it} | X_{it}] = \exp(\alpha_i + \delta_t + \beta \times D_{it})$$

The estimator remains consistent even when the data does not follow a strict Poisson distribution, and is robust to the overdispersion that characterizes patent data (Santos Silva & Tenreyro, 2006). The estimated coefficient is interpreted by exponentiating it to obtain the Incidence Rate Ratio (IRR). An IRR of 1.30 would mean that treated firms file 30% more patents per year than comparable control firms, after accounting for firm and year fixed effects.

The estimation uses the `ppmlhdf` command in Stata, which is built to handle models with many fixed effects (Correia, Guimaraes, & Zylkin, 2020). Standard errors are clustered

at the firm level because patent filings within the same firm are correlated over time (Bertrand, Duflo, & Mullainathan, 2004). As documented in Chapter 3, firms that never filed any patent across the entire observation period are excluded from the estimation due to the complete separation problem, restricting the PPML sample to 132 firms with at least one patent filing. After the `ppmlhdfe` algorithm drops singleton observations and observations separated by fixed effects, the estimation sample is further reduced.

4.1.3. Event Time and Observation Window

Because firms entered the program at different calendar years, the analysis converts calendar time into event time, measured in years before and after each firm's individual treatment start. A firm funded in 2012 and observed in 2015 is three years into the post-treatment period ($k = 3$), while a firm funded in 2016 and observed in 2015 is one year before treatment ($k = -1$). The year immediately before treatment ($k = -1$) serves as the reference point in all regressions. Control firms are assigned a pseudo-event year based on the treated firm they were matched to, ensuring comparability at the same point in the event timeline.

Rather than estimating a separate effect for each individual year, which would be unreliable given the small sample, event time is grouped into three broader phases. The pre-treatment phase covers all years before the funding starts and serves as the baseline. The active treatment phase covers $k = 0$ to $k = 2$, encompassing the typical two-year project duration of a KMU-innovativ grant (Aschhoff et al., 2012) plus one additional year to allow for any lag between project completion and patent filing. The post-treatment phase covers the years after the project phase ends ($k \geq 3$), capturing whether any effect persists once the funding has stopped.

The main analysis restricts the observation window to five years before and five years after the treatment start ($k = -5$ to $k = 4$). This symmetric window avoids imbalances arising from firms treated early in the program having few pre-treatment years and firms treated late

having few post-treatment years. The event study specification, which requires sufficient observations at each individual event time, uses a slightly narrower window of $k = -4$ to $k = 3$.

4.1.4. Parallel Trends Test and Robustness Checks

The DiD approach rests on the assumption that without the KMU-innovativ program, treated and control firms would have followed the same patent trajectory over time (Angrist & Pischke, 2009). This cannot be verified for the post-treatment period, but it can be tested in the pre-treatment period. The test is conducted using a dynamic version of the model that allows the difference between treated and control firms to vary freely in each year before and after treatment (Schmidheiny & Siegloch, 2023):

$$E[Y_{it} | \alpha_i, \delta_t] = \exp(\alpha_i + \delta_t + \sum_{k=-4, k \neq -1}^3 \beta_k \times D_{it}^k)$$

The indicator for the year immediately prior to treatment ($k = -1$) is omitted to serve as the baseline reference period, so that all estimated coefficients are interpreted as deviations from the treated-control difference at $k = -1$. The pre-treatment coefficients for k in $\{-4, -3, -2\}$ constitute the parallel trends test. If these coefficients are close to zero and statistically insignificant, this is conventionally taken as evidence supporting the assumption (Schmidheiny & Siegloch, 2023). However, Roth (2022) cautions that failing to reject parallel pre-trends does not guarantee that the assumption holds, as these tests frequently suffer from low statistical power. On top of that, only running the main analysis when the pre-test passes can itself introduce bias, making the problem worse if parallel trends are actually violated. The pre-trends test conducted here should therefore be interpreted as a necessary but not sufficient condition for the validity of the identification strategy.

To address the serial correlation concerns documented by Bertrand, Duflo, and Mullainathan (2004), standard errors throughout all specifications are clustered at the firm level. As an additional robustness check, a placebo test in the spirit of Bertrand et al. (2004) is conducted. Their work demonstrated that standard DiD models can severely over-reject the null hypothesis due to serial correlation in panel data, finding spurious significance in up to 45% of cases in Monte Carlo simulations with repeated random treatment assignments. The placebo test conducted here follows a simpler design: a single fake treatment is assigned to firms in the control group, using the same distribution of treatment years as the real program, and the model is re-estimated on these control firms only. Since none of these firms actually received KMU-innovativ funding, the model should find no effect. If it does, that would indicate the estimation method is detecting spurious patterns unrelated to the program.

A further robustness check tests whether the PPML model is correctly specified, following Pregibon (1980). This test checks whether the model's own predictions fully explain the outcome, or whether there is unexplained structure left over. Technically, the test adds the square of the model's predicted values as an additional regressor. If the model is correctly specified, this squared term should have no explanatory power. A significant squared term would indicate that the exponential mean function assumed by PPML does not fit the data, and that the results cannot be relied upon.

4.2. Empirical Results

This section presents the empirical results in five stages: the parallel trends validation, the main PPML estimates, the patenting consistency analysis, the robustness checks, and the firm size threshold analysis.

4.2.1. Parallel Trends Validation

Before reporting the main treatment effect estimates, the parallel trends assumption must be examined. The dynamic event study specification described in Section 4.1.4 estimates a separate coefficient for each year relative to treatment onset, with $k = -1$ as the omitted reference period. Table 4 reports the event study coefficients and Figure 2 visualizes the dynamic treatment effects with 95% confidence intervals.

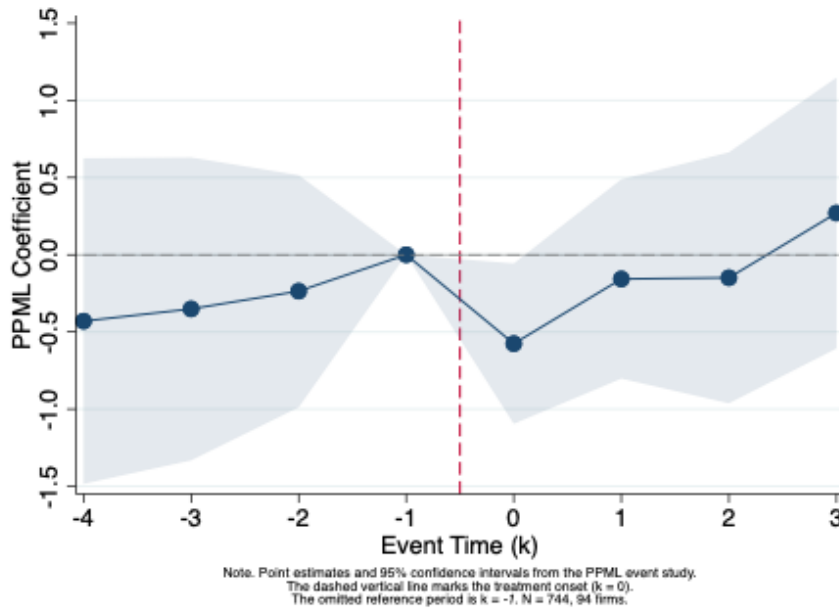
Table 4

Event Study Coefficients

Event Time (k)	Coefficient	IRR	SE	p-value	95% CI
-4	-0.430	0.651	0.538	0.425	[-1.484, 0.625]
-3	-0.351	0.704	0.500	0.483	[-1.332, 0.630]
-2	-0.236	0.789	0.384	0.538	[-0.988, 0.515]
-1	(baseline)				
0	-0.576	0.562	0.265	0.029	[-1.095, -0.057]
1	-0.157	0.855	0.329	0.634	[-0.802, 0.488]
2	-0.149	0.861	0.415	0.719	[-0.962, 0.663]
3	0.270	1.310	0.447	0.546	[-0.607, 1.147]

Note. PPML event study estimates. Dependent variable: annual patent count. Omitted reference period: $k = -1$. Firm and year fixed effects included. Standard errors clustered at the firm level. $N = 744$ (after singleton drops), 94 firms. Wald chi-squared(7) = 17.14, $p = 0.017$. Pseudo R-squared = 0.493.

Figure 2



Dynamic Treatment Effects on Patent Output

Note. Point estimates and 95% confidence intervals from the PPML event study. The dashed vertical line marks the treatment onset ($k = 0$). The omitted reference period is $k = -1$.

The pre-treatment coefficients at $k = -4$, $k = -3$, and $k = -2$ are all negative and statistically insignificant, with confidence intervals that comfortably encompass zero. The pre-treatment coefficients therefore support the parallel trends assumption. But as Roth (2022) warns, passing this test does not guarantee that parallel trends actually hold. The test may simply lack the power to pick up a small divergence between treated and control firms.

Roth (2022) proposes a formal sensitivity analysis that constructs bias-adjusted confidence intervals for the treatment effect under specified departures from perfect parallel trends. This approach would provide a more rigorous assessment of how robust the null result is to plausible pre-trend violations. However, the framework requires sufficient precision in the pre-treatment coefficient estimates to generate informative bounds. With only 94 firms in the event study specification and pre-treatment standard errors ranging from 0.384 to 0.538, the resulting bias-adjusted intervals would be wide enough to encompass both large

positive and large negative effects, offering little additional diagnostic value beyond what the unadjusted confidence intervals already convey. A formal Roth (2022) sensitivity analysis was therefore not conducted, though future work with larger samples could usefully apply this framework.

The post-treatment coefficients at $k = 1$, $k = 2$, and $k = 3$ are all statistically insignificant, with IRRs close to or below 1.0. There is no evidence of a positive treatment effect at any post-treatment horizon. The only significant coefficient is at $k = 0$ ($p = 0.029$), the year of treatment onset. This coefficient is negative (IRR = 0.562), indicating that treated firms filed approximately 44% fewer patents than comparable controls in the year they received the grant. A short-term dip in patent filings at the onset of a collaborative R&D project is consistent with firms redirecting resources toward project execution rather than IP filing during the initial project phase. It does not indicate a treatment benefit. However, given that this is a single significant result at the boundary between pre- and post-treatment, and that the confidence interval is wide, it should be interpreted with caution rather than as strong evidence of a specific mechanism. It is also worth noting that the event study estimates seven coefficients simultaneously. Under a Bonferroni correction for multiple comparisons, the adjusted significance threshold is $0.05/7 = 0.007$. The $k = 0$ p-value of 0.029 does not meet this threshold. The finding should therefore be regarded as suggestive rather than statistically established. The pooled PPML estimates reported in Section 4.2.2 provide the primary test of the treatment effect.

A power analysis based on the estimated standard errors provides further context for interpreting the results. At 80% power and the 5% significance level, the minimum detectable effect in the post-treatment period ranges from an IRR of approximately 3.19 at $k = 2$ to an IRR of 3.50 at $k = 3$. This means the analysis is powered to detect only large treatment effects, on the order of a doubling or tripling of patent output. A moderate but policy-

relevant effect, for instance a 50% increase in patent filings, would not be detectable given the sample size and the variance of patent counts in this population. Similarly, the pre-trends test can only detect pre-treatment divergences of substantial magnitude (MDE coefficients between 1.08 and 1.51), confirming that the failure to reject parallel trends is a limited diagnostic. These power constraints are a direct consequence of the small number of treated firms in the patenting subsample and the high variance of the patent count variable, both of which are structural features of the evaluation context rather than design choices that could be easily remedied.

To assess the sensitivity of these findings to potential violations of the parallel trends assumption, the maximum absolute pre-treatment deviation can serve as an informal benchmark. The largest pre-treatment coefficient is -0.430 at $k = -4$, which is itself far from significance ($p = 0.425$). Even if a post-treatment violation of similar magnitude were present but undetected, the overall pattern of largely insignificant post-treatment effects would remain unchanged, suggesting that the findings are robust to plausible departures from perfect parallel trends.

4.2.2. Main PPML Estimates

Table 5 reports the results of the primary PPML specification, which pools the post-treatment periods into a single coefficient to maximize statistical power. The estimation uses the full matched panel of 387 firms, restricted to the 132 firms with at least one patent filing and the symmetric window of $k = -5$ to $k = 4$.

Table 5

Main PPML Results

Coefficient	IRR	SE	p-value	95% CI
-------------	-----	----	---------	--------

Active treatment	-0.120	0.887	0.276	0.665	[-0.660, 0.421]
phase (k = 0 to 2)					
Post-treatment	0.239	1.270	0.396	0.545	[-0.536, 1.014]
phase (k >= 3)					

Note. N = 1,029 (after singleton drops), 104 firms. Pseudo R-squared = 0.504. Wald chi-squared(2) = 1.17, p = 0.557. Firm and year fixed effects included. Standard errors clustered at the firm level.

During the active treatment phase, the estimated effect on patent output is essentially zero. The IRR of 0.887 means treated firms filed approximately 11% fewer patents per year than comparable control firms during this period, but this difference is far from statistical significance (p = 0.665). The negative direction lines up with the significant k = 0 coefficient in the event study. It suggests that firms may shift attention away from patent filing while they are working on the funded project. The confidence interval is wide, meaning the data are consistent with both a substantial decrease and a substantial increase in patent output during the funding period.

After the treatment period ends, the coefficient turns positive but is nowhere near statistically significant. The IRR of 1.270 suggests treated firms filed approximately 27% more patents than controls after the grant period ended, but the parametric p-value of 0.545 is far from any conventional significance threshold and the confidence interval crosses 1.0 comfortably. The Wald test confirms that the two treatment phase indicators jointly explain none of the variation in patent output (p = 0.557).

The post-treatment estimate draws on 98 treated firm-year observations from 29 unique treated firms. The two highest-volume filers are Werth Messtechnik (109 total post-treatment patents) and Laser Zentrum Hannover (41 total post-treatment patents), which together account for approximately 40% of the 371 total post-treatment treated patents. Although

these two firms contribute a disproportionate share of the volume, the post-treatment estimate is spread across 29 firms rather than being dominated by a handful of outliers. The result is nevertheless not significant by any measure.

The pseudo R-squared of 0.504 reflects the fact that firm and year fixed effects alone explain a large share of the variation in annual patent filings, which is expected in a panel with many repeated observations of the same firms.

Taken together, the results point to a clear conclusion: the KMU-innovativ program did not produce a detectable effect on patent output among treated firms. Neither the active nor the post-treatment coefficient approaches statistical significance. This finding does not mean the program had no effect on the firms' R&D activity, their collaborative networks, or their technological capabilities. It means that no effect on patent counts is detectable in this sample. Given the structural features of patent activity in this population documented in Chapter 3, where the majority of treated firms never file a patent and those that do tend to do so occasionally and irregularly, this absence of a detectable effect is consistent with the metric mismatch argument developed in Chapter 2.

4.2.3. Patenting Consistency Analysis

Because the patent count variable is severely skewed, an alternative outcome measure is examined to confirm that the null result is not an artifact of this skew. Patenting consistency measures the proportion of years within a given window in which a firm filed at least one patent, regardless of volume. By aggregating the time-series data into single pre- and post-treatment measures for each firm, this approach inherently eliminates the serial correlation problem that plagues standard panel DiD estimators, implementing one of the primary solutions recommended by Bertrand et al. (2004). Two post-treatment windows are compared. Window A ($k = 0$ to $k = +4$) includes the active treatment years alongside the early post-treatment period. Window B ($k = +3$ to $k = +7$) excludes the active treatment years entirely

and captures only the true post-funding period, addressing the concern that gestation lags may suppress observable patent output during the project phase.

Table 6*Patenting Consistency Before and After Treatment*

	Window A		Window B	
	Control	Treated	Control	Treated
Pre-treatment (k = -5 to -1)	6.3%	14.2% (p = 0.002)	6.3%	14.3% (p = 0.002)
Post-treatment	8.2%	13.8% (p = 0.044)	8.3%	10.6% (p = 0.400)
Change (post minus pre)	+1.9 pp	-0.4 pp	+2.0 pp	-3.7 pp
DiD	-2.3 pp (p = 0.269)		-5.7 pp (p = 0.046)	

Note. Full sample of 387 firms (96 treated, 291 control). Consistency is the proportion of years within the respective window in which the firm filed at least one patent. Window A post = k = 0 to +4 (includes active treatment years). Window B post = k = +3 to +7 (true post-funding period only). Group-level p-values are from two-sample t-tests of treated versus control means. DiD p-values are from a first-differenced regression of the change in consistency on treatment status.

Treated firms are significantly more consistent patent filers than control firms before the program in both windows, with treated firms filing patents in 14.2 to 14.3% of pre-treatment years compared to 6.3% for control firms. This pre-existing gap reflects positive selection into the program rather than a treatment effect, consistent with Hypothesis 2.

The critical finding is in the change rows. In Window A, the change in consistency from pre- to post-treatment is -0.4 percentage points for treated firms and +1.9 pp for control firms. The DiD of -2.3 pp is not statistically significant (p = 0.269). Window B provides a sharper test: treated firms' consistency falls further to 10.6% while controls remain at 8.3%. The DiD in Window B is -5.7 pp in favor of controls (p = 0.046), driven entirely by treated firms declining by 3.7 pp while controls rise by 2.0 pp.

Both windows tell the same story: no evidence that KMU-innovativ increased the regularity of patent filings. If anything, treated firms' patenting consistency erodes relative to controls once the active project period ends. This pattern is consistent with positive selection on transitory pre-treatment patenting spikes followed by regression toward the sample mean. As Angrist and Pischke (2009) note, when treatment assignment is correlated with transitory past outcomes, the assumption of time-invariant fixed effects is violated. In the context of this consistency measure, treated firms may have been patenting at an unusually high rate in the years immediately preceding their grant application, making them appear as strong candidates in the competitive selection process. When their patenting rate naturally returns to its long-run baseline in the post-treatment period, it creates a spurious negative DiD. The significant negative estimate in Window B should therefore be interpreted as a methodological artifact of mean reversion rather than evidence that the program reduced patenting activity. This finding reinforces the main PPML result: the program had no detectable effect on patent output, and the apparent negative effect in the consistency analysis reflects the limits of the metric rather than a genuine adverse program impact.

4.2.4. Robustness Checks

The link test of Pregibon (1980) was applied to verify that the PPML exponential mean function is correctly specified. When regressing the outcome on the model's predicted linear index and its square within the fixed effects framework, the coefficient on the linear index was exactly 1.000, confirming that the model reproduces its own predictions. The coefficient on the squared term was essentially zero ($p = 1.000$), indicating that no unexplained nonlinear structure remains in the data. There is no evidence of functional form misspecification, meaning the main result cannot be attributed to an incorrectly specified conditional mean.

As an additional check, a placebo treatment was randomly assigned to firms in the control group, with pseudo-treatment timing drawn from the same distribution as actual

treatment years. The PPML specification was then re-estimated on the control group only, substituting the real treatment variable with the placebo indicator and clustering standard errors at the firm level.

Table 7

Placebo Test Results

	Coefficient	p-value
Placebo active phase	0.312	0.432
Placebo post phase	1.084	0.050

Note. Control group only. N = 653 (after 214 singleton drops), 66 firms. Pseudo-treatment timing drawn from the same distribution as actual treatment years. Standard errors clustered at the firm level.

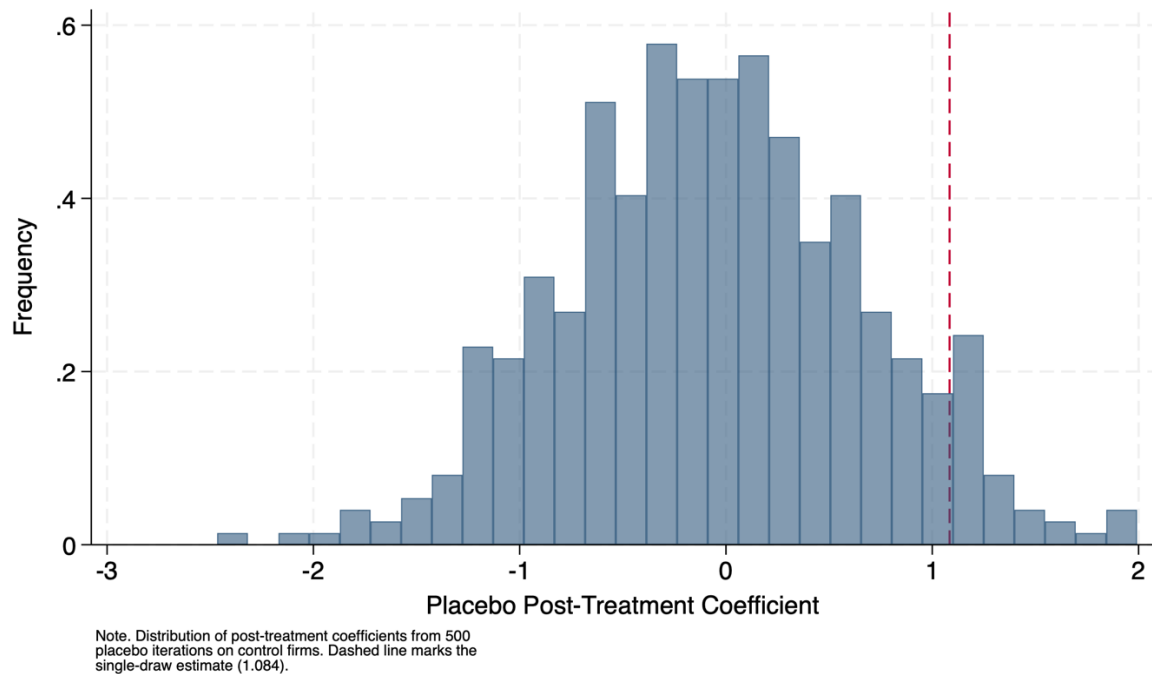
The placebo active coefficient is not significant ($p = 0.432$), so there is no false signal during the pseudo-active phase. The placebo post coefficient from the single random draw is large (1.084) and borderline significant ($p = 0.050$).

To assess whether this single-draw result is unusual, a permutation-based placebo test was conducted. Treatment dates were randomly reassigned to control firms 500 times, and the PPML specification was re-estimated for each permutation. Figure 3 shows the distribution of the resulting placebo post coefficients. The mean across 500 iterations is -0.057 with a standard deviation of 0.721, confirming that the expected placebo effect is approximately zero. Of the 500 permutations, 33 produced a post coefficient at least as large as the single-draw estimate of 1.084, yielding a permutation p-value of 0.066. The single-draw result therefore sits in the upper tail of the distribution but is not exceptional: roughly one in fifteen random treatment assignments produces a pseudo-effect of this magnitude.

This confirms that the borderline significant placebo reflects the high variance of patent counts in this population rather than a systematic flaw in the estimation procedure. Patent

counts are noisy enough in this sample that random treatment assignment can produce large pseudo-effects, which reinforces the metric mismatch argument developed in Chapter 2.

Figure 3



Distribution of Placebo Post-Treatment Coefficients (500 Permutations)

Note. Each bar represents the frequency of PPML post-treatment coefficients from 500 random placebo assignments on control firms. The dashed line marks the single-draw estimate (1.084). Of the 500 permutations, 33 (6.6%) produced a coefficient at least as large.

If the estimation method fires on control firms with a random treatment date, the null result for actual treated firms is more likely to reflect low statistical sensitivity than a genuine absence of any program effect. The robustness checks confirm that the main finding of no detectable treatment effect on patent output is genuine, and that the estimation method, while correctly specified in terms of functional form, operates in a setting where the outcome variable is too noisy to serve as a reliable program evaluation metric.

As an additional robustness check, an alternative matching specification was explored in which treated and control firms were matched on pre-treatment patent trajectories in addition to NACE code and baseline employee count. Control firms were selected as nearest

neighbours on average pre-treatment patent output, with zero-patent firms matched to zero-patent controls in a first tier and non-zero patent firms matched by continuous distance in a second tier. This procedure produced a 2:1 matched sample of 288 firms (96 treated, 192 control) and achieved strong parallel pre-trends across the full sample (joint pre-trend F-test $p = 0.968$). The PPML specification was re-estimated on the patenting subsample of this alternative panel. The main treatment effect is $IRR = 1.364$ ($p = 0.160$), which is not statistically significant. The null result is therefore robust to this more restrictive matching and is not driven by imbalances in pre-treatment patenting propensity between the treated and control groups. The alternative matching was not adopted as the primary specification because the tighter matching reduced sample coverage substantially and limited the scope for additional robustness checks.

4.2.5. Firm Size Threshold

The logistic regression estimated in Chapter 3 (Section 3.5.3) established that the predicted probability of ever patenting crosses 50% at approximately 150 employees. Of the 96 treated firms, 86 had fewer than 150 employees at baseline. These firms are structurally unlikely to generate patent output even when conducting genuine R&D, reflecting the well-documented cost and complexity barriers that small firms face in the patenting process (Hogersson, 2012; Athreye et al., 2021).

This threshold has direct implications for the interpretation of the results. The PPML estimation sample is reduced by singleton drops from 132 patenting firms to 104, and the post-treatment estimate draws on 98 firm-year observations from 29 unique treated firms. The majority of treated firms never enter the estimation at all, not because they are uninnovative, but because they operate below the firm size at which patent filing becomes a statistically regular activity. For evaluations targeting firms below 150 employees, alternative outcome measures such as R&D expenditure, process innovation, or productivity growth

would provide a more appropriate and sensitive test of program effectiveness. The 150-employee threshold should be interpreted as a rough empirical guideline derived from this specific R&D-active sample rather than a universal rule, but it offers a practical benchmark for future evaluation design.

5. Discussion

This chapter interprets the empirical results presented in Chapter 4 in light of the theoretical framework and hypotheses developed in Chapter 2. The central research question of this thesis is whether receiving KMU-innovativ funding causes automotive SMEs to file more patents. The answer, across all specifications and robustness checks, is that no such effect is detectable in the available data. This section discusses what that null result means, why it arises, and what it implies for the evaluation of collaborative SME R&D programs.

The discussion is organized as follows. Section 5.1 evaluates the results against each of the three hypotheses. Section 5.2 examines the treatment-onset finding and the potential collaboration paradox that may explain it. Section 5.3 considers why a minority of treated firms do patent despite the overall null result. Section 5.4 synthesizes the metric mismatch argument. Section 5.5 addresses the limitations of the study.

5.1. Hypothesis Evaluation

Hypothesis 1 predicted that KMU-innovativ funding would be associated with an increase in annual patent output among treated firms relative to matched controls in the post-treatment period. This hypothesis is not supported. The pooled PPML model estimates an active treatment coefficient of -0.120 (IRR = 0.887, $p = 0.665$) and a post-treatment coefficient of 0.239 (IRR = 1.270, $p = 0.545$). Neither coefficient approaches conventional significance thresholds. The event-study specification reinforces this conclusion: of the seven event-time dummies estimated, only $k = 0$ reaches significance, and it is negative rather than positive. The patenting consistency analysis confirms the null result using an alternative outcome measure, with no evidence that funded firms increased the regularity of their patent filings relative to controls.

This finding is consistent with the broader pattern documented in the R&D subsidy literature, where output additionality is substantially harder to establish than input additionality (Czarnitzki & Licht, 2006; Zúñiga-Vicente et al., 2012). The Aschhoff et al. (2012) system evaluation confirmed that KMU-innovativ generates a roughly 30% increase in annual R&D expenditure among participating firms, but the present analysis shows that this additional R&D spending does not translate into detectable patent output. The gap between input and output additionality is not an anomaly specific to this study. It reflects a structural feature of the firm population being evaluated, as the following discussion of Hypothesis 3 and 4 makes clear.

Hypothesis 2 predicted that funded firms would exhibit a higher pre-treatment patenting propensity than matched control firms, consistent with positive selection into the program. This hypothesis is supported. In the patenting consistency analysis, treated firms filed patents in 14.2% of pre-treatment years compared to 6.3% for control firms, a difference that is statistically significant ($p = 0.002$). The correlation analysis in Appendix A shows that the size-patent correlation is 0.444 within the treated group compared to 0.202 within the control group, indicating that treatment selection favors more patent-active firms within the SME size range.

This selection pattern is consistent with the picking-winners literature (Almus & Czarnitzki, 2003) and with the institutional design of the KMU-innovativ application process. The competitive two-stage selection procedure documented in Section 2.3.1 favors firms with prior R&D experience and familiarity with formal innovation management, precisely the firms most likely to have filed patents before applying. However, the parallel trends test shows that this selection does not generate differential pre-treatment trends: the event-time coefficients at $k = -4$, $k = -3$, and $k = -2$ are all statistically insignificant and small in magnitude, confirming that treated and control firms followed parallel patent trajectories

before treatment. The selection effect is therefore a level difference, not a trend difference, and the DiD identification strategy remains valid.

Hypothesis 3 predicted that the majority of KMU-innovativ recipients in the automotive sector fall below the firm size threshold at which patenting becomes a statistically reliable measure of innovation output. This hypothesis is strongly supported. The logistic regression reported in Section 3.5.3 estimates that the predicted probability of ever patenting crosses 50% at approximately 150 employees. Of the 96 treated firms, 86 had fewer than 150 employees at baseline. These firms are structurally unlikely to generate patent output even when conducting genuine R&D, because the fixed costs of patenting, legal fees, prosecution complexity, and ongoing maintenance weigh disproportionately on smaller organizations (Holgersson, 2012; Athreye et al., 2021).

The practical consequence is that the PPML estimation rests on a severely reduced sample. After the necessary restriction to firms with at least one lifetime patent and the removal of singletons by the estimator, the main specification retains 104 firms contributing 1,029 observations, with only 29 unique treated firms contributing patent filings in the post-treatment period. The majority of the 96 treated firms never enter the causal estimation at all, not because they are uninnovative, but because they operate in a size range where patent filing is not a regular feature of firm behavior. For these firms, the outcome metric is structurally incapable of registering the kind of innovation output that a collaborative R&D program might plausibly generate: process improvements, prototype development, knowledge absorption, or the establishment of research partnerships.

This finding has implications that extend beyond the present study. Any patent-based evaluation of an SME R&D program whose recipients fall predominantly below 150 employees is testing the metric as much as the program. The null result cannot distinguish two possibilities: the program had no effect on innovation, or the effect manifests in forms that

patent counts cannot capture. The descriptive evidence on firm size and patenting propensity strongly favors the second interpretation.

The exploratory analysis of consortium size and patent output yields a descriptive pattern that is consistent with the theoretical expectation but falls short of causal evidence. The pairwise correlation between project partners and total patents within the treated group is -0.212 and statistically significant, while the correlation between partners and baseline employees is essentially zero (-0.002), indicating that the negative association is not simply a proxy for firm size. However, this is a bivariate correlation across 96 observations with no controls for other firm characteristics. It cannot distinguish whether larger consortia cause lower patent output, whether firms with lower patenting propensity self-select into larger consortia, or whether both are driven by an unobserved third factor. The pattern is noted here as a descriptive regularity that is consistent with the IP coordination costs discussed in Section 5.2, but it should not be interpreted as evidence for a causal mechanism.

5.2. The Treatment-Onset Effect and a Possible Collaboration Paradox

The event-study specification produces one result that departs from the general null pattern: the coefficient at $k = 0$, the year of treatment onset, is -0.576 and statistically significant (IRR = 0.562, $p = 0.029$, 95% CI $[-1.095, -0.057]$). As noted in Section 4.2.1, this p -value does not survive a Bonferroni correction for seven simultaneous tests (adjusted threshold: 0.007), so the result should be treated as suggestive evidence rather than a statistically established finding. This means that treated firms filed approximately 44% fewer patents in the year they received the grant, relative to what would have been expected based on the control group's trajectory. No other event-time coefficient is individually significant.

This finding is consistent with what may be termed a collaboration paradox in the context of publicly funded R&D consortia. The KMU-innovativ program requires collaborative project structures, typically involving three or more partners including other SMEs and

university or research institute partners. When firms enter such multi-partner consortia, they face immediate transaction costs that do not arise in single-firm R&D: negotiating joint IP ownership, clarifying who files and who licenses, and aligning the patent strategies of partners with divergent commercial interests. These coordination costs are well-documented in the literature on collaborative R&D. Hottenrott and Lopes-Bento (2016) find that while R&D collaboration generally enhances innovation outcomes, there is a tipping point beyond which additional partner-ship complexity reduces measurable output. Clarysse et al. (2009) make a related argument specifically in the context of behavioral additionality: participating in collaborative research networks is highly beneficial for organizational learning, but it has a negative impact on measurable output additionality such as patents.

The significant negative coefficient at treatment onset is consistent with these collaboration costs. Rather than navigating complex co-ownership agreements, firms entering KMU-innovativ consortia may rely on trade secrets or non-disclosure agreements to protect their contributions during the project phase, particularly when the division of intellectual property among consortium partners is ambiguous. The active treatment phase coefficient in the pooled model ($IRR = 0.887$, $p = 0.665$) points in the same direction, though it does not reach significance when the active phase years are pooled together. This suggests that the suppression of patent filing is concentrated at the moment of entry into the consortium rather than persisting throughout the project duration.

An alternative interpretation is that the $k = 0$ result reflects a reallocation of managerial attention rather than IP coordination costs. In the year a firm receives a grant, its R&D team shifts focus toward setting up the funded project: hiring, configuring equipment, initiating collaboration with partners, and meeting the administrative requirements of the funding agency. This reallocation may temporarily crowd out the time and attention required to prosecute patent applications already in the pipeline. Under this interpretation, the effect is a

short-run disruption rather than a structural feature of collaborative R&D. The fact that the coefficients at $k = +1$ through $k = +3$ revert to insignificance is consistent with both interpretations, but the reallocation account would predict a compensatory increase in patent filings once the project is established, which is not observed.

A third possibility is that the $k = 0$ coefficient captures Ashenfelter's dip, a well-known phenomenon in program evaluation where treated units experience a temporary decline in outcomes just before or at treatment, which is what motivates their program participation (Ashenfelter, 1978). If some firms applied for KMU-innovativ funding partly because their patenting activity had declined, the $k = 0$ effect could reflect this self-selection rather than a causal impact of the program. However, the parallel trends test provides evidence against this account: the pre-treatment coefficients at $k = -4$, $k = -3$, and $k = -2$ are all insignificant and show no systematic downward trend toward $k = -1$, which is the pattern Ashenfelter's dip would produce. The absence of a pre-treatment decline makes the collaboration cost interpretation more plausible than the self-selection account.

5.3. Patenting Behavior Among Treated Firms

The overall null result raises a secondary question: why do some treated firms patent at all? Of the 96 treated firms, 29 contributed patent filings in the post-treatment period, accounting for 371 total patents. Two firms alone, Werth Messtechnik (109 patents) and Laser Zentrum Hannover (41 patents), account for approximately 40% of this total. The remaining 27 firms contributed patent filings at substantially lower volumes. Understanding what distinguishes these firms from the majority that never patent is informative for the interpretation of the null result.

For the minority of treated firms that do patent, the primary motivation may be less about excluding competitors from using a specific technology and more about signaling technical competence to potential investors, OEM partners, or acquirers. Hughes and Mina (2010) and

Holgersson (2013) document that patents serve an increasingly important signaling function for high-tech SMEs seeking venture capital or strategic partnerships. In the German automotive supply chain, where OEM qualification processes increasingly require demonstrated IP capability, a patent portfolio may function as a credential rather than as a competitive advantage. This signaling motive is unrelated to the question of whether the subsidy caused additional patenting, and its presence among a subset of treated firms does not contradict the null finding.

The concentration of post-treatment patents among a small number of firms also illustrates a broader point about the distributional properties of patent data in SME populations. The two highest-volume filers contribute 40% of the total volume, meaning that the point estimates in the PPML model are sensitive to the behavior of a handful of firms. This is not a flaw in the estimation; it is a structural feature of patent activity in any small-firm sample. It does, however, limit the statistical power of any patent-based evaluation and underscores the practical difficulty of detecting moderate treatment effects in populations with highly skewed outcome distributions.

5.4. The Metric Mismatch: Synthesizing the Evidence

The results of this thesis, taken together, point to a coherent conclusion that extends beyond the specific question of whether KMU-innovativ affects patent output. The combination of a clean null result in the main specification, the structural evidence on firm size and patenting propensity, and the noisy placebo test constitutes a case study in metric mismatch: the application of an outcome measure to a population for which that measure is structurally inappropriate.

Three pieces of evidence support this interpretation. First, the logistic regression establishes that 86 of 96 treated firms fall below the 150-employee threshold at which patenting probability first crosses 50% in this R&D-active sample. For these firms, the absence of

patent output is the expected baseline regardless of their innovation activity. Second, the placebo test reveals that even among untreated control firms, random assignment of pseudo-treatment dates generates a large post-period coefficient (1.084, $p = 0.050$). A permutation test with 500 random reassignments confirms that this is not an isolated fluke: 6.6% of random draws produce a coefficient of this magnitude or larger (permutation $p = 0.066$). This indicates that patent counts are sufficiently noisy in this population to produce spurious effects of substantial magnitude, which in turn means that the null result for treated firms is not evidence of a well-powered test failing to detect an effect. Rather, it is evidence that the test itself lacks the statistical sensitivity to distinguish a moderate program effect from background noise. Third, the consistency analysis shows that the patenting behavior of both treated and control firms is characterized by irregular, occasional filing rather than systematic activity, which is exactly the pattern that makes patent counts an unreliable metric for program evaluation.

There is an unavoidable tension in this argument. The PPML estimation can only be run on the subset of firms that actually file patents, which means the causal estimate describes the program's effect on its least typical recipients. The 29 treated firms that contribute post-treatment filings are disproportionately large and IP-active relative to the full treated group. The estimation and the metric mismatch argument therefore operate on different populations: the argument applies most strongly to the 67 treated firms that never patent, while the estimation can only speak to the 29 that do. These two pieces of evidence do not simply sit side by side. They pull against each other, and the honest position is that the thesis cannot resolve that tension with the available data. The null result holds for the patenting subsample. Whether the program affected the non-patenting majority through channels that patent counts cannot capture remains an open question.

A reader may reasonably ask why this thesis builds its entire empirical analysis around a metric it ultimately argues is inappropriate for the population under study. The answer is that the inadequacy of patent counts for this population is an empirical finding, not a starting assumption. It could not have been established without first conducting the patent-based evaluation and documenting where and why it breaks down. The logistic regression in Section 3.5.3, the PPML null result in Section 4.2.2, and the noisy placebo in Section 4.2.4 are all necessary steps in building the metric mismatch argument. An alternative design that used a different outcome variable from the outset would have sidestepped this tension but would also have forfeited the ability to demonstrate it empirically. The circularity is real, but it is productive: the thesis uses the standard evaluation metric to establish the limits of the standard evaluation metric.

The metric mismatch argument developed here is not a criticism of patent data per se. For populations of larger firms with established IP management functions, patent counts remain a valuable and widely used innovation indicator (Griliches, 1990). The argument is that the reliability of this indicator deteriorates sharply below a firm size threshold, and that the KMU-innovativ population falls predominantly below that threshold. Evaluating the program through patent counts is like measuring the effect of a nutrition intervention with a ruler marked in centimeters rather than millimeters. The instrument is too rough to detect changes of the magnitude that the program could plausibly produce.

5.5. Limitations

Several limitations qualify the interpretation of the results and define the boundaries of the conclusions that can be drawn.

First, the sample size constrains statistical power. The causal estimation ultimately rests on 104 firms in the pooled specification and 94 firms in the event-study specification, with only 29 unique treated firms contributing post-treatment patent filings. A genuine positive

treatment effect of moderate size may exist but remain undetectable in a sample of this scale. Formal power calculations for a specific effect size are not reported because the relevant effect size is unknown *ex ante*, though the minimum detectable effects reported in Section 4.2.1 illustrate the scale of the power constraint. The wide confidence intervals on the treatment coefficients, spanning from $IRR = 0.517$ to $IRR = 1.524$ for the active treatment phase, make clear that the data cannot rule out economically meaningful positive or negative effects. Relatedly, the formal sensitivity analysis proposed by Roth (2022) for assessing robustness to parallel trends violations was not implemented. While the pre-treatment coefficients are small and insignificant, the standard errors are large enough that bias-adjusted confidence intervals would be uninformative at this sample size.

Second, the PPML estimation is restricted to the 132 firms that filed at least one patent during the observation period, which discards 66% of the full matched sample. This restriction is necessary because the PPML estimator cannot identify firm fixed effects for firms with zero patents in every year. However, it means that the causal estimate is based on a subsample that is structurally different from the program's target population: the retained firms are on average larger and more IP-active than the firms that are excluded. This tension deserves to be stated plainly. The 29 treated firms that contribute post-treatment patent filings are almost certainly the largest, most IP-active, and least typical firms in the treated group. The treatment effect estimated by the PPML model therefore characterizes the program's impact on its most unusual recipients, not its average ones. The null result should be interpreted as the absence of a detectable effect among the subset of treated firms that patent at all, rather than as a statement about the full set of 96 treated firms. It cannot tell us what the program did for the other 67 treated firms that never filed a patent.

Third, the matching strategy, while producing well-balanced groups on observable characteristics, cannot eliminate unobservable differences between treated and control firms.

Firms that apply for and win competitive grants may differ in managerial quality, strategic orientation, or absorptive capacity in ways that are not captured by employee count and industry classification. The parallel trends test provides some reassurance that these unobservables do not generate differential pre-treatment trends, but it cannot guarantee that the parallel trends assumption holds in the post-treatment period.

Fourth, the patent data come from three different databases (PATSTAT, LENS.ORG, and DPMA), and where the counts disagree, the highest number is used. This is meant to avoid undercounting, but it cannot catch patents filed under subsidiary names, joint ventures, or parent companies that are not linked to the firm identifiers in the dataset. If there is a systematic gap between how firms are identified in the patent databases and how they appear in Orbis, this creates measurement error in the dependent variable. In a count data setting, that kind of error pushes the estimates toward zero.

Fifth, the observation window of three to four post-treatment years may be insufficient to capture the full effect of the program on patent output. The translation of R&D investment into patent filings involves substantial lags: the research itself typically requires one to three years, followed by the drafting, filing, and prosecution of patent applications, which adds another one to two years. For a program with project durations of two to three years, the full patent response may only become visible five to seven years after treatment onset. The event-study coefficient at $k = +3$, while not significant, is the only positive post-treatment estimate (IRR = 1.310), which is at least consistent with the possibility that effects emerge at longer horizons.

Sixth, the study evaluates a single outcome metric. Patent counts capture only one dimension of innovation output, and the literature reviewed in Chapter 2 makes clear that small firms frequently innovate through channels that do not produce patents: process improvements, trade secrets, lead-time advantages, and tacit knowledge. The null result on patent

output cannot be interpreted as evidence that KMU-innovativ had no effect on firm-level innovation broadly defined. What it can establish is that no effect on the specific margin of formal patent filing is detectable. Evaluating the program against a broader set of innovation indicators, including R&D expenditure, product announcements, and survey-based innovation measures from instruments such as the Community Innovation Survey, would provide a more comprehensive assessment.

Seventh, the external validity of the results is limited to the automotive technology field within the KMU-innovativ program. The automotive sector has specific features, including long product development cycles, tiered supplier structures, and OEM-imposed IP requirements, that may not generalize to other technology fields funded under the same program. Whether the null result extends to KMU-innovativ projects in biotechnology, information technology, or other funded domains remains an open empirical question.

6. Conclusion

The question this thesis set out to answer is whether KMU-innovativ funding causes automotive SMEs to file more patents. Based on a matched panel of 387 firms and a staggered DiD design with PPML estimation, the answer is that no such effect is detectable in the patent data. Both the active and post-treatment coefficients are far from significant, and the event study shows no positive treatment effect at any horizon. The one result that reaches nominal significance is at $k = 0$, where treated firms filed fewer patents in the year they entered the program, though this does not survive correction for multiple comparisons. The Pregibon link test and a placebo test on the control group indicate that the estimation is correctly specified and that the null is not driven by functional form issues.

What does this thesis add to the literature? Most directly, it is the first study to estimate the causal effect of KMU-innovativ on patent output within a single technology sector, rather than pooling across the program's technology fields as previous evaluations have done. Beyond that, the analysis provides quantitative evidence for what is often only stated as a general caveat in the SME innovation literature: that patent counts lose their reliability as an innovation indicator below a certain firm size. The logistic regression puts that threshold at around 150 employees in this sample, and 86 of 96 treated firms fall below it. Finally, the event study turns up a suggestive finding at treatment onset, a significant drop in patent filings in the year firms enter the consortium, that is consistent with the IP coordination costs described by Hottenrott and Lopes-Bento (2016) and Clarysse et al. (2009). Whether this reflects a genuine collaboration cost or simply noise in a low-powered test cannot be settled with the present data.

For the evaluation of programs like KMU-innovativ, the implication is that patent counts are not well suited as an outcome metric for this population. When a program targets firms that are predominantly below 150 employees, and when the vast majority of those firms fall

below the threshold at which patenting becomes statistically regular, a patent-based evaluation is testing the sensitivity of the metric as much as the effectiveness of the program. The clean null for treated firms, together with the noisy placebo among controls ($p = 0.050$), suggests that the outcome variable simply does not have enough signal in it to separate a moderate program effect from background variation.

The Oslo Manual (OECD/Eurostat, 2018) points toward two better approaches. First, behavioral additionality: did the funded firms actually improve how they manage R&D? Did they build lasting partnerships? Did they change how they collaborate? These are things a program like KMU-innovativ is supposed to produce, and they can be measured through surveys and interviews with participants (Gök & Edler, 2012). Second, survey-based innovation indicators like the Community Innovation Survey can capture the share of revenue that comes from new or improved products. That is a much more direct measure of innovation output than counting patent filings. These approaches would provide a more appropriate and sensitive test of program effectiveness for the SME population targeted by KMU-innovativ.

Several directions for future research follow from the limitations discussed in Chapter 5. The most pressing is the extension of the observation window. The translation of R&D investment into patent filings involves substantial lags, and the event-study coefficient at $k = +3$, while not significant, is the only positive post-treatment estimate (IRR = 1.310). A study with access to patent data extending seven to ten years beyond treatment onset would provide a more definitive test of whether delayed effects exist. Second, future work should investigate whether the null result extends to other technology fields funded under KMU-innovativ. The automotive sector has specific structural features, including OEM-imposed IP requirements and long product development cycles, that may not generalize. A multi-

sector comparison would establish whether the metric mismatch documented here is universal across the program or specific to the automotive context.

Third, access to firm-level R&D expenditure data would enable a direct test of the input-to-output additionality gap. The Aschhoff et al. (2012) evaluation confirms input additionality at the program level, but matching firm-level R&D spending to patent output would reveal whether the firms that increase their R&D spending the most are the same firms that produce the most patent output, or whether the two margins are disconnected. Fourth, a mixed-methods approach combining quantitative patent analysis with qualitative case studies of funded firms would illuminate the mechanisms behind the null result. In particular, interviews with IP managers at treated firms could clarify whether the absence of patent filings reflects strategic decisions to rely on alternative appropriation mechanisms, capacity constraints in patent prosecution, or a genuine absence of patentable output.

If there is a broader takeaway from this analysis, it is about methodology rather than policy. A null result in program evaluation only tells you something useful if the outcome measure is actually capable of picking up the kind of effect the program might produce. When it is not, the null says more about the measurement tool than about the program. As public R&D funding increasingly targets small firms working in collaborative, early-stage settings, the evaluation methods need to keep up. Patent counts were developed as an innovation indicator for a different kind of firm. For the population that programs like KMU-innovativ actually serve, the field will need different metrics.

References

- Almus, M., & Czarnitzki, D. (2003). The Effects of Public R&D Subsidies on Firms' Innovation Activities: The Case of Eastern Germany. *Journal of Business & Economic Statistics*, 21(2), 226-236. <http://www.jstor.org/stable/1392458>
- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. <https://doi.org/10.2307/j.ctvcem4j72>
- Arrow, K. J. (1962). The Economic Implications of Learning by Doing. *The Review of Economic Studies*, 29(3), 155-173. <https://doi.org/10.2307/2295952>
- Arundel, A., & Kabla, I. (1998). What percentage of innovations are patented? empirical estimates for European firms. *Research Policy*, 27(2), 127-141 [https://doi.org/10.1016/S0048-7333\(98\)00033-X](https://doi.org/10.1016/S0048-7333(98)00033-X)
- Aschhoff, B., Astor, M., Crass, D., Eckert, T., Heinrich, S., Licht, G., Rammer, C., Riesen-berg, D., Rüffer, N., Strohmeier, R., Tonoyan, V., & Woywode, M. (2012). *Systemevaluierung "KMU-innovativ": Abschlussbericht; Begleit- und Wirkungsforschung zur Hightech-Strategie* (ZEW-Dokumentation No. 12-04). Zentrum für Europäische Wirtschaftsforschung. <https://hdl.handle.net/10419/59003>
- Ashenfelter, O. (1978). Estimating the Effect of Training Programs on Earnings. *The Review of Economics and Statistics*, 60(1), 47-57. <https://www.jstor.org/stable/1924332>
- Athreye, S.S., Fassio, C. & Roper, S. (2021). Small firms and patenting revisited. *Small Business Economics*, 57(2), 513–530. <https://doi.org/10.1007/s11187-020-00323-1>
- Beck, T., & Demirguc-Kunt, A. (2006). Small and medium-size enterprises: Access to finance as a growth constraint. *Journal of Banking & Finance*, 30(11), 2931-2943. <https://doi.org/10.1016/j.jbankfin.2006.05.009>

- Bellemare, M. F., & Wichman, C. J. (2020). Elasticities and the Inverse Hyperbolic Sine Transformation. *Oxford Bulletin of Economics and Statistics*, 82(1), 50-61.
<https://doi.org/10.1111/obes.12325>
- Bertrand, M., Duflo, E., Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, 119(1), 249-275.
<https://doi.org/10.1162/003355304772839588>
- Brouwer, E., & Kleinknecht, A. (1999). Innovative output, and a firm's propensity to patent.: An exploration of CIS micro data. *Research Policy* 28(6) 615-624.
[https://doi.org/10.1016/S0048-7333\(99\)00003-7](https://doi.org/10.1016/S0048-7333(99)00003-7)
- Bundesamt für Kartographie und Geodäsie. (2026). Verwaltungsgebiete 1:250 000 aktualisiert. *BKG Wir geben Orientierung*. https://www.bkg.bund.de/SharedDocs/Produktinformationen/BKG/DE/P-2025/250911_VG250.html
- Bundesministerium für Bildung und Forschung. (2024). Änderung der Richtlinie zur Förderung von Zuwendungen für "KMU-innovativ." Bundesanzeiger
<https://www.bundesanzeiger.de/pub/publication/8pGbhrF4Htzq3Dw59on/content/8pGbhrF4Htzq3Dw59on/BAnz%20AT%2009.08.2024%20B4.pdf?inline>
- Bundesministerium für Forschung, Technologie und Raumfahrt. (2026). Projekte.
<https://www.zukunft-der-wertschoepfung.de/projekte/>
- Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, 225(2), 200-230.
<https://doi.org/10.1016/j.jeconom.2020.12.001>
- Clarysse, B., Wright, M., Mustar, P. (2009). Behavioural additionality of R&D subsidies: A learning perspective. *Research Policy*, 38(10), 1517-1533.
<https://doi.org/10.1016/j.respol.2009.09.003>

- Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1), 128-152. <https://doi.org/10.2307/2393553>
- Commission Regulation (EU). (2014). Declaring certain categories of aid compatible with the internal market in application of Articles 107 and 108 of the Treaty. *Official Journal of the European Union*, L 187, No 651, 1-78. <https://eur-lex.europa.eu/eli/reg/2014/651/oj>
- Correia, S., Guimaraes, P., Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. *The Stata Journal*, 20(1), 95-115. <https://doi.org/10.1177/1536867X20909691>
- Czarnitzki, D., & Licht, G. (2006). Additionality of public R&D grants in a transition economy: The case of Eastern Germany. *Economics of Transition*, 14(1), 101-131. <https://doi.org/10.1111/j.1468-0351.2006.00236.x>
- David, P. A., Hall, B. H., Toole, A. A. (2000). Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. *Research Policy*, 29(4-5), 497-529. [https://doi.org/10.1016/S0048-7333\(99\)00087-6](https://doi.org/10.1016/S0048-7333(99)00087-6)
- Deutsches Patent- und Markenamt. (2025). Publikationen und Register für Patente, Gebrauchsmuster, Marken und Designs. *DPMAregister*. <https://www.dpma.de>
- European Patent Office. (2025). PATSTAT. *EPO*. <https://www.epo.org/en/searching-for-patents/business/patstat>
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254-277. <https://doi.org/10.1016/j.jeconom.2021.03.014>

- Gök, A., & Edler, J. (2012) The use of behavioural additionality evaluation in innovation policy making. *Research Evaluation*, 21, 306-318
<https://doi.org/10.1093/reseval/rvs015>
- Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, 28(4), 1661-1707. <http://www.jstor.org/stable/2727442>
- Hall, B. H., & Lerner, J. (2009). The financing of r&d and innovation. *National Bureau of Economic Research Working Paper 15325*. <https://doi.org/10.3386/w15325>
- Hayter, C. S., Link, A. N., Scott, J. T. (2018). Public-sector entrepreneurship. *Oxford Review of Economic Policy*, 34(4), 676-694 <https://doi.org/10.1093/oxrep/gry014>
- Holgersson, M. (2012). Innovation and Intellectual Property: Strategic IP Management and Economics of Technology. *PhD thesis, Chalmers University of Technology*, <https://ssrn.com/abstract=2563297>
- Holgersson, M. (2013). Patent management in entrepreneurial SMEs: a literature review and an empirical study of innovation appropriation, patent propensity, and motives. *R&D Management*, 43(1), 21-36. <https://doi.org/10.1111/j.1467-9310.2012.00700.x>
- Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, 81(396), 945-960. <https://doi.org/10.2307/2289064>
- Hottenrott, H., & Lopes-Bento, C. (2014). (International) R&D collaboration and SMEs: The effectiveness of targeted public R&D support schemes. *Research Policy*, 43(6), 1055-1066. <https://doi.org/10.1016/j.respol.2014.01.004>
- Hottenrott, H. and Lopes-Bento, C. (2016). R&D Partnerships and Innovation Performance: Can There Be too Much of a Good Thing?. *Journal of Product Innovation Management*, 33(6), 773-794. <https://doi.org/10.1111/jpim.12311>
- Hughes, A., & Mina, A. (2010). *The impact of the patent system on SMEs* (Working Paper 411). Centre for Business Research, University of Cambridge.

https://www.researchgate.net/publication/228977675_The_Impact_of_the_Patent_System_on_SMEs

Krzywdzinski, M., Lechowski, G., Ferdinand, J., & Schnei, D. (2023). The german path to electromobility and its impacts on automotive production and employment. *On the way to electromobility - a green (er) but more unequal future*, 179-206.

LENS. (2025). Explore global science and technology knowledge. *LENS.ORG*.
<https://www.lens.org>

Link, A. N., Scott, J. T. (2018). Propensity to Patent and Firm Size for Small R&D-Intensive Firms. *Review Industrial Organization*, 52(4), 561-587.
<https://doi.org/10.1007/s11151-018-9617-0>

Moody's. (2025). Orbis Database: Global Private Company Data & Intelligence. *Moody's Corporation*. <https://www.moody's.com/web/en/us/capabilities/company-reference-data/orbis.html>

OECD/Eurostat. (2018). *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation (4th ed.)*. The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing.
<https://doi.org/10.1787/9789264304604-en>.

Pregibon, D. (1980). Goodness of Link Tests for Generalized Linear Models. *Journal of the Royal Statistical Society*, 29(1), 15-23. <https://doi.org/10.2307/2346405>

Roth, J. (2022). Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends. *American Economic Review: Insights*, 4(3), 305-322.
<https://doi.org/10.1257/aeri.20210236>

Rubin, D. B. (1974). Bayesian Inference for Causal Effects: The Role of Randomization. *The Annals of Statistics*, 6(1), 34-58. <https://www.jstor.org/stable/2958688>

- Santos Silva, J. M. C., & Tenreyro, S. (2006) The Log of Gravity. *The Review of Economics and Statistics*, 88 (4), 641–658. <https://doi.org/10.1162/rest.88.4.641>
- Schmidheiny, K., & Sieglöcher, S. (2023). On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization. *Journal of Applied Econometrics*, 38(5), 695-713. <https://doi.org/10.1002/jae.2971>
- Simon, H. (1996). You Don't Have to be German to be a "Hidden Champion". *Business Strategy Review*, 7(2), 1-13. <https://doi.org/10.1111/j.1467-8616.1996.tb00118.x>
- Verband der Automobilindustrie. (2024). On the move. VDA.
<https://www.vda.de/en/topics/automotive-industry>
- Verband der Automobilindustrie. (2021) Clean. Safe. Digital. Innovations in 2021. *VDA*.
<https://www.vda.de/en/topics/innovations/innovations-2021>
- Wu, A. (2017). The signal effect of Government R&D Subsidies in China: Does ownership matter?. *Technological Forecasting and Social Change* 117, 339-345.
<https://doi.org/10.1016/j.techfore.2016.08.033>
- Zúñiga-Vicente, J. Á., Alonso-Borrego, C., Forcadell, F. J., Galán, J. I. (2012). Assessing the effect of public subsidies on firm r&d investment: A survey. *Journal of Economic Surveys* 28(1) 36-67 <https://doi.org/10.1111/j.1467-6419.2012.00738.x>

Appendix

Appendix A:

Table A1

Pairwise Correlations: Full Sample (N=387)

	Total Patents	Baseline Emp.	Treated
Total Patents	1.000		
Baseline Employees	0.258*	1.000	
Treated	0.036	-0.005	1.000

Note. denotes significance at the 5% level.

Table A2

Pairwise Correlations: Treated Group (N=96)

	Total Patents	Baseline Emp.	Project Partners
Total Patents	1.000		
Baseline Employees	0.444*	1.000	
Project Partners	-0.212*	-0.002	1.000

Note. denotes significance at the 5% level.

Table A3

Pairwise Correlations: Control Group (N=291)

	Total Patents	Baseline Emp.
Total Patents	1.000	
Baseline Employees	0.202*	1.000

Note. denotes significance at the 5% level.

Across the full sample, firm size is the only significant predictor of total patent output ($r = 0.258$). Treatment status is essentially uncorrelated with total patents ($r = 0.036$) and with firm size ($r = -0.005$), confirming that the matching procedure produced groups with no pre-existing innovation advantage. Within the treated group, the size-patent correlation strengthens to 0.444, compared to 0.202 within the control group, consistent with positive selection into the program.

Appendix B:

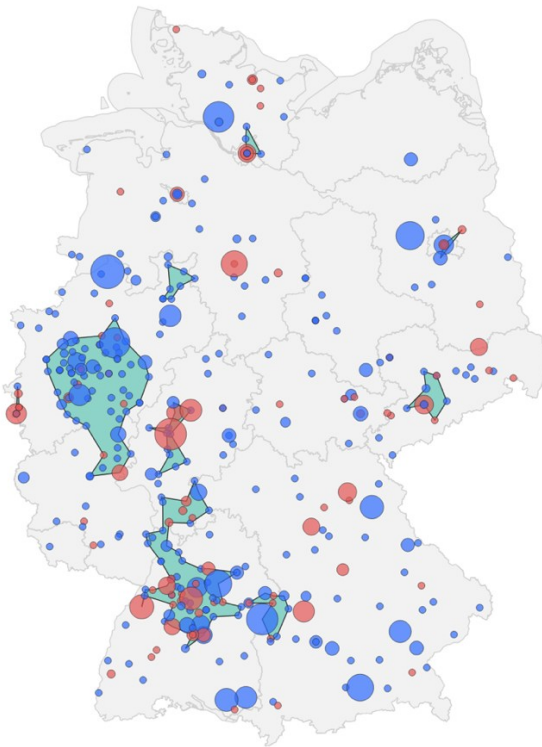
Table B1

Balance Statistics of the Patenting Subsample

Variable	Treated (N=45)	Control (N=87)	Diff.	t-stat	p-value
Baseline Em- ployees	81.2	93.0	-11.8	-1.00	0.319
Competition Index	20.2	21.1	-0.91	-0.50	0.615
Pre-Trend Patents	4.1	4.3	-0.14	-0.07	0.945

Note. Restricted to firms with at least one patent filing across the observation period.

Figure B



Geospatial Distribution of the Matched Sample

Note. Red markers indicate treated firms; blue markers indicate control firms.

Table B2*Geographic Distribution and Patent Output*

Test	Statistic	p-value	Interpretation
State-level effect on patents	ANOVA $F = 0.44$	0.966	No significant relationship
Treated-control geographic balance	Pearson chi-squared = 13.70	0.548	Proportional distribution

Note. The 387 firms are distributed across 16 German federal states. Nordrhein-Westfalen (92 firms, 23.8%), Baden-Württemberg (75 firms, 19.4%), and Bayern (50 firms, 12.9%) account for approximately 56% of the sample.

Appendix C:**Table C***Patenting Consistency by Treatment Status*

	Control (N=291)	Treated (N=96)	Diff.	p-value
Pre-Treatment Consistency	0.063	0.142	-0.078	0.002
Post-Treatment Consistency	0.082	0.138	-0.055	0.044
Consistency Change	0.019	-0.006	0.023	0.269

Note. Consistency is the proportion of years within the respective five-year window in which the firm filed at least one patent. Full sample of 387 firms. The change in consistency from pre- to post-treatment is not significantly different between treated and control firms, consistent with positive selection rather than a treatment effect on filing regularity.

Appendix D:

Table D1

Characterization of the Treated Sample by Collaboration Type

Project Type	Firm Size (Employees)	Network (Partners)	Size	Pre-Treatment Patents (t-4 to t-1)	N
Industry-Only	69.2	3.1		2.2	78
University-Linked	53.0	2.7		0.6	18
Total Treated	66.2	3.0		1.93	96

Note. Differences between groups across all three metrics are not statistically significant.

Table D2

Partner Count and Patent Output: Key Correlations (Treated Group, N=96)

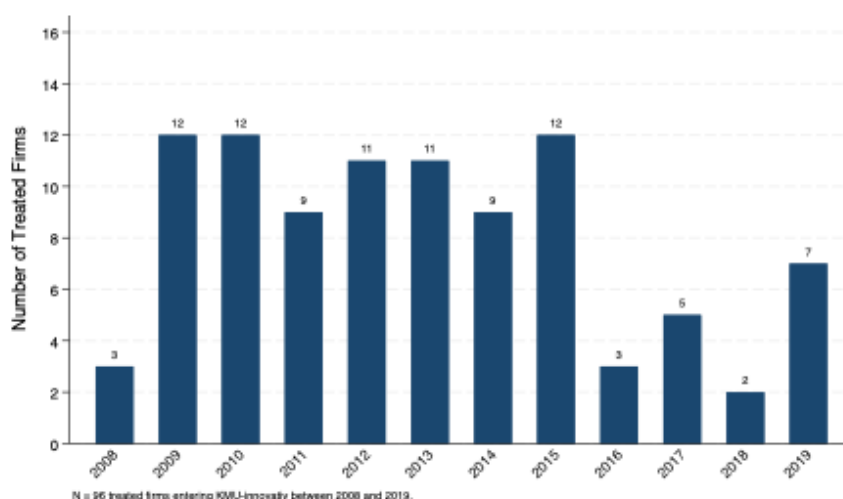
Correlation	r	p-value
Partner count vs. pre-treatment patents	-0.230	< 0.05
Partner count vs. total lifetime patents	-0.212	< 0.05
Partner count vs. baseline employees	-0.002	n.s.

Note. The negative relationship between partner count and patent output is independent of firm size, suggesting that IP coordination costs in larger consortia dampen the conversion of collaborative R&D into formal patent filings.

Appendix E:

Figure E

Distribution of Treatment Cohorts



Note. N = 96 treated firms entering KMU-innovativ between 2008 and 2019. Each bar shows the number of firms beginning their funded project in that calendar year.

The staggered entry of treated firms into the KMU-innovativ program is a defining feature of the identification strategy. Figure E shows that treatment timing is spread across the full 2008 to 2019 window, with no single cohort dominating the sample. This distribution motivates the event-time specification used in Chapter 4, which measures all firms on a common relative timeline rather than in calendar years. The absence of extreme clustering in any single year reduces the risk that the treatment effect estimate is driven by a specific macroeconomic shock coinciding with a large cohort's entry.

AI Tools Disclaimer

This thesis was developed with support from artificial intelligence tools at multiple stages of the research and writing process. The following declaration describes how these tools were used, what role they played, and where the boundaries of my own contribution lie.

Tools Used:

ChatGPT (OpenAI): Used in the early stages for brainstorming, exploring potential research angles, and generating initial ideas about how to frame the research question and structure the analysis.

Claude (Anthropic): Used for structuring and rewriting, code generation, and iterative refinement of the thesis text. This included generating Stata do-files based on my ideas, rewriting my draft sections, structuring arguments, organizing notes on the academic literature, identifying connections between sources, checking numerical consistency between tables and prose such as grammatical errors or misspelling.

What I Did

The intellectual direction of this thesis is my own. Specifically:

- I identified the research question, selected the KMU-innovativ program as the object of evaluation, and chose the automotive sector as the empirical focus.
- I collected and assembled the dataset manually, including extracting project records from the BMBF Förderkatalog, identifying and matching control firms from Orbis, and triangulating patent data across three databases (PATSTAT, LENS.ORG, DPMA).
- I designed the empirical strategy, including the choice of a staggered Difference-in-Differences framework with PPML estimation, the definition of treatment phases and observation windows, and the selection of robustness checks.

- I ran all Stata estimations myself. The do-files were developed in collaboration with AI (primarily Claude), but every regression was executed, reviewed, and validated by me in Stata on my laptop.
- I decided which results to include, how to interpret them, and how to frame the null result. The metric mismatch argument, the collaboration paradox interpretation, and the firm size threshold analysis reflect my own reading of the literature and my judgment about what the data show.
- I pre-wrote draft versions of the chapters, which were then refined and rewritten with AI assistance. The structure, argumentation, and analytical choices in every chapter originate from me.
- I conducted a complete proofreading pass in which I corrected factual errors, removed passages that did not fit, adjusted the tone, and ensured consistency between the text and the empirical output.

What AI Did

AI tools played a role in the following areas:

- Literature processing: Organizing notes on the academic literature and suggesting relevant references I had not yet considered.
- Code generation: Writing and debugging Stata do-files for data preparation, estimation, verification of results, and figure generation. The logic behind each specification was mine, but the syntactic implementation was AI-assisted.
- Writing and rewriting: Rewriting my drafts and notes, such as restructuring paragraphs for clarity.
- Numerical verification: Cross-checking every empirical claim in the thesis against Stata output, identifying discrepancies between tables and prose, and flagging inconsistencies that I then corrected.

Boundaries

AI did not make analytical decisions. It did not choose which hypotheses to test, which specifications to run, how to interpret ambiguous results, or what to conclude. When AI-generated text contained claims I disagreed with or framings I found inaccurate, I revised or discarded them. The final text reflects my understanding of the literature, my interpretation of the results, and my judgment about what this thesis contributes.

Alignment with University of Vienna Guidelines

This declaration follows the University of Vienna's guidelines on the use of artificial intelligence in teaching and scientific work (Guidelines der Universität Wien zum Umgang mit Künstlicher Intelligenz (KI) in der Lehre, 2nd edition, September 2024). The guidelines require that the use of AI tools in final theses be made transparent and comprehensible, specifying which tools were used, how, to what extent, and for what purpose. This declaration addresses each of these points in the sections above.

The guidelines further require that student work maintain a sufficient level of original intellectual achievement (Schöpfungshöhe) so that it is not classified as plagiarism or ghostwriting. The research question, the empirical design, the data collection, the interpretation of results, and the analytical conclusions of this thesis are entirely my own work. AI tools were used to support and optimize the working process, not to replace the intellectual contribution. The distinction between my own contribution and AI-assisted elements is documented in the sections "What I Did" and "What AI Did" above.

No personal data of third parties was entered into any AI tool during the course of this project.