How Finance Shapes Careers of Highly Skilled Individuals

Inaugural Dissertation
to Obtain the Academic Degree of
a Doctor in Business Administration
at the University of Mannheim

submitted by

Clemens Mueller

Paris, France

Dean: *Joachim Lutz*
Referent: *Prof. Ernst Maug, PhD*
Correferent: *Prof. Dr. Alexandra Niessen-Ruenzi*

Day of oral examination: May 25th 2023

für meine Familie

Acknowledgements

I was and I continue to be fortunate in life. I owe a lot to many people I was lucky to cross paths with. In the following, I want to thank especially those who supported me during my time as a PhD student in Mannheim.

Since 2015, I worked as a student assistant for Prof. Erik Theissen, who was not just great at teaching investments and empirical finance and fostered my interest in academic research, but also provided guidance and assistance during my first two years as a PhD student. I thank Prof. Peter Albrecht for supporting me and giving me the opportunity to teach financial mathematics, which has been a privilege to me. To this, I owe the second chapter in this dissertation. I want to thank my advisor Prof. Ernst Maug for continued guidance, advice, crucial feedback, and for trust and freedom to pursue my research agenda. I feel fortunate to have had Ernst as my mentor and the quality and rigorousness of my work would not be the same without him. I thank my second advisor Prof. Alexandra-Niessen Ruenzi, who was not just a great supervisor who supported my research and teaching. Alexandra also offered advice and always pushed me to strive for more high quality and policy relevant research.

I thank Johan Hombert for the invitation to visit the finance department at HEC Paris, which has been an invaluable experience for me. Bernhard Ganglmair, thank you for support and for teaching me the economics of innovation. I am indebted to Jiri Tresl and Markus Huggenberger who provided advice and emotional support, and I also thank all doctoral students in the finance area. I always appreciated the support of all other team members in the finance area such as Alex, Angelika, Bettina, Marion and Mesude as well as on the university level Sophie and Filomena, and many people in the IT department. My PhD time was characterized by the best GESS CDSB cohort in existence. Thank you Alison, Amir, Hala, Hamed, Hossein, Marius, Matthew, Mengnan, Mohammad, Theresa, and Yen, as well as every member of the Mensa and hiking group. I owe special thanks to Santanu, who is not just a friend, but also co-author and with whom I was lucky to work closely in the last years. I owe a lot to Dani, Kilian, and Lukas, as without you, Mannheim would not have been the same.

I thank my home in Mannheim, Ali and Musti. Ali, you are my inspiration and you are teaching me something every single day. You are the person with the biggest heart who I have ever met and your love towards others has no boundaries. I will forever admire you.

I am indebted to my family, who provided unconditional support throughout my life. Thank you Sigrid, Michael, Lukas and Oma und Opa for everything. This dissertation is dedicated to you.

Table of Contents

Acknowledgements . . . . **iv**

Table of Contents . . . . **vi**

I. Introduction **1**

II. Reacting to Failing Finance in University **8**

2.1. Introduction 10

2.2. Institutional background and data 15

2.2.1. Institutional background 15

2.2.2. Data 16

2.2.3. Descriptive statistics 18

2.2.4. Survey data 21

2.3. Regression discontinuity evidence 21

2.3.1. Baseline regression discontinuity scatterplot 21

2.3.2. Are students able to manipulate the running variable? 23

2.3.3. Are students on either side comparable? 28

2.3.4. Baseline regression 32

2.4. What explains the reaction of male students? 36

2.4.1. Heterogeneity 36

2.4.2. Expectations, overconfidence, and failure 37
2.4.3. Competitiveness 41
2.4.4. Retake behavior 44
2.5. Conclusion 46
2.6. Appendix 48
References 53

III. Non-Compete Agreements and Labor Allocation Across Product Markets 56

3.1. Introduction 58
3.2. Data 63
3.2.1. Employment Histories of Corporate Inventors 63
3.2.2. Institutional Details and Data on Enforcement Changes 64
3.2.3. Sample Construction and Descriptive Statistics 67
3.3. Staggered State-Level Changes in Non-Compete Enforcement 70
3.3.1. Event Study and Dynamic Effects 70
3.3.2. Non-Compete Agreements and Product Market Similarity 73
3.3.3. Within State-Year: Is the Effect Stronger in the Presence of Non-Compete Agreements? 75
3.3.4. Does Increased Non-Compete Enforcement *Cause* Industry Mobility? 77
3.3.5. Heterogeneity: Outside Opportunities 78
3.3.6. Inventors move to Employers which rely less on NCAs 80
3.4. NCA-Constrained Industry Moves Lead to Lower Productivity 82
3.5. Channels 85
3.5.1. NCA Enforcement leads to Worse Inventor-Firm Matching Quality 85

3.5.2. Non-compete Agreement Enforcement leads to Longer Employment Gaps 87
3.6. Industry Mobility and Productivity Across Product and Technology Markets 89
3.7. Firm-Level Productivity Regressions 92
3.8. Conclusion 93
3.9. Appendix 95
References 102

IV. Angels and Demons 108
4.1. Introduction 110
4.2. Data 115
4.2.1. Angel Employees 115
4.2.2. Employment Histories 115
4.2.3. Sample Construction and Independent Variables 116
4.2.4. Innovation Output 118
4.2.5. Other Control Variables 120
4.2.6. Descriptive Statistics 120
4.3. Empirical Results 125
4.3.1. Baseline Panel Regression 125
4.3.2. Within Firm-year: Angel Employees and Innovation Output Across States 127
4.3.3. Within Firm-year: Event Study and Dynamic Effects 130
4.3.4. Instrumental Variable Regression: Competition from Venture Capital 133

4.4. Channels . . . 137

4.4.1. Agency conflicts . . . 138

4.4.2. Angel Employee Exit: Loss of Human Capital . . . 144

4.5. Angel Employees are Beneficial for Start-ups . . . 145

4.6. Robustness . . . 149

4.6.1. Evidence from Angel Roles: Innovation-related Angels . . . 149

4.6.2. Non-Patent-Based Measures of Innovation . . . 151

4.6.3. Private Firms . . . 153

4.6.4. Excluding Recent IPO Years . . . 154

4.6.5. Outsourcing . . . 156

4.6.6. Welfare Analysis . . . 156

4.7. Conclusion . . . 158

4.8. Appendix . . . 160

References . . . 168

Curriculum Vitae . . . 174

Chapter I

Introduction

Human capital is an essential driver to maximize the well-being of all individuals on the planet. A large share of differential growth rates globally is due to human capital.[1,2,3] Human capital is also a crucial input for innovation, one specific first order driver of economic prosperity.[4] In the absence of perfect capital versus labor substitution (e.g. improved artificial intelligence), the importance of human capital for society is likely to grow.

Human capital in the form of highly skilled individuals will be at the core of each chapter in this dissertation. I focus on individuals and how their careers interact with finance. Within these two broad topics, I specifically analyze causes and consequences of individuals' choices in relation to finance. In the following, I focus on important and strategic decisions of individuals which shape their lives and careers.

[1]Barro (1991), Economic Growth in a Cross Section of Countries. Quarterly Journal of Economics, Vol. 106, No. 2, pages 407-443.

[2]Acemoglu (2009), Introduction to Modern Economic Growth, Princeton University Press.

[3]Aghion, P., and & Howitt, P.W. (2009), The Economics of Growth, MIT Press.

[4]Romer (1990), Endogenous Technological Change, Journal of Political Economy, Vol. 98, No. 5, pages 71-102.

This introduction will in the following chronologically track careers of highly skilled individuals. For each chapter, I describe which career stage I analyze, I motivate why the strategic choice I analyze is meaningful, and then I sketch an empirical design, provide visual evidence, and finally conclude.

Chapter 2 starts at an early career stage of an individual, financial education in university. Educational attainment, e.g. whether to complete university, is a choice with potentially large impacts on career trajectories. The positive (signalling) effects of obtaining a high school or university degree are well documented.[5,6,7,8]

My research question is whether students decide to drop out of university when they fail their very first finance exam which, in the setting I study, happens to be students' very first university exam. I estimate causal effects by comparing the following two groups of students: those just marginally above to those just marginally below the passing threshold of their very first university exam. The assumption underlying this framework is that these two groups of students are virtually identical, however one group fails the exam and the other does not.

[5] Tyler, J.H., and Murnane, R.J., and Willett, J.B. (2000), Estimating the labor market signaling value of the GED. The Quarterly Journal of Economics 115 (2), pages 431–468.

[6] Canaan, S., and Mouganie, P. (2018), Returns to education quality for low-skilled students: Evidence from a discontinuity. Journal of Labor Economics 36 (2), pages 395–436.

[7] Clark, D., and Martorell P. (2014), The signaling value of a high school diploma. Journal of Political Economy 122 (2), pages 282–318.

[8] Machin, S., and McNally, S., and Ruiz-Valenzuela J. (2020), Entry through the narrow door: The costs of just failing high stakes exams. Journal of Public Economics 190, 104224.

Figure I.1.: Graduation Probability and Failing the First University Exam
The x-axis is the distance to the passing threshold in a finance exam, students very first exam in university. On the y-axis is the average university completion probability.

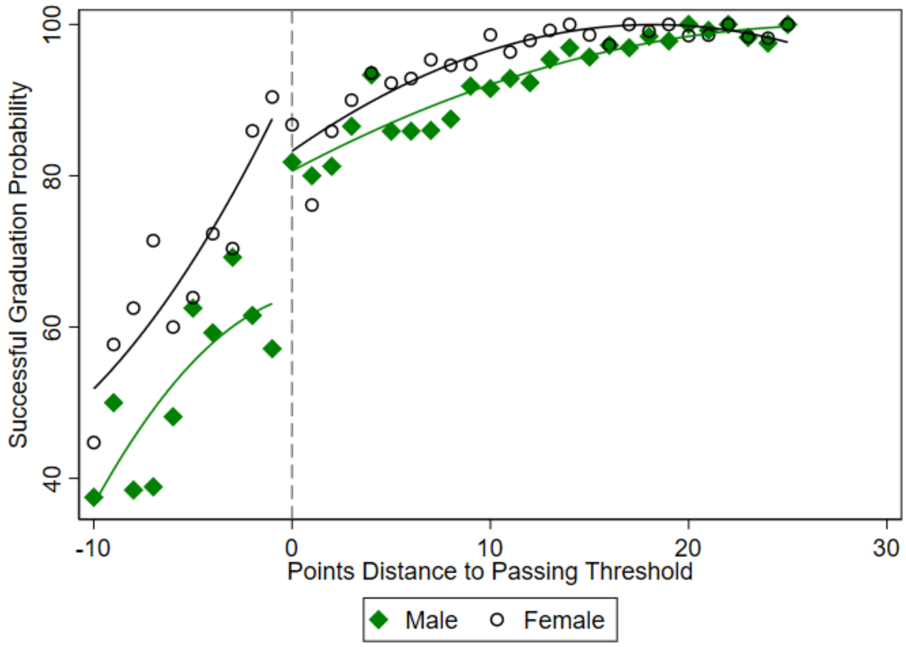


Figure I.1 shows the baseline effect. Female students are completely unaffected and more resilient. Male students who marginally fail their very first university exam are 15% less likely to successfully obtain a university degree. I complement the analysis with survey evidence. Differences in attitudes towards competitiveness and overconfidence between male and female students are potential mechanisms. Finance, in the form of financial education, thus shapes careers early onward, here in the form of experiencing failure and educational choices.

Chapter 3 analyzes individuals in their early/mid stage of their careers. I analyze employment choices, e.g. which employer to work for, which are decisions every individual faces. Which firm to work for and what to work has impacts on personal fulfillment, financial stability, as well as overall well-being. For the economy, efficient allocation of labor can lead to sizable efficiency gains.[9,10]

My research question is on non-compete agreements, which restrict individuals' within-industry employment choice set. I use state-level increases in non-compete agreement enforcement in a staggered difference-in-differences event time regression.

Figure I.2.: Non-Compete Agreements Cause Across-Industry Mobility
The x-axis is time, in years relative to an increase in non-compete agreement enforcement. The y-axis is the probability an employee moves to a new employer in a different industry.

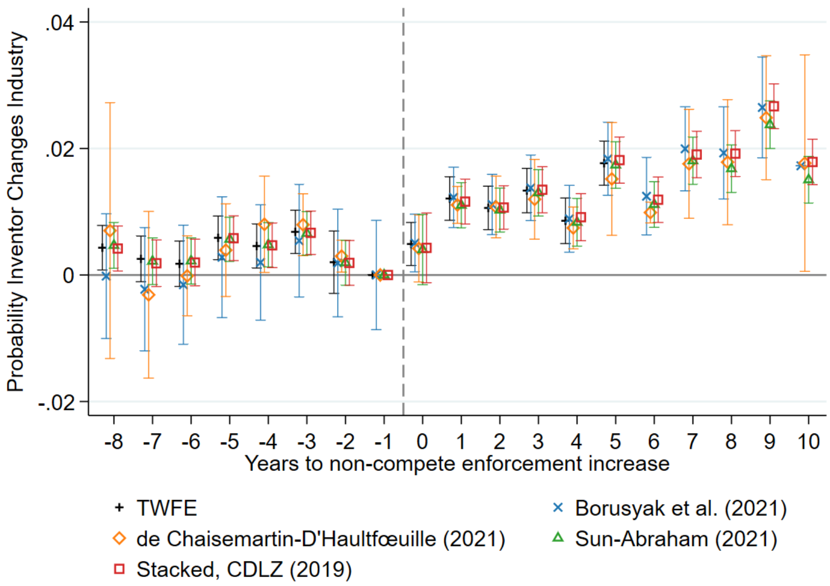


Figure I.2 shows the baseline effect. I show significant reallocation of human capital as non-compete agreements cause some individuals to move to a new employer in another industry. 1.5 out of 100 inventors move to a new employer in another industry *per year*. Reallocated individuals are subsequently less productive, as measured by their innova-

[9]Hsieh, C.T., Hurst, E., Jones, C.I., and Klenow, P.J. (2019), The Allocation of Talent and U.S. Economic Growth. Econometrica, 87: 1439-1474.

[10]Hsieh C.T., Klenow P.J. (2009), Misallocation and Manufacturing TFP in China and India, The Quarterly Journal of Economics, Volume 124, Issue 4, Pages 1403–1448.

tion output. Financial regulation, here in the form of restricting the within-industry employment choice set, thus has profound effects on career trajectories of individuals, as well as on their productivity.

Chapter 4 turns to seasoned industry professionals and their private investments. I analyze investment choices of individuals and how these decisions shape individuals' careers. I focus on the early-stage financing market, which is of importance from the point of view of investors because these investment choices are sizable and risky. Analyzing this market from a social planner point of view is important as innovative and radical ideas (e.g. to combat climate change) need risk-seeking investors to overcome the uncertain life cycle of an early-stage firm.[11,12]
My research question is on angel investments, which are personal equity investments in early-stage firms. Together with my coauthor Santanu Kundu, we analyze broader effects of angel investors in the economy. We empirically show that when employees invest their private wealth into startups, the innovation output of their employer declines.

[11] Lerner, J., and Ramana N. (2020), Venture Capital's Role in Financing Innovation: What We Know and How Much We Still Need to Learn. Journal of Economic Perspectives 34, no. 3: 237–261.
[12] Howell, S.T. (2017), Financing Innovation: Evidence from R&D Grants. The American Economic Review 107, no. 4: 1136–64.

Figure I.3.: Innovation Declines when Firms' Employees conduct Angel Investments
The x-axis is the time, in years relative to when employees start investing their personal wealth in early-stage firms. The y-axis is the economic value of patents of the employers.

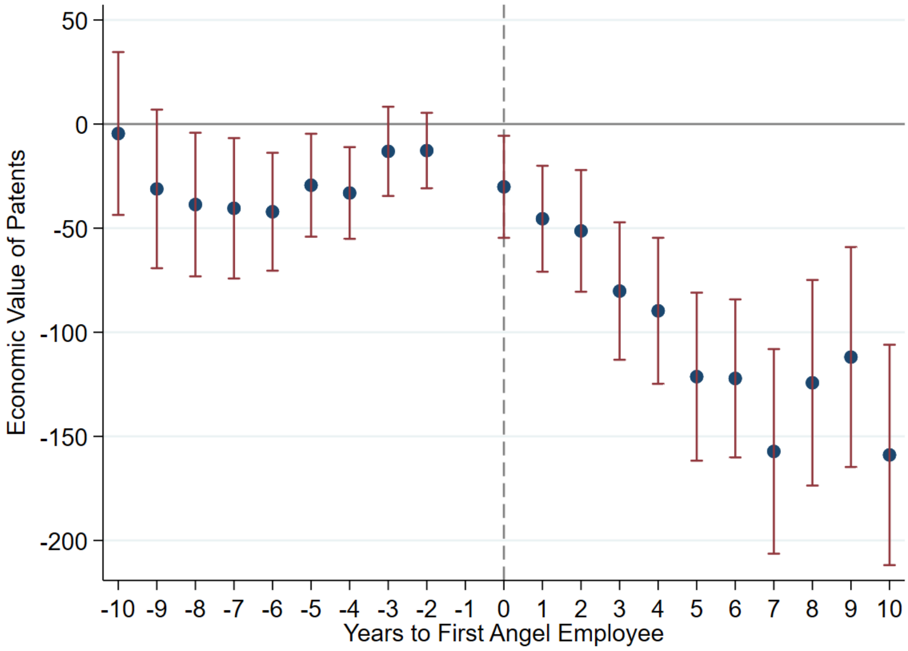


Figure I.3 shows the baseline effect. I analyze two novel channels for these results: The first channel is an agency conflict as angel investors trade-off exerting effort at their employer and their portfolio early-stage firm. The second channel is the loss of human capital for their employers, as these highly skilled employees are more likely to leave. On the positive side, early-stage firms benefit when they receive financing from seasoned industry professionals. I thus show that finance, here in the form of private investments by seasoned industry professionals, impacts their own careers and finances, but also has large impacts on their employers as well as their portfolio early-stage firms.

The chapters of this dissertation have now traversed across career stages of highly skilled individuals and their choices and interactions with finance. Individuals react to failing a finance exam in university, individuals' employment choices are affected by labor regulation, and seasoned individuals' private investments have broader effects on their employers and early-stage firms. In each, finance fundamentally affects lives of human beings, and thus also the lives of their friends and families, employers, and the economy.

Chapter II

Reacting to Early Failure in University: Evidence from a Regression Discontinuity Design

Reacting to Early Failure in University: Evidence from a Regression Discontinuity Design

Clemens Mueller[1]

Abstract

This paper investigates gender differences in persistence in educational attainment. I ask whether students successfully complete their university degrees when they suddenly experience failure in university, specifically, failing their very first university exam. I identify causal effects using university administration data and a sharp discontinuity at the passing threshold of the first university exam of 8,500 undergraduate students. Male students who marginally fail are 15% less likely to successfully obtain a university degree. Female students are completely unaffected and more resilient towards early failure in university. I add survey evidence to show that overconfidence and competitiveness explain the negative reaction of male students. I provide causal evidence of one reason why male students are less resilient in educational attainment: They have a strong negative reaction to early failure in university.

[1]Comments appreciated. Please contact Clemens Mueller, University of Mannheim, clemens.mueller@uni-mannheim.de, +49 621 181 1362. I would like to thank Vicki L. Bogan, Ernst Maug, Alexandra Niessen-Ruenzi, and Alison Schultz as well as seminar participants at the University of Mannheim for helpful comments.

2.1. Introduction

Women are more persistent and resilient when it comes to educational attainment. Female students are more likely to complete high school (Murnane, 2013) as well as university (Bailey and Dynarski, 2011; Goldin, Katz, and Kuziemko, 2006). Conventional economic models are unable to explain these differences between male and female students in educational attainment. Factors such as socioeconomic status and ethnicity do not vary as male and female students grow up in the same families and attend the same schools. This paper aims to provide causal evidence on one reason why male students are less likely to complete university: Male students have a strong negative reaction to failure.

In this paper, I ask whether students decide to drop out of university when they fail their very first university exam. I use a sharp regression discontinuity design and compare students who marginally fail to those who marginally pass their very first university exam. The analysis is based on administrative data and detailed records of the first university exam of around 8,500 students of a mid sized German university. The exam I analyze is an introduction to financial mathematics course, mandatory in the first semester for all economics, business, and law majors. The final grade is fully determined by an exam 6 weeks after semester start, which means that the course is generally students' very first university exam.

Female students marginally above and marginally below the passing threshold successfully obtain a university degree with a probability of 89%. Male students marginally above the passing threshold complete university with a probability of 80%. Male students marginally below the passing threshold only complete university with a probability of 65%. Thus, there is an economically large and statistically significant discontinuity of 15%. These results are robust to various functional form specifications, bandwidth selection, and other econometric choices.

The empirical results hold under the assumption that students in a local area around the passing threshold are similar to each other, except for assignment into treatment (failing their very first university exam). I argue that marginally passing, or failing the first university exam is to some extent randomly determined. Precise manipulation of selection into treatment is an unlikely explanation. Students might very well be able to aim for a certain number of points achieved in the exam. However, the precise passing threshold was unknown to students and course administrators alike before the exam. Grading was cross-sectional in nature such that 15% of the course failed the exam. Thus, the passing threshold was determined by the exam difficulty and competition in each respective cohort. The passing threshold varied across the years 2008-2018 as follows: {22.5, 17, 19, 20, 16.5, 16.5, 18.5, 18, 15, 16.5, 20.5}.

Reassuringly, there is no discontinuity of the distribution of male students around the passing threshold. Students on either side of the threshold are also similar on observable characteristics such as high school GPA, as a measure for student ability, and student age. Covariate continuity furthermore is balanced *within* gender, which means that, e.g. female students around the passing threshold have very similar high school grades and age. Also among students who failed their first university exam, female and male students look indistinguishable.

I next analyze heterogeneous effects and show that only German students react to failing their first exam by dropping out of university. There is no effect for Non-German students. This is consistent with an opportunity cost based explanation. Non-German students might face immigration or other restrictions and do not react to failing an exam. Next, only relatively older students react to failing their first exam by dropping out of university. Relatively older students are those who have worked or were involved in other activities before university, and as such likely face higher opportunity costs.

It is puzzling that male students show a strong negative reaction to early failure in university whereas female students do not. To analyze potential channels, I administer a survey among 927 students in the same course out-of-sample, in the first week of the fall semester 2022. I elicit expectations and attitudes towards failure and competition and, since the course is an introductory math class, some financial mathematics specific questions. There are several benefits of this exercise. First, this allows to measure attitudes and opinions which are not available in archival data. Second, I can link survey responses to students' realized performance in the exam. This allows to focus on students close to the passing threshold. These students are most important as they are closest to the regression discontinuity sub-population.

Expectations and overconfidence are likely a channel why male students drop out after failing an exam. The literature has shown that both men and women are overconfident, however men are more overconfident than women (Barber and Odean, 2001; Niederle and Vesterlund, 2007). In the survey, I asked students directly about their expected grade and on average I measure substantial overconfidence. Consistent with the literature, male students are more overconfident than female students. I next link survey responses to realized exam performance to show that male students are also more overconfident when *conditioning* on the realized performance in the exam. Male students around the passing threshold are in fact most overconfident. For male students, failing the exam might thus be an informational shock to their perceived ability.

Male students also self-assess as being less afraid and more prepared compared to female students. They are more likely to agree that they would be surprised to fail the course. They are less likely to agree that it would be a burden for them to fail the exam. The survey evidence indicates that male students are likely less emotionally prepared for early failure in university. Thus, failing the very first exam is likely to be a much more surprising event for male students compared to female students. They might react to this sudden shock of new information by dropping out of university.

Lastly, I look at competitiveness as a possible channel. The literature has shown that women shy away from competition, while men embrace it (Niederle and Vesterlund, 2007; Buser, Niederle, and Oosterbeek, 2014; Flory, Leibbrandt, and List, 2015; Reuben, Sapienza, and Zingales, 2015). I confirm this empirically in the survey. Male students are more competitive compared to female students and are more likely to compare their performance with peers. It is also more important for male students to be better than their peers. The observed gender gap in attitudes towards competitiveness might have some explanatory power for why male students drop out when faced with failure, whereas female students do not. Consistent with this, it is precisely those students who face the strongest competition, relatively bad students, who react by dropping out of university.

I also analyze students' retake behavior. 81.7% of students attempt the retake exam in the future. Male students are 5% less likely to do so. This explains a third of the observed reaction and indicates that the reaction by male students is quick. Conditional on attempting the retake exam, male students are also 5% less likely to pass the retake exam. While there is a math gender gap in the main exam, this gap reverses in the retake exam and male students perform worse. This indicates that male students might exert less effort in the retake compared to female students.

By and large, students who fail their very first university exam pass the exam at the second attempt. Among those who marginally failed, only 5.6% fail the retake exam. This puts an upper bound on the mechanical component of the causal effect of failing the university exam on university completion. However, this mechanical component should be the same for both female as well as male students and thus cannot explain the baseline results.

I contribute to several strands of literature. First I add to the literature on educational attainment (Denning et al., 2022). I provide one causal channel on why male students have lower educational attainment. Male students, but not female students, seem to react negatively to early failure in university.

Second, I contribute to how men and women react differently to feedback (Möbius et al., 2022). Previous research has shown that female students are more responsive than male students to positive incentives in the form of scholarships (Dynarski, 2008). Lindo, Sanders, and Oreopoulos (2010) analyze the reaction of students when they are placed in probation. The results are in contrast to (Wasserman, 2021) and (Wasserman, 2022). In these papers, female politicians are less (or equally) likely to persist after facing an electoral defeat.

The paper highlights the role expectations and belief updating can play in educational attainment as well as more generally (Möbius et al., 2022; Thaler, 2021; Giustinelli, 2022). In this paper, failing an exam leads to a sudden informational shock and this in turn leads to belief updating among students who fail the exam. Surprisingly, only male students react to this informational shock by dropping out of university, female students persist and do not react.

2.2. Institutional background and data

2.2.1. Institutional background

The data comes from the administration of a mid-sized German business school. All undergraduate students who major in economics, business, or law are mandated to take a course called "Financial Mathematics" in their first semester. The course covers basic concepts such as compound interest, net present value, annuity calculation and the rate of return of assets. 34% of students are majoring in business, 19% law, 16% in business and culture, 15% economics, and 14% in business education.

An important and distinct institutional feature is that the grade is 100% determined by a final exam and this exam already takes place after six weeks. This means that the course is the very first university exam and the first time students receive a signal on their relative performance in university. On average, around 1,200 undergraduate students take the class every year. Students have more than one opportunity to pass the exam. There are 3 credits, so-called European Credit Transfer and Accumulation System (ECTS) awarded upon successful completion. As a comparison, students need 180 ECTS to successfully complete a three-year undergraduate degree. Given the low number of credits awarded and the fact that students have several opportunities to pass the exam, the exam is low stakes. Students however perceive the exam as high stakes and there is substantial fear and uncertainty ahead of students very first university exam.

Of particular relevance for the methodology of this paper is the structure and grading of the exam. Grading of students was historically done in a cross-sectional fashion. Every year, the passing threshold was determined ex-post, such that around 15% of students fail the exam. After the exam is written, the course administrators compute the point distribution and then determine the point threshold such that 15% of the cohort does not pass the exam. E.g. in the year 2011, the passing threshold was set to 20 out of 45 points. This means that a student who obtains precisely 20 points just marginally passes the course and receives the lowest passing grade. A student with 19 points, in contrast, just marginally fails and receives a failing grade. The points needed to successfully pass the exam is not constant over the years. The passing threshold varied across the years 2008-2018 as follows: {22.5, 17, 19, 20, 16.5, 16.5, 18.5, 18, 15, 16.5, 20.5}. Since passing/failing is cross-sectional in nature, the points needed to pass the exam were therefore not known ahead of time to students and course administrators alike. The passing threshold is thus determined by 1) the difficulty of the exam and 2) the performance of the cohort. Given the uncertainty about the points needed to pass the exam, it is unlikely that students can precisely determine whether they pass or not. Later on in the paper, I revisit the ability of students to manipulate selection into treatment (failing the exam).

2.2.2. Data

I obtain the dependent variable, *Degree*, from the university administration. The variable is a dummy variable equal to one if a student has successfully obtained her degree. The variable is equal to zero if she has not successfully obtained her degree and is not currently enrolled anymore. Due to data protection reasons, students who are still enrolled cannot be considered.

I calculate the independent variable, *Points*, from historical grading data, collected from past course administrators. The data includes the total points, as well as the grade obtained in the financial mathematics exams of all students. The data also includes the passing threshold for each exam. The variable I compute is defined as the total point of the student minus the passing threshold in the respective exam. I refer to this variable as the point distance to the passing threshold. *Points* is equal to zero if the student just barely passes her exam with the exact points needed. It is equal to -1 for those who marginally fail the exam and equal to 1 for those who pass with a buffer of one point.

I obtain the gender of each student as recorded by the university administration. To control for student ability, I obtain the high school GPA. The high school GPA is by far the most important criteria for university admission in Germany. I also obtain a dummy variable equal to one if the nationality of the student is German, and zero otherwise. Lastly, to compute student age, I obtain the birth date.

The sample starts in the year 2010 as this marks the first year when grading data could be collected. I analyze all students who took the exam until the year 2017. The standard duration of an undergraduate degree is three years. I collect the information on university completion until the end of 2022, which means that students have at least five years to successfully complete their studies.

I apply the following filters to the data set. I only keep the very first exam for each student in the data. This means that if a student failed her exam, but passes it at the second try, I only keep the first failed exam in the data. I drop students who do not write the exam at the semester start (2% of the sample), are sick, or do not show up on the exam day (2% of the sample). I also exclude students who deliberately cross out and thus fail the exam (1% of the sample). These filter steps guarantee that I look at each students' very first university exam. This could contaminate the research design if students receive a signal on their quality beforehand. The filters also guarantee that the observed effect is not mechanical, as students have subsequent tries to pass the exam.

I end up with a sample of 8,588 students. The sample is purely cross-sectional, and every student only appears in the data set once. Every student has a certain point distance to the passing threshold in their first university exam. The outcome variable then measures whether this student successfully obtained her degree or not.

2.2.3. Descriptive statistics

Descriptive statistics are shown in table 2.1. Panel A shows the descriptive statistics for the full sample of students. 53% of students are female. 92% of undergraduate students are German. The remaining students are Chinese, Turkish, Bulgarian, among many other nationalities. The large fraction of German students is explained by the fact that for every undergraduate student, at least some courses are fully taught in German.
The average age of students at the time of the exam is equal to 19.9 at the mean and 20 at the median. The average German high school GPA is equal to 1.8 at the mean and 1.7 at the median. The high school GPA is by far the most important factor for university admission. The German grading scale ranges from 1.0 (best) to 5.0 (worst and a failing grade) and is inverted compared to an US-based GPA system. The grading is usually in increments of 0.3 as follows: 1.0, 1.3, 1.7, 2.0, 2.3, 2.7, 3.0, 3.3, 3.7, 4.0, 5.0. The best passing grade is a 4.0 and 5.0 is a failing grade.
The point distance to the passing threshold is 10.8 points at the mean and 12 at the median. The average grade achieved in the financial mathematics exam is 2.7 at the mean and median. Based on descriptive statistics, students seem to perform much worse in university compared to high school. This is a typical feature of good students enrolled in a competitive university with cross-sectional grading.
Consistent with previous research, I observe a substantial female financial math gap in the data. Panel B shows only female and Panel C only male students. Female students achieve on average 9.2 points distance to the passing threshold. Male students achieve

12.6.

The average probability to successfully complete university is 92%. Female students are more persistent and graduate at higher levels compared to male students. 92% of female students and 91% of male students successfully obtain an undergraduate degree. This is consistent with previous literature (Bailey and Dynarski, 2011; Goldin, Katz, and Kuziemko, 2006).

Table 2.1.: Summary statistics
The unit of observation is on a student level. Variable definitions are provided in the Appendix.

Panel A: Summary Statistics for All Students

Variable	N	Mean	SD	Min	25%	50%	75%	Max
Successful Degree	8,588	0.92	0.28	0	1	1	1	1
Point Distance to Pass	8,588	10.81	8.82	-10	5	12	18	25
Financial Math Grade	8,588	2.71	1.14	1.0	2.0	2.7	3.3	5.0
Male	8,588	0.47	0.50	0	0	0	1	1
School GPA	8,566	1.84	0.63	1	1.3	1.7	2.3	4
Age	8,588	19.93	1.77	16	19	20	20	42
German	8,588	0.92	0.27	0	1	1	1	1
Business Major	8,588	0.34	0.47	0	0	0	1	1
Law Major	8,588	0.19	0.39	0	0	0	0	1
Business and Culture Major	8,588	0.16	0.37	0	0	0	0	1
Economics Major	8,588	0.15	0.36	0	0	0	0	1
Business Education Major	8,588	0.14	0.35	0	0	0	0	1
Other Major	8,588	0.01	0.09	0	0	0	0	1

Panel B: Only Female Students

Variable	N	Mean	SD	Min	25%	50%	75%	Max
Successful Degree	4,539	0.92	0.27	0	1	1	1	1
Point Distance to Pass	4,539	9.19	8.70	-10	4	10	16	25
Financial Math Grade	4,539	2.92	1.13	1.0	2.0	2.7	3.7	5.0
School GPA	4,529	1.90	0.63	1	1.4	1.8	2.3	3.9
Age	4,539	19.91	1.78	17	19	20	20	40
German	4,539	0.90	0.30	0	1	1	1	1

Panel C: Only Male Students

Variable	N	Mean	SD	Min	25%	50%	75%	Max
Successful Degree	4,049	0.91	0.29	0	1	1	1	1
Point Distance to Pass	4,049	12.61	8.84	-10	7	14	20	25
Financial Math Grade	4,049	2.49	1.11	1.0	1.7	2.3	3.0	5.0
School GPA	4,037	1.78	1.13	1	1.7	1.6	2.2	4
Age	4,049	19.95	1.75	16	19	20	21	42
German	4,049	0.94	0.24	0	1	1	1	1

2.2.4. Survey data

I administered a survey out-of-sample, in the fall semester 2022. I asked 927 undergraduate students 16 questions related to students' attitudes towards competitiveness, failure, and expectations. I list all questions in the Appendix. The survey was administered in the first week of class before any contents were introduced. The goal of the survey is to shed light on potential channels for the baseline results.

I match survey responses to students' realized performance in the exam. This allows to compare male to female students conditional on realized grades. This allows to focus on students close to the passing threshold.

2.3. Regression discontinuity evidence

2.3.1. Baseline regression discontinuity scatterplot

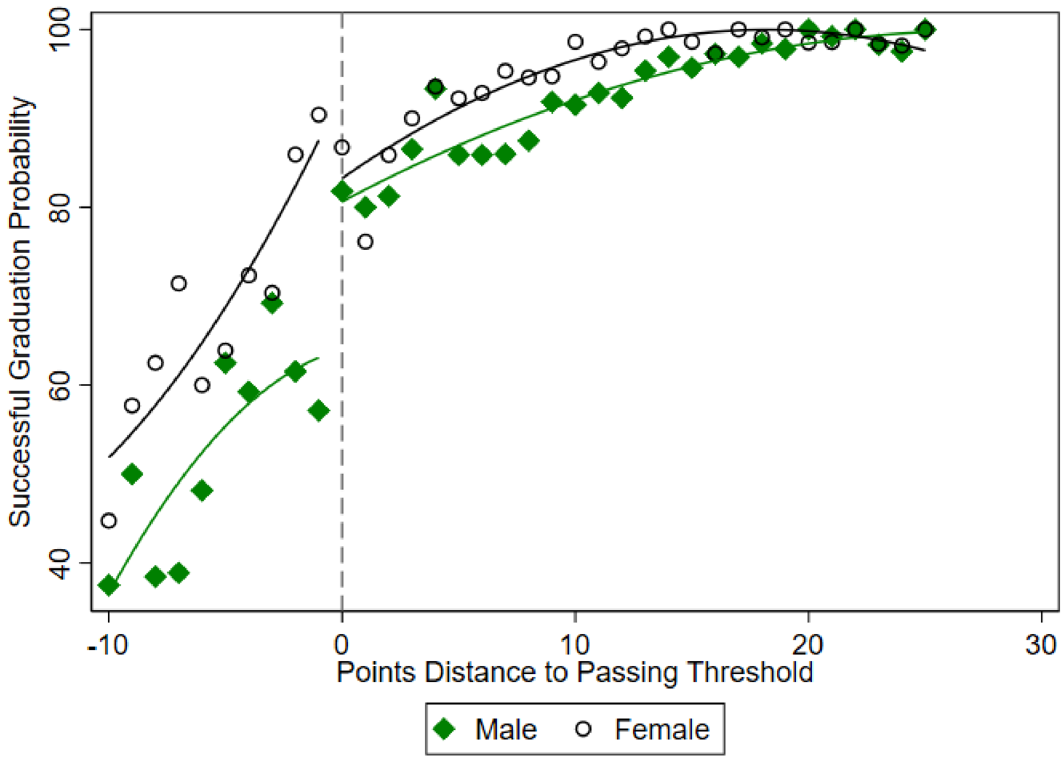
I visualize the average probability of obtaining a university degree, conditional on the point distance to the passing threshold in the financial mathematics exam, in a binned scatterplot and grouped by gender in figure 2.1.

For female students there is no discontinuity around the passing threshold. On both sides around the threshold of zero, female students have a probability of successfully obtaining a degree of 89%.

Male students marginally above the passing threshold successfully obtain a degree with a probability of 80%. Male students marginally below the passing threshold obtain a degree with a probability of 65%. Based on the raw data alone, there seems to be a sharp drop of 15% in the probability of completing the university degree for male students. This is an economically sizable effect.

Female students are completing university at higher rates throughout the distribution, conditional on their results in a financial mathematics exam. Female students thus seem to be more resilient in educational attainment than male students. Additionally, failing their very first mathematics exam does not seem have any effect on the resilience of female students.

Figure 2.1.: Regression discontinuity: baseline results
This figure visualizes the raw data in a binned scatterplot. On the x-axis is the distance to the passing threshold in the financial mathematics exam. On the y-axis is the average university completion probability. Students are binned per point distance and capped at the extreme end at -10 and 25 points respectively. Male students are visualized in green squares and female students in white circles. The graph includes a polynomial of second order to both sides of the passing threshold and separately for each subgroup.



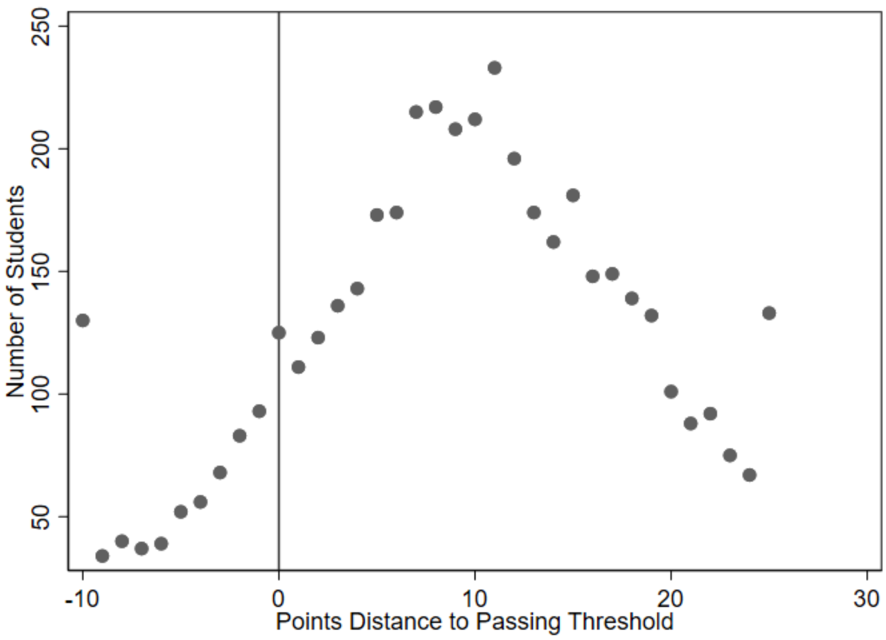
2.3.2. Are students able to manipulate the running variable?

The crucial assumption for analyzing causal effects is whether marginally passing, or marginally failing the first university exam is to some extent randomly allocated. A concern for a causal interpretation is whether students can precisely determine whether they pass the exam or not. This would be problematic if particularly skilled students manage to obtain just marginally enough points in order to pass the exam. If these students are also more likely to complete university, a causal interpretation is not valid. As a first test of this assumption, I visualize the distribution of students over all instances of the point distance to the passing threshold in figure 2.2.

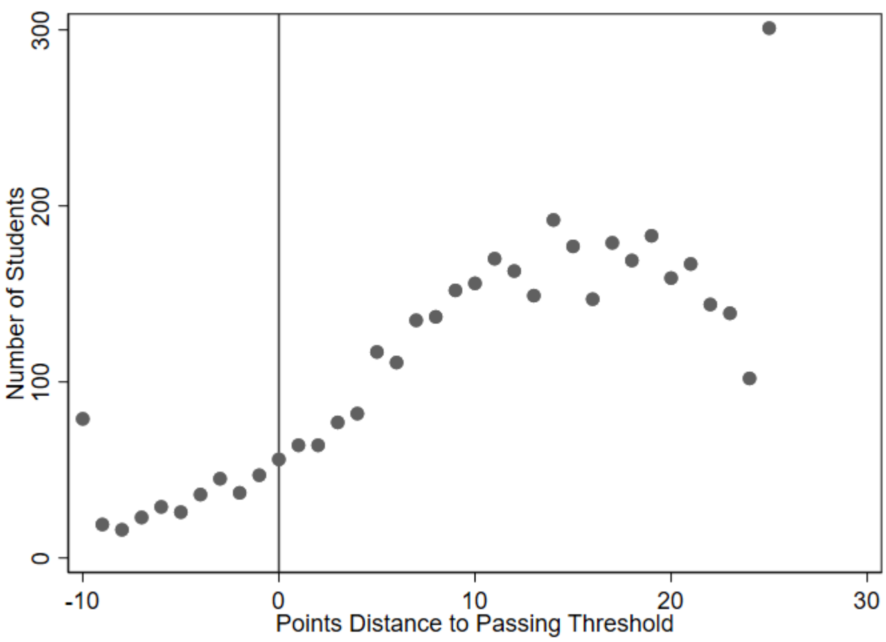
There is evidence for slight bunching above the passing threshold only for female students, however similar jumps appear throughout the distribution. Reassuringly, there is no discontinuity of the distribution of male students around the passing threshold.

Figure 2.2.: Threshold manipulation
This figure visualizes the distribution around the threshold. On the x-axis is the distance to the passing threshold in the financial mathematics exam. On the y-axis is the number of students. Students are binned per point distance and capped at the extreme end at -10 and 25 points respectively. Panel A shows the results only for female students. Panel B shows the results only for male students.

Panel A: Only Female Students



Panel B: Only Male Students



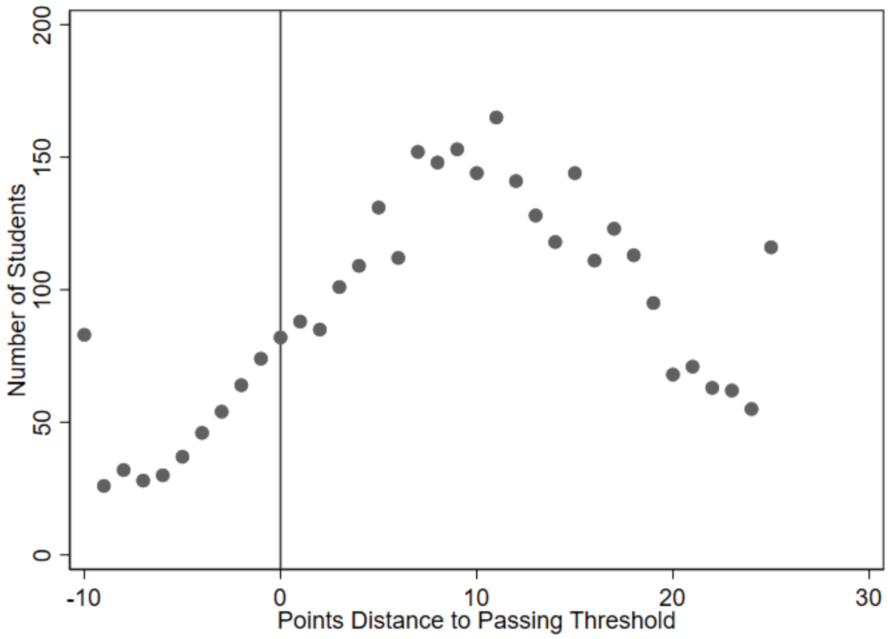
I argue that manipulation of the running variable, point distance to the passing threshold, by students is unlikely for several reasons: First, students might very well be able to determine how many points they achieve in the exam. However, it is probably not possible to do so with very high confidence, as grading by course instructors might be subjective. More importantly, the variable point distance to the passing threshold includes an additional component: the passing threshold. The passing threshold was not constant and unknown for students at the time of writing the exam. The precise points needed to pass the exam was also unknown to the instructors. The passing threshold varied across the years 2008-2018 as follows: {22.5, 17, 19, 20, 16.5, 16.5, 18.5, 18, 15, 16.5, 20.5} out of a maximum of 45 points. In every year, the passing threshold was set such that 15% of the students did not pass the course. The passing threshold was thus determined by outside factors such as the difficulty of the exam and the performance of each cohort of students. A student in the year 2010 needs to aim for precisely 19 points to pass the exam. If she would write the exam instead in the year 2008 or 2011, this would be a failing grade. The uncertainty involved in the passing threshold prohibits students to precisely manipulate the points they obtain in the exam to just marginally pass. To some degree, there is a random component in whether a student passes or fails her first university exam. Students are better off performing as good as they possibly can, and this is the most consistent explanation given the point distribution.

Second, no bunching occurs for male students around the passing threshold. There is limited bunching for female students, however this does not directly imply that students are able to precisely determine the point distance to the passing threshold. It might also be course administrators who push marginal students above the passing threshold. Indeed, some past course administrators corrected exams of students who marginally failed one additional time. Points were in some marginal cases adjusted upwards. This was only done for those who marginally failed the exam and never for those who marginally passed. Bunching above the threshold is thus not necessarily evidence in favor of running variable manipulation by students.

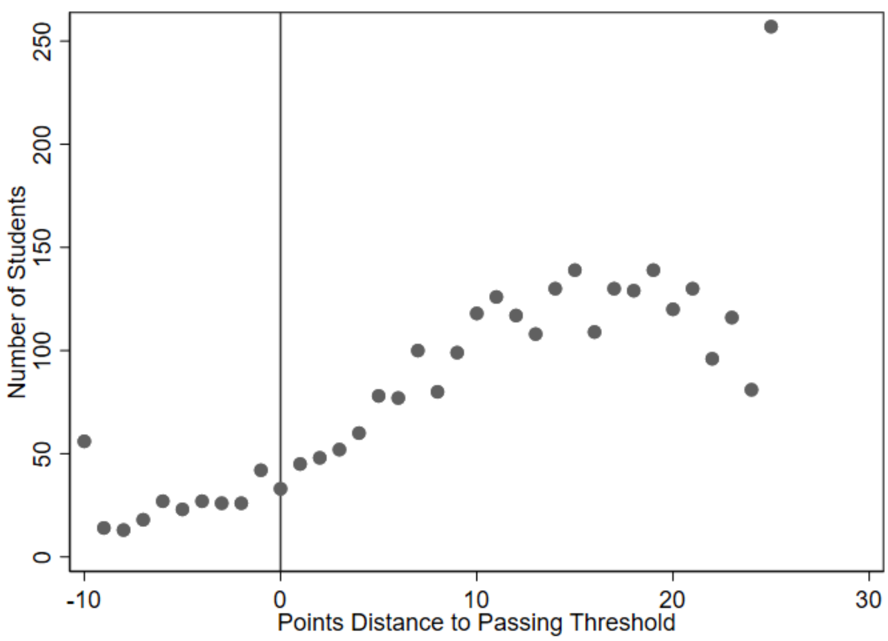
Nevertheless, the slight bunching of female students might be problematic. To mitigate this, I perform a robustness exercise. The discontinuity originates from two years in the sample. Specifically, in the years 2011 and 2015, the course instructors regraded all exams marginally below the passing threshold. I repeat all analyses when excluding these two years. The results are shown in figure 2.3. The distribution is smooth overall for male as well as female students and there is no visible bunching around the threshold. All results are unchanged when excluding these two years.

Figure 2.3.: Threshold Manipulation- Without 2011 and 2015
This figure visualizes the distribution around the threshold. Two years were omitted from this graph, the years 2011 and 2015. Only in these two years, students who marginally failed were regraded by the course instructor. Students in those two years are therefore arguably more distant in terms of their performance compared to other years. On the x-axis is the distance to the passing threshold in the financial mathematics exam. On the y-axis is the number of students. Students are binned per point difference and capped at the extreme end at -10 and 25 points respectively. Panel A shows the results only for female students. Panel B shows the results only for male students.

Panel A: Only Female Students



Panel B: Only Male Students



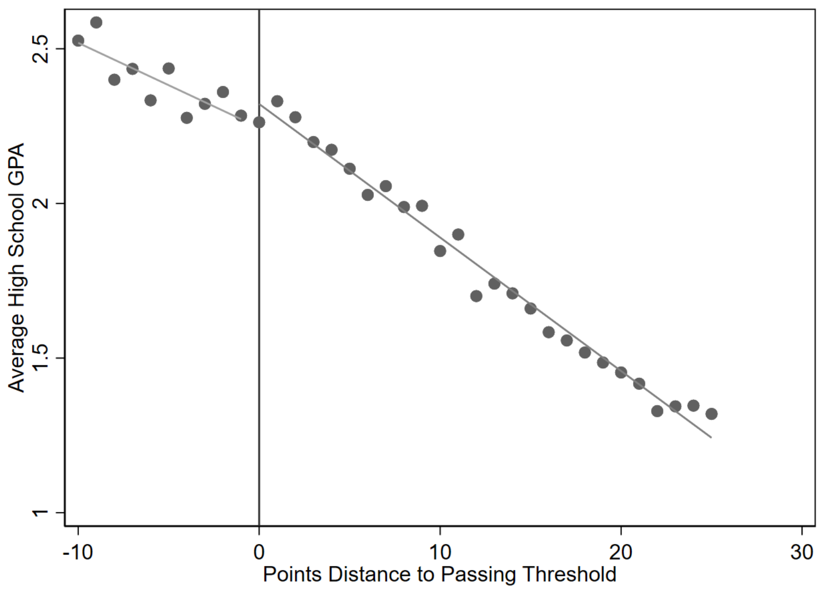
2.3.3. Are students on either side comparable?

Next, I analyze whether students on either side of the threshold are different when it comes to observable characteristics. As an imperfect proxy for student ability, the first characteristic I visualize is the high school GPA achieved. I perform a similar exercise as before, but instead of the number of students, I compute the average high school GPA for every point distance to the passing threshold.

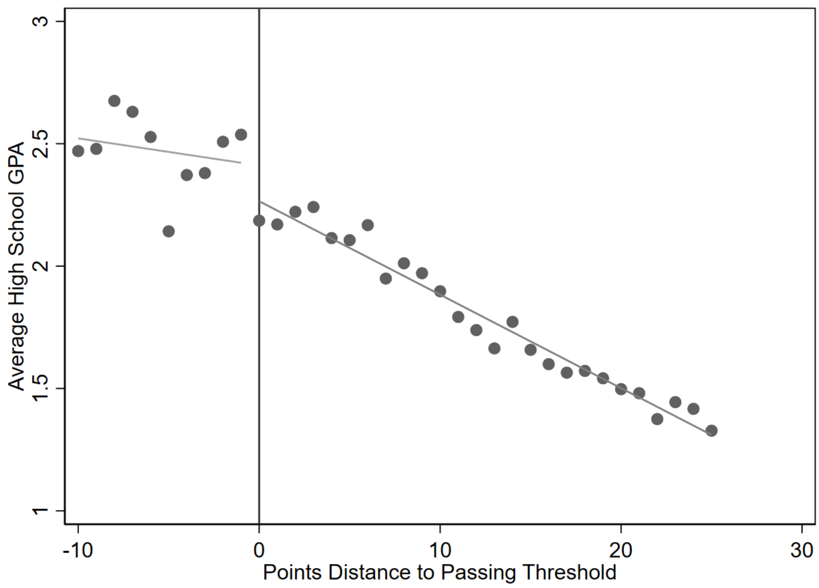
The result is visualized in figure 2.4. The average high school grade looks relatively smooth around the passing threshold. For female students there is no visible discontinuity. For male students, there is a statistically insignificant jump of around 0.2 GPA around the threshold. However, similar jumps appear at other instances.

Figure 2.4.: High school GPA around threshold
This figure visualizes the high school average GPA around the threshold. On the x-axis is the distance to the passing threshold in the financial mathematics exam. On the y-axis is the average high school GPA, which is ranging from 1.0 (best) to 4.0 (worst). Students are binned per point distance and capped at the extreme end at -10 and 25 points respectively. Panel A shows the results only for female students. Panel B shows the results only for male students.

Panel A: Only Female Students



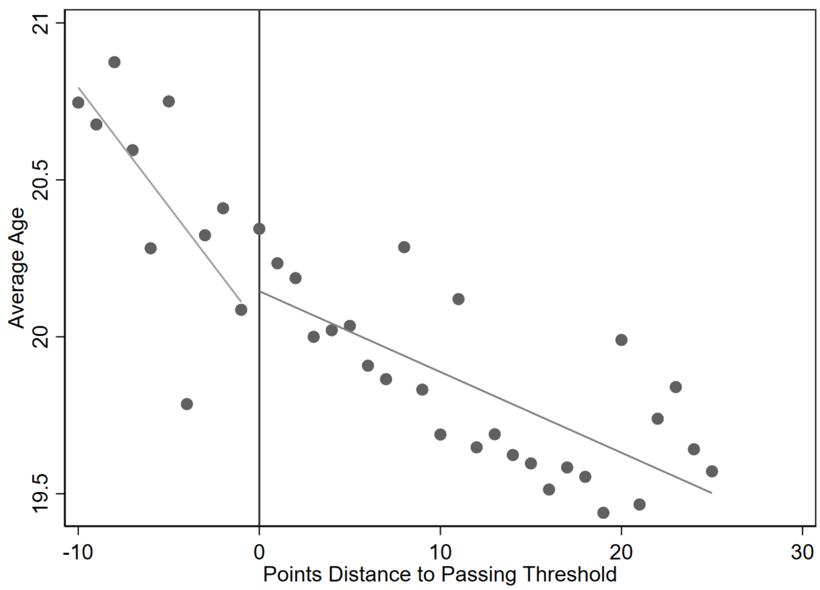
Panel B: Only Male Students



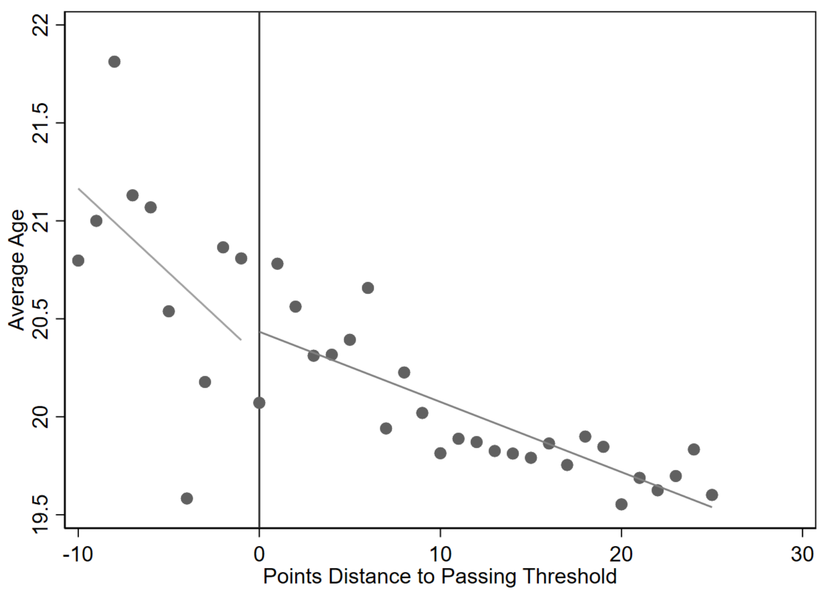
Lastly, I visualize the average age of students around the threshold in figure 2.5. The average age looks continuous around the threshold, but the variable is noisy and I see frequent jumps in the distribution. For female students there is no visible discontinuity. For male students, those at the precise cutoff are somewhat younger compared to those below. However, widening the bandwidth by one point, students look very similar.

Figure 2.5.: Age around threshold
This figure visualizes the average age around the threshold. On the x-axis is the distance to the passing threshold in the financial mathematics exam. On the y-axis is the average age at the time of the exam. Students are binned per point distance and capped at the extreme end at -10 and 25 points respectively. Panel A shows the results only for female students. Panel B shows the results only for male students.

Panel A: Only Female Students



Panel B: Only Male Students



2.3.4. Baseline regression

I estimate the following baseline specification:

$$\begin{aligned} Degree_i =& \alpha Male_i + \beta Fail_i + \gamma(Male_i \times Fail_i) + f(P_i) + \theta(Fail_i \times f(P_i)) + \\ & \zeta(Male_i \times Fail_i \times f(P_i)) + \phi_t + \epsilon_i \end{aligned} \tag{2.1}$$

where $Degree$ is defined as a dummy variable equal to one if student i successfully obtains an undergraduate degree, and zero if not. Year t is defined as the year of university entry which coincides with the year of the financial mathematics exam. The variable $Male$ is equal to one if the student is male, as indicated by university administration data. The variable $Fail$ is equal to one if the student failed her first university exam. The running variable, point distance to the passing threshold is included as a function either as a linear term or using higher order polynomials.

$Male_i$ captures a level shift between the average passing probability of male relative to female students. $Fail_i$ captures the intercept shift for female students who fail the exam. The variable of interest is thus the interaction term $Male_i \times Fail_i$ which picks up the effect of failing the first university exam for male students. The function $f(P_i)$ captures the effect of the point distance to the passing threshold for female students. The interaction with $Fail_i \times f(P_i)$ allows to include a different slope for female students who fail the exam. The interaction with $Male_i \times f(P_i)$ allows to include a different slope for male students who pass the exam. Lastly, the triple interaction $Male_i \times Fail_i \times f(P_i)$ captures a different slope estimate for male students who failed their first exam. I include Major $\times$ Year fixed effects and cluster standard errors on the running variable (Lee and Card, 2008).

To ease interpretation, I invert the running variable. I multiply the point distance to the passing threshold by -1 and subtract a constant of 0.000005 to the students with a value equal to zero. Treatment is defined as failing the very first university exam, so after this modification, I can interpret the treatment indicator $Male_i \times Fail_i$ as the causal effect of failing the very first university exam for male students.

The results are shown in table 2.2. When male students marginally fail their very first university exam, the probability of successfully obtaining a degree decreases by between 14% to 32% depending on the specification. Male students are on average less likely to obtain a degree compared to female students. When using local linear functions, women appear to be more likely to successfully finish university when they fail the exam, however this effect disappears when looking at either full sample linear, parametric regressions, or choosing a local linear non-parametric specification with a bandwidth of 2. Only male students seem to significantly react to failing their very first university exam. Overall the coefficients on $Male_i \times Fail_i$ is well aligned with the visual results presented earlier. Older students are less likely to successfully finish their undergraduate degree. I also see a strong relationship between high school performance and students' likelihood to complete university. Lastly, German students are much more likely to complete university compared to non-German students. These results are consistent with results in the literature.

The table also reports parametric regression discontinuity specifications using a second-order polynomial of the point distance to the passing threshold using the full sample of students. The optimal bandwidth is calculated as equal to 3 points around the threshold (Calonico, Cattaneo, and Titiunik, 2014). The regression output thus reports non-parametric local linear regression with the optimal bandwidth of 3 points, as well as using either 2 or 4 points around the threshold.

The results are robust to using local randomization regression discontinuity approaches. In the context of this research question, the running variable is not continuous, but can be seen as discrete. This leads to a moderate number of distinct masspoints. The number of discrete instances of the point distance to the passing threshold is equal to 36 unique values in the interval [-10,25]. Since the optimal bandwidth in this context is not necessarily appropriate, I refer to economic intuition. The most stringent bandwidth would be one point. This equates to comparing students just above to just below the threshold. The results are unchanged to using this most stringent comparison or widening the interval to either two or three points around the cutoff. This is essentially a trade-off between sample size and the assumption of random assignment into treatment in a narrow window around the threshold.

Table 2.2.: Baseline regression discontinuity: reacting to early failure
This table reports the regression discontinuity of equation 2.1. The dependent variable is equal to one if the student successfully finished her undergraduate degree. Column (1) shows a linear regression using the full sample. Column (2) shows parametric regressions using a fully interacted model including a second order polynomial of the running variable. Columns (3) to (5) display non-parametric local linear regressions with a bandwidth of 2, 3, and 4 respectively. The optimal bandwidth is calculated as equal to 3 following Calonico, Cattaneo, and Titiunik (2014). Variable definitions are provided in the Appendix. The regression includes Major × Year fixed effects. Standard errors are clustered on the level of the running variable: point distance to the passing threshold. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)	(3)	(4)	(5)
Sample:	Full Sample		Bandwidth: 2	Bandwidth: 3	Bandwidth: 4
Polynomial Order:	1	2	1	1	1
$Male$	-0.04**	-0.03	-0.07	-0.05	-0.06
	(-2.50)	(-1.37)	(-1.44)	(-0.76)	(-1.12)
$Fail$	0.02	0.06	0.06	0.17**	0.14**
	(0.47)	(1.31)	(1.67)	(2.55)	(2.57)
$Male \times Fail$	-0.14**	-0.28***	-0.17**	-0.32***	-0.24***
	(-2.13)	(-5.54)	(-3.60)	(-4.05)	(-3.77)
$SchoolGPA$	-0.05***	-0.05***	-0.12**	-0.08**	-0.08**
	(-5.26)	(-5.20)	(-3.54)	(-3.09)	(-2.69)
Age	-0.01***	-0.01***	-0.02**	-0.02**	-0.02***
	(-4.26)	(-4.20)	(-3.88)	(-3.59)	(-4.57)
$German$	0.08***	0.08***	0.19**	0.14**	0.15***
	(5.59)	(5.45)	(3.55)	(3.16)	(4.04)
Observations	8,563	8,563	797	1,121	1,438
R-squared	0.24	0.24	0.15	0.13	0.14
Major × Year FE	YES	YES	YES	YES	YES

2.4. What explains the reaction of male students?

In the following, I will first analyze heterogeneity in the data and second why male students might react strongly to early failure in university, while female students do not. I do so using two complementary datasets. First, I rely on sources of heterogeneity in the data. Students differ along various characteristics, which might indicate why some drop out and others do not. The benefit of relying on the regression discontinuity sample is that it relies on a revealed choice: dropping out. The drawback is that I have little data and imperfect proxies. The second dataset comes from a survey I administered out-of-sample among 927 students who took the course in the fall semester 2022. The benefit of the survey is that I could elicit expectations and self-assessments on potential channels that are unobservable in the archival data. I match survey responses to the realized exam performance, which allows to analyze gender differences in survey responses particularly for students around the passing threshold. This is the local student population most relevant for the research design and I particularly focus on gender differences in this local subset. The drawback is that since the survey includes the out-of-sample cohort of 2022, it is impossible to analyze who eventually drops out of university.

2.4.1. Heterogeneity

I explore two separate sources of heterogeneity in the data. First, I split the sample into German and Non-German students. Non-German students might have visa restrictions and face more legal and financial mobility restrictions compared to German students. Consistent with this, the effects are confined to male students who are German. Non-German students do not seem to react by dropping out of university.

The second source of heterogeneity I explore is student age. I split the sample at the median into relatively older and younger students. Only male students who are relatively old drop out of university. Relatively older students are more likely to be involved in some other activity before starting university. Such students might have worked, finished an apprentice program, etc. Older students might thus have higher opportunity costs of continuing education. Or to phrase it differently, it might be easier for them to switch to another activity besides full time studying. Upon early failure in university, they might go back to their previous job or switch to another university.

Table 2.3.: Heterogeneity: Old and German students
This table reports heterogeneity regressions similar to equation 2.1. The dependent variable is equal to one if the student successfully finished her undergraduate degree. The sample is composed with a bandwidth of 4 points around the cutoff. The sample is split around the local median into two parts. In column (1) and (2), the students are split into those below the age of 20.2 (young students) and those above (old). In column (3) and (4), the students are split into German and Non-German. Variable definitions are provided in the Appendix. Standard errors are clustered on the level of the running variable: point distance to the passing threshold. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)	(3)	(4)
Sample:	Young students	Old students	German	Non-German
$Male$	-0.07	-0.10***	-0.07*	-0.17
	(-1.50)	(-4.16)	(-2.29)	(-0.85)
$Fail$	0.20**	0.13*	0.16**	0.14
	(2.97)	(2.16)	(3.20)	(1.13)
$Male \times Fail$	-0.22	-0.42***	-0.39***	0.38
	(-1.79)	(-12.94)	(-6.93)	(1.52)
Observations	677	675	1,221	130
R-squared	0.09	0.12	0.10	0.24
Controls	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

2.4.2. Expectations, overconfidence, and failure

Expectations might play an important role and explain the differing response of male students compared to female students. To analyze this potential channel, I add evidence from the survey. The sample is composed of 927 students in the out-of-sample 2022 cohort. I specifically asked students at the beginning of the semester what grade they expect to earn in financial mathematics. Students could select every grade step from 1.0 (best) to 5.0 (worst, and a failing grade). Students expect significantly better grades (0.5 grade points on average) than they ended up achieving.

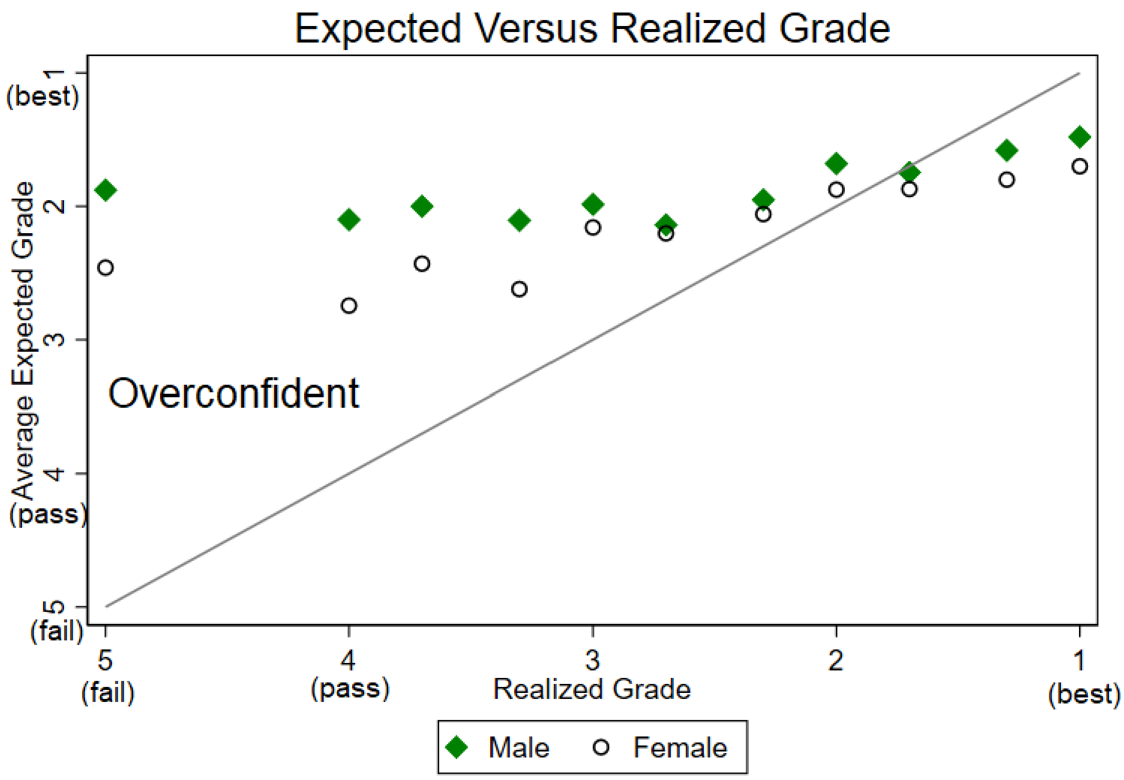
There is significant sorting of students into majors. Different majors differ on how competitive they are. By far the most important criteria to enter a certain major is the high school average grade. Because of this, all regressions include major fixed effects and thus for example compare male economics students to female economics students.

In table 2.4 column (1), male students on average expect 0.16 better grades compared to female students. The results are a first indication that male students are overconfident compared to female students. Next, I condition on the realized exam performance. I match the survey responses to realized grades to visualize the gender gap in expectations. Figure 2.6 visualizes the results.

On the x-axis, students are grouped by what grade they achieved in the course, separate by gender. On the y-axis is the average expected grade of these students. On average both female and male students expect worse grades only for the three best grades 1.0, 1.3, and 1.7. From grade 2.0 onwards, students expect better grades.

Along the complete realized grade distribution, male students expect better grades compared to female students. Thus, even conditioning on the performance in financial mathematics, male students seem to be overconfident. Overconfidence is small at the upper end of the distribution and smallest in the middle. It is largest at the tail end of the distribution, precisely in the area close to the passing threshold. Thus in the local area relevant for the baseline results, male students are most overconfident compared to female students. Overconfidence might thus explain some of the response to failing an exam for male students.

Figure 2.6.: Expected versus realized grade
This figure visualizes the average expected grade on the y-axis and the realized grade on the x-axis.



Next I look at the general attitude of male students towards failing. The goal is to elicit whether male compared to female students differ in their expectations towards failure specific to the financial mathematics exam. I ask students to what extent they agree to the following statements: "I would be surprised to fail the financial mathematics exam", "I am afraid of financial mathematics", "It is a burden to fail this course", and "I would consider dropping out of university if I fail this course".

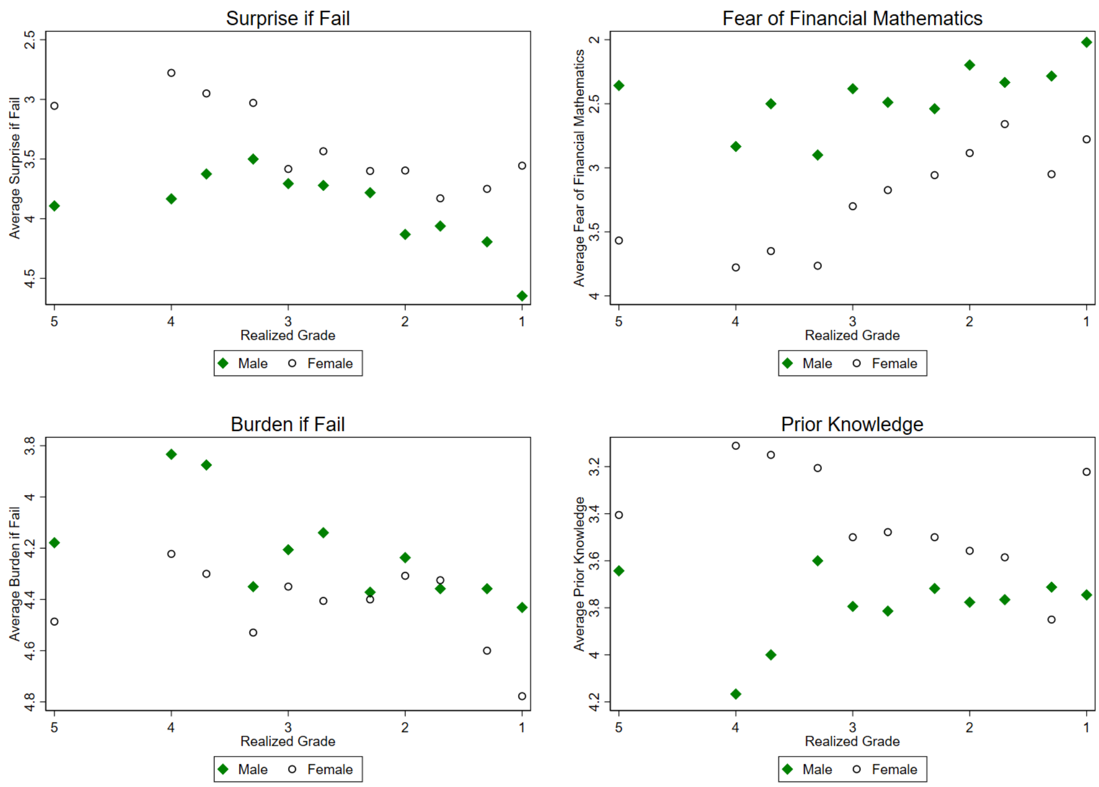
Table 2.4 columns (2) to (5) shows the results. Male students are significantly more likely to be surprised if they would fail the course. Male students are also much less likely to be afraid of financial mathematics. They are also slightly less likely to say failing the exam would be a burden. There is no difference in their personal perception of whether they would consider dropping out of university if they would fail the exam.

Table 2.4.: Survey evidence: expectations and failure
This table shows survey results of the out-of-sample cohort of the fall semester 2022. In column (1) I ask respondents what grade they expect, ranging from 1.0 (best) to 4.0 (worst). In the following columns, I ask respondents on a 5 point Likert scale to what extent they agree with the following statements: (2) I will be surprised if I fail this exam. (3) I am afraid of this course. (4) It will be a burden for me if I fail this course. (5) I will consider dropping out if I fail this course. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)	(3)	(4)	(5)
	Expected Grade	Surprised if fail exam	Afraid of course	Burden if fail	Drop out if fail
Male	-0.16***	0.40***	-0.66***	-0.10*	0.08
	(-4.00)	(6.59)	(-8.94)	(-1.81)	(1.20)
Major FE	YES	YES	YES	YES	YES
Observations	927	927	927	926	926
R-squared	0.23	0.16	0.20	0.02	0.00

Similar as before, I visualize the answers to four of these questions in figure 2.7. Male students are most overconfident at the bottom of the realized performance, close to the passing threshold. There is a large gender gap as male students are more surprised if they would fail the exam, have much less fear of the exam, claim that it would not be burden for them to fail, and that prior knowledge will help them master the course.

Figure 2.7.: Overconfidence and failure
This figure visualizes survey results of the out-of-sample cohort of the fall semester 2022. On the x-axis, students are grouped by their realized grade in the course. I asked respondents to what extent they agree to various statements. Top left: "I would be surprised to fail the financial mathematics exam", top right: "I am afraid of financial mathematics", bottom left: "It is a burden to fail this course", and bottom right: "prior knowledge will help me master this course". The responses were on a 5 point Likert scale.



2.4.3. Competitiveness

Another channel I investigate is the role of competition and attitudes towards competitiveness. In the context of the results, the realization of how (little) competitive one is might be a shock. Students might suddenly realize that they are facing strong competition when they fail their first university exam.

I first use evidence of the survey and see in table 2.5 that male students are significantly more likely to compare their performance to peers. It is also more important for male students to be better than their peers. They are more competitive and self-assess as more likely to want to win a game. This is consistent with the literature in which men are consistently seen as more competitive compared to female students.

Table 2.5.: Survey evidence: competition
This table shows survey results of the out-of-sample cohort of the fall semester 2022. I ask respondent on a 5 point Likert scale to what extent they agree to the following statements: (1) I often compare my results with my peers. (2) It is important for me to be better than my peers. (3) When I play a game I want to win. (4) My performance is important to my self-worth. (5) I often think about my own performance. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)	(3)	(4)	(5)
	Compare results with peers	Important to be better than peers	Want to win game	Performance important self-worth	Think often about performance
Male	0.17***	0.20***	0.41***	-0.18***	-0.18***
	(2.60)	(2.94)	(7.27)	(-3.17)	(-3.59)
Major FE	YES	YES	YES	YES	YES
Observations	927	927	927	926	927
R-squared	0.04	0.11	0.06	0.01	0.03

I add evidence from the archival data to this. Students in the sample were among the best in high school, but are suddenly compared to other equally high achieving students in university. I hypothesize that relatively worse students, those who suddenly face more fierce competition, are reacting more negatively to early failure in university.

To analyze this question in the regression discontinuity sample, I split students into two subgroups. Depending on their high school GPA, relatively worse students, who I argue face much more competition and relatively good students who face less competition.

Indeed, dropping out of university is strongly concentrated in the subgroup of male students who are relatively bad and face strong competition in university.

Table 2.6.: Regression discontinuity: competition
This table reports heterogeneity regressions similar to equation 2.1. The dependent variable is equal to one if the student successfully finished her undergraduate degree. The sample is defined with a bandwidth of 4 points around the cutoff. The sample is split around the local median into two parts. In column (1) and (2), the students are split into relatively good (above a GPA of 2.2.) and relative bad students, respectively. Variable definitions are provided in the Appendix. Standard errors are clustered on the level of the running variable: point distance to the passing threshold. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)
Sample:	Good students	Bad students
$Male$	-0.04	-0.11***
	(-0.76)	(-7.15)
$Fail$	0.16***	0.19**
	(4.04)	(2.71)
$Male \times Fail$	0.00	-0.59***
	(0.06)	(-8.23)
Observations	745	608
R-squared	0.09	0.12
Controls	YES	YES
Year FE	YES	YES

2.4.4. Retake behavior

Lastly, I analyze how male and female students differ in their retaking behavior after failing the exam. Do male and female students attempt the retake at similar rates? And conditional on retaking the exam, how do male and female students perform? To analyze these questions, I construct data which captures retake behavior of students who failed their first exam. I first calculate a dummy equal to one if the student attempts the retake exam, which 81.7% of students do. I then analyze gender differences in table 2.7.

Male students are 5% less likely to attempt the retake exam compared to female students, significant at a 10% level, and marginally insignificant at a 5% level. About one third of the baseline effect can thus be explained by the fact that male students do not attempt to retake the exam. Male students seem to react quickly and drop out of university.

Second, I analyze the performance in the retake conditional on retaking. Conditional on retaking, male students are 5% less likely to pass the retake. Looking at the point distance in the retake, there is no statistically significant difference between male and female students. Noteworthy is that while there is a gender gap overall in the financial mathematics exam, when looking at the subset of students who fail the exam, the gender gap reverses and male students seem to perform worse than female students. This might indicate that female students exert more effort compared to male students in the retake exam.

Table 2.7.: Exam retake behavior

This table shows regressions on students retake behavior. In column (1) the dependent variable is a dummy equal to one if the student attempts a retake exam in the future. In column (2), the variable is a dummy whether the student passes the retake. In column (3) the dependent variable is the point distance to the passing threshold in the retake exam. The sample only includes students who failed their first attempt. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively. t-statistics are displayed in parenthesis.

	Attempt Retake	Pass Retake	Points Retake
$Male$	-0.05*	-0.05*	-0.81
	(-1.89)	(-1.81)	(-1.38)
$SchoolGPA$	-0.04	-0.09***	-3.09***
	(-1.48)	(-3.82)	(-5.47)
Age	-0.01	-0.00	0.03
	(-1.57)	(-0.80)	(0.22)
$German$	-0.02	0.08**	3.69***
	(-0.42)	(2.04)	(4.23)
Observations	984	805	805
R-squared	0.04	0.06	0.13
Major FE	YES	YES	YES
Year FE	YES	YES	YES

2.5. Conclusion

This paper analyzes the question whether failing the very first university exam causes students to drop out of university. I exploit university administration data of around 8,500 students and a sharp discontinuity at the passing threshold of the very first university exam. Male students who marginally fail their very first university exam are 15% less likely to successfully obtain a university degree. Female students on the other hand are much more resilient to failure in university. The channels are consistent with the explanation that overconfidence and attitudes to competitiveness explain the reaction of male students. The results provide causal evidence of one explanation on why male students are less likely to successfully obtain a university degree: male students react strongly negative to early failure in university.

APPENDIX

2.6. Appendix

Variable Definitions The data is on a pure cross-sectional student level.

1. *Points (Distance to Passing Threshold)* – Number of points relative to the passing threshold. 0 indicates that the student has just passed the exam. -1 equals that one additional point was needed to pass the exam. +1 indicates that the students passed the exam with a buffer of one point. The variable points is binned at the two extremes at -10 and at +25.

2. *(Financial Mathematics) Grade* – Grade captures what grade the student achieved in her first university exam. The German grading scale ranges from 1.0 (best) to 5.0 (fail), usually in increments of 0.3 as follows: 1.0, 1.3, 1.7, 2.0, 2.3, 2.7, 3.0, 3.3, 3.7, 4.0, 5.0. The best passing grade is a 4.0 and 5.0 is a failing grade.

3. *Fail* – Dummy variable equal to one if the student did not pass her very first university exam: financial mathematics.

4. *(Successful) Degree* – Dummy variable equal to one if the student has successfully completed her undergraduate university degree.

5. *School GPA* – High school average grade which is used for university admission. The German educational system does not use standardized tests, thus high school GPA is by far the most important criteria for university admission. The German grading scale ranges from 1.0 (best) to 5.0 (worst and a failing grade), usually in increments of 0.3 as follows: 1.0, 1.3, 1.7, 2.0, 2.3, 2.7, 3.0, 3.3, 3.7, 4.0, 5.0. The best passing grade is a 4.0 and 5.0 is a failing grade.

6. *Age* – Age of the student at the time of the exam.

7. *German* – Dummy equal to one if the student is a German national.

8. *Major* – Students are enrolled in one of the following majors: Business, Business Law, Business and Culture, Economics, Business Education, or others (such as Psychology, Sociology, or History)

Survey questions

1. What grade are you expecting to earn in this course? The German grading scale ranges from 1.0 (best) to 5.0 (fail), usually in increments of 0.3 as follows: 1.0, 1.3, 1.7, 2.0, 2.3, 2.7, 3.0, 3.3, 3.7, 4.0, 5.0. The best passing grade is a 4.0 and 5.0 is a failing grade.

2. I often compare results with my peers. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

3. It is important for me to be better than my peers. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

4. If i play a game, I want to win. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

5. My university performance is important for my self worth. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

6. I often think about my university performance. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

7. It is important for me to be good in financial mathematics. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

8. Financial mathematics is of interest to me. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

9. Financial mathematics is an important subject for me. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

10. The contents in this course will be helpful for me later on. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

11. Prior knowledge will help me master the course. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

12. My peers think that financial mathematics is interesting. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

13. I would be surprised to fail the financial mathematics exam. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

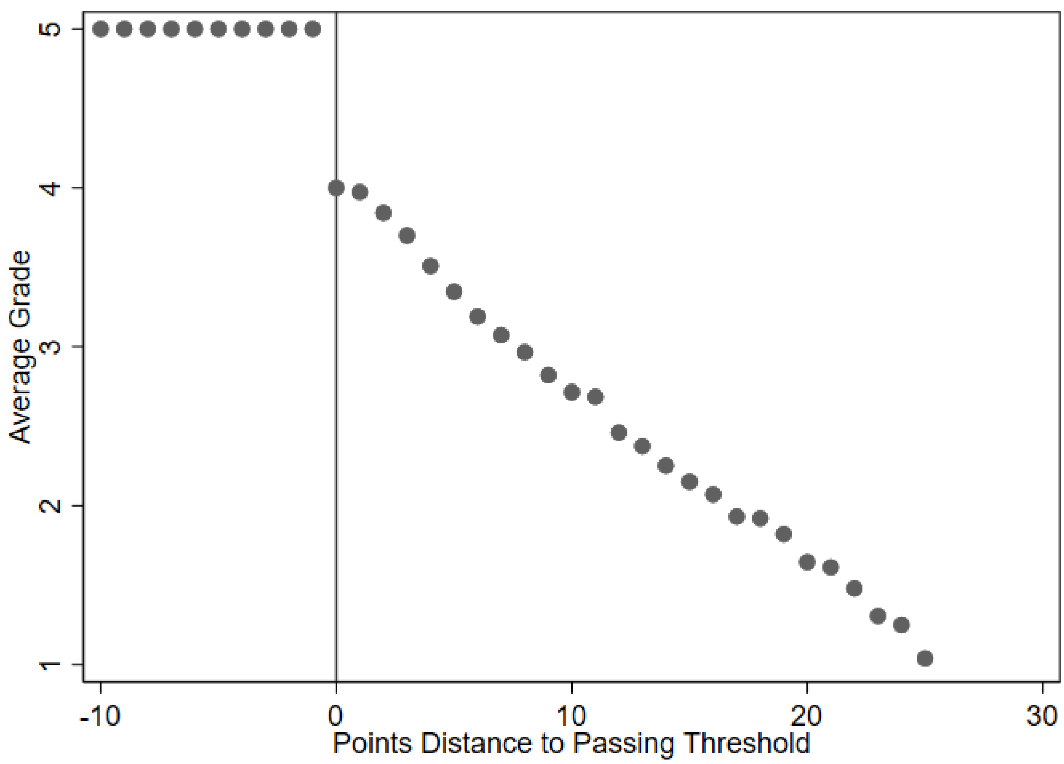
14. I am afraid of financial mathematics. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

15. Men are better at solving mathematical problems compared to women. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

16. It is a burden to fail this course. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

17. I would consider dropping out of university if I fail this course. Likert scale ranging from 1 (I disagree strongly) to 5 (I agree strongly).

Figure A1.: Visualizing the Sharp Discontinuity
This figure visualizes the sharp discontinuity which is exploited in the analysis. On the x-axis is the distance to the passing threshold in the first university exam. On the y-axis is the average grade, which is a function of the point distance in the exam. Students are binned per point difference and capped at the extreme end at -10 and 25 points respectively. The German grading scale ranges from 1.0 (best) to 5.0 (worst and a failing grade), usually in increments of 0.3 as follows: 1.0, 1.3, 1.7, 2.0, 2.3, 2.7, 3.0, 3.3, 3.7, 4.0, 5.0. The best passing grade is a 4.0 and 5.0 is a failing grade. Marginally failing the exam results in a sharp drop from grade 4.0 to 5.0.



References

Bailey, Martha and Susan Dynarski (2011). *Gains and Gaps: Changing Inequality in U.S. College Entry and Completion.* Cambridge, MA.

Barber, B. M. and T. Odean (2001). "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment". In: *The Quarterly Journal of Economics* 116.1, pp. 261–292.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek (2014). "Gender, Competitiveness, and Career Choices". In: *The Quarterly Journal of Economics* 129.3, pp. 1409–1447.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik (2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs". In: *Econometrica* 82.6, pp. 2295–2326.

Denning, Jeffrey T., Eric R. Eide, Kevin J. Mumford, Richard W. Patterson, and Merrill Warnick (2022). "Why Have College Completion Rates Increased?" In: *American Economic Journal: Applied Economics* 14.3, pp. 1–29.

Dynarski, Susan (2008). "Building the Stock of College-Educated Labor". In: *The Journal of Human Resources* 43.3, pp. 576–610.

Flory, J. A., A. Leibbrandt, and J. A. List (2015). "Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions". In: *The Review of Economic Studies* 82.1, pp. 122–155.

Giustinelli, Pamela (2022). "Expectations in Education: Framework, elicitation, and evidence". In: *SSRN Electronic Journal*.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko (2006). "The Homecoming of American College Women: The Reversal of the College Gender Gap". In: *Journal of Economic Perspectives* 20.4, pp. 133–156.

Lee, David S. and David Card (2008). "Regression discontinuity inference with specification error". In: *Journal of Econometrics* 142.2, pp. 655–674.

Lindo, Jason M., Nicholas J. Sanders, and Philip Oreopoulos (2010). "Ability, Gender, and Performance Standards: Evidence from Academic Probation". In: *American Economic Journal: Applied Economics* 2.2, pp. 95–117.

Möbius, Markus M., Muriel Niederle, Paul Niehaus, and Tanya S. Rosenblat (2022). "Managing Self-Confidence: Theory and Experimental Evidence". In: *Management Science*.

Murnane, Richard J. (2013). "U.S. High School Graduation Rates: Patterns and Explanations". In: *Journal of Economic Literature* 51.2, pp. 370–422.

Niederle, Muriel and Lise Vesterlund (2007). "Do Women Shy Away From Competition? Do Men Compete Too Much?" In: *The Quarterly Journal of Economics* 122.3, pp. 1067–1101.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales (2015). *Taste for Competition and the Gender Gap Among Young Business Professionals*. Cambridge, MA.

Thaler, Michael (2021). "Gender differences in motivated reasoning". In: *Journal of Economic Behavior & Organization* 191, pp. 501–518.

Wasserman, Melanie (2021). "Up the Political Ladder: Gender Parity in the Effects of Electoral Defeats". In: *AEA Papers and Proceedings* 111, pp. 169–173.

— (2022). "Gender Differences in Politician Persistence". In: *Review of Economics and Statistics, forthcoming.*

Chapter III

Non-Compete Agreements and Labor Allocation Across Product Markets

Non-Compete Agreements and Labor Allocation Across Product Markets

Clemens Mueller[1]

Abstract

I analyze the effect of non-compete agreements (NCAs) on career trajectories of inventors in the US. NCAs constrain the within-industry employment choice set of inventors. I show causal effects that 1.5 in 100 inventors annually bypass their NCAs by moving to new employers in more distant product markets. Reallocated inventors are subsequently less productive. Inventors move to new employers which are less reliant on NCAs and there is a lower quality match between inventors and their new employers. Firms affected by labor outflows grow less whereas firms with labor inflows grow more. I highlight regulatory frictions which lead to unintended detrimental reallocation of human capital in the economy.

[1]Comments highly appreciated. Please contact Clemens Mueller at clemens.mueller@uni-mannheim.de. I would like to thank Anthony Cookson (discussant), Matt Marx, Ernst Maug, Daniel Metzger (discussant), Maurizio Montone (discussant), Alexandra Niessen-Ruenzi, Alison Schultz, and Jiri Tresl for helpful comments. I also want to thank conference participants at the Northern Finance Association, American Economic Association (poster session), Corporate Finance Day, IZA Workshop on Labor Market Institutions, and the Swiss society for Financial Market Research. I also want to say thanks to seminar participants at Erasmus University Rotterdam, and at the University of Mannheim. This version subsumes an earlier draft with the title "How Reduced Labor Mobility Can Lead to Inefficient Reallocation of Labor". I acknowledge support of the high performance cluster computing facilities and the German Research Foundation (DFG) through grant INST 35/1134-1 FUGG.

3.1. Introduction

Non-compete agreements (NCA) are covenants which restrict employees from working for competitors during and after employment. Employers commonly use NCAs to retain valuable human capital within firm boundaries and to protect trade secrets. There is an ongoing debate in economics and finance about benefits and drawbacks of these agreements,[1] and the FTC, on January 5th 2023, has proposed a ban on NCAs.[2] On the one hand, NCAs can benefit employees, because of increased incentives for employers to retain and invest in employees' human capital (Garmaise 2011). However, the cost is reduced wages (Lipsitz and Starr, 2021) as well as lower labor mobility (Marx, Strumsky, and Fleming 2009).

In this paper, I add a novel and important dimension to the literature: product markets. Non-compete agreements effectively constrain the within-industry employment choice set of inventors. Inventors who want to move to a new employer thus face the following trade-off: either 1) terminate the employment contract and wait until the NCA expires to be able to move to a competitor or 2) "bypass" the NCA and immediately work for a new employer, however in a more distant product market. The evidence provided in this paper supports the existence of this trade-off and extensively analyzes consequences of the latter.

I use data of around 600,000 US corporate inventors from 1976 to 2018. Patent data provides a suitable laboratory to study NCAs and allocation of labor for several reasons: First, patents provide the precise location of inventors and as patent ownership rights are assigned to their employers, they provide detailed employment histories. Second,

[1] See among others (Chen et al., 2022; Shi, 2023; Garmaise, 2011; Starr, 2019; Marx and Fleming, 2012; Samila and Sorenson, 2011; He, 2021).

[2] https://www.ftc.gov/news-events/news/press-releases/2023/01/ftc-proposes-rule-ban-noncompete-clauses-which-hurt-workers-harm-competition

corporate employers of these inventors provide measures of industry affiliation. Third, inventors are highly skilled individuals and, as such, are likely affected by NCAs. Fourth, patent data provides measures of a technology dimension as well as a time series measure of productivity (e.g. citations received and the economic value based on employers' market reactions to patent grants) on a granular level.

Staggered changes of NCA enforceability across U.S. states provide variation for estimating causal effects. In a staggered difference-in-differences event time regression, increases in NCA enforcement are positively related to the probability that an inventor moves to a more distant product market. In terms of economic magnitude, on average 1.5 out of 100 inventors move to another industry *per year*, an increase of 35%. These results hold using several industry definitions such as SIC and NAICS codes as well as textual-based definitions of product markets. The baseline regression uses inventor and year fixed effects, and thus exploits the staggered timing of 9 NCA enforcement increases across states either in the form of precedent-setting court cases or state laws. There is no effect for decreases in NCA enforcement.

Econometric theory provides guidance on the event study design: I compare treated inventors (i.e. those exposed to an increase in NCA enforceability) to never-treated in an event time framework (Baker, Larcker, and Wang 2022; Borusyak, Jaravel, and Spiess 2021, Chaisemartin and d'Haultfoeuille 2021, Callaway and Sant'Anna 2021, Sun and Abraham 2021). I match treated inventors to control inventors based on their quality as measured by number of patents and the number of citations received, as well as the technology they patent in. Inventors move to more distant product markets after NCAs become more enforceable. Consistent with a causal interpretation of the results, there are no pre-trends.

It would be problematic, if the introduction of state-level legislation is due to economic and potentially unobserved reasons. I address the potential endogeneity of state-level shocks by using within state-year variation in the *intensity* in treatment. Specifically, I construct a firm-level proxy, based on 10-K and 10-Q filings, whether an employer heavily relies on NCAs. If inventors indeed bypass their NCA and move to more distant product markets, then inventors employed at firms which heavily rely on such agreements should be more affected. I include state-year fixed effects, and show that the effect is confined to inventors whose employers do rely on NCAs. This is in line with a causal interpretation of the results.

The effect is confined to inventors with more available outside opportunities. Inventors who move after an increase in the enforcement of NCAs subsequently work for firms that are less likely to rely on NCAs. Inventors thus seem to avoid NCAs in their future employment.

The natural follow-up question to ask is: What is the effect of NCA-constrained reallocation on the productivity of inventors, measured by the economic value of patents and citation-weighted patents? On one hand, it might be beneficial to society if increased inter-industry mobility leads to more idea recombination, and thus more innovation. On the other hand, inventors might perform worse after a NCA-constrained industry move. In a difference-in-differences analysis, those inventors who move (i.e. leave) to more distant product markets subsequently perform worse compared to those who do not (i.e. stay). I compare all inventors who are affected by more enforceable NCAs, however one subgroup decides to stay and another leaves to more distant product markets. This result thus does not allow for a causal interpretation as it relies on a revealed choice. Those inventors who stay patent with similar quality before and after, however those inventors who move subsequently perform worse. There is little evidence of negative selection into moving: inventors who move and those who stay are virtually identical and patent with similar quality before. Only afterwards a performance gap emerges.

To shed some light on this finding, I subsequently analyze what characterizes the observed industry mobility. I compare NCA-constrained to other, unconstrained, industry movers. I calculate a measure for matching quality between inventors and their new employers based on patent technologies. The technological similarity between inventor and her new employer is reduced by 20% after an increase in NCA enforcement. Regulatory frictions in the form of NCA enforcement and the associated limited choice set of inventors thus leads to a lower matching quality in the labor market.

NCAs usually expire 1-2 years after the termination of the employment contract. I find evidence of the existence of the trade-off to either move immediately after contract termination to a firm which is further away in the product market or to terminate the employment contract and wait until the NCA expires to join a close competitor. The duration between two employment spells increases after an increase in NCA enforcement, especially for inventors who move to close industry competitors.

Generally, unconditional across industry mobility is associated with *higher* future productivity. Inventors are subsequently even more productive when there is a high product market as well as technology similarity. From a social planner point of view, to the extent employers retain incentives to invest in their human capital, regulation should therefore foster inventor mobility of closely related employers.

I analyze firm-level effects and show that inflows of human capital due to NCAs are associated with higher future firm productivity growth. Outflow of human capital is associated with lower future firm growth. NCAs thus not just shape career trajectories of inventors but also have a first order effect on firm boundaries and firm productivity.

The results emphasize an important distinction between ex-ante and ex-post effect of labor market regulation. Ex-ante, NCAs are designed to incentivize employers to invest in their employees. Ex-post however, NCAs create a hold-up problem and shift bargaining power to employers. Inventors cannot credibly threaten to move to another firm and retain their industry-specific human capital. It might thus be optimal for them to leave and retain a higher share of their productivity output.

This paper contributes to several strands of literature. First, on real effects of labor market frictions (Bena, Ortiz-Molina, and Simintzi, 2021; Shen, 2021). Previous research has shown that NCAs lead to lower labor mobility (Fallick, Fleischman, and Rebitzer, 2006; Marx, Strumsky, and Fleming, 2009; Jeffers, 2017; Garmaise, 2011; Balasubramanian et al., 2020), as well as a brain drain of enforcing states (Marx, Singh, and Fleming, 2015). In contrast to lower labor mobility, by focusing on a product market dimension, I instead show *higher* labor market mobility. The paper is thus closely related to Marx (2011), who provides survey evidence consistent with the empirical results presented in this paper. My setting allows to analyze long run employment outcomes and an important outcome for society: productivity of labor, in this context innovation output. This paper is closely related to two theoretical papers on NCAs. Chen et al. (2022) theoretically and empirically argue that current regulatory restrictions are near optimal for growth. Shi (2023) on the other hand suggests that a complete ban on NCAs is the optimal policy.

I also add to the allocation of labor literature (Babina, Ouimet, and Zarutskie, 2020; Babina, 2020; Hombert and Matray, 2017; Hombert and Matray, 2018; Hacamo and Kleiner, 2022)). I show how labor market frictions can lead to some reallocation of labor in the economy, which is likely an unintended consequence for policy makers in the context of NCAs. Lastly, I add to the literature on firm and industry boundaries (Seru, 2014). NCAs have profound impact on career choices of employees, shape firm boundaries, and affect firm productivity. While unconstrained inter-industry mobility seems to be beneficial for society, NCA-constrained industry mobility is detrimental.

3.2. Data

3.2.1. Employment Histories of Corporate Inventors

I obtain data on corporate innovation from 1976 until 2020 from two sources. I obtain patents matched to firms from Kogan et al. (2017), commonly referred to as KPSS. This list is complemented with the DISCERN database of Arora, Belenzon, and Sheer (2021).[3] The first dataset is thus a list of patent numbers and an associated unique corporate identifier.

The next step is to match individual inventors to these patents. The United States Patent and Trademark Office (USPTO) provides detailed data on patents such as who invented which patents, the location of the inventor, and the application year which is used to proxy for innovation generation. Most importantly, the USPTO provides disambiguated inventor-level data.[4] Disambiguated data allows researchers to track individual inventors over time. I obtain this data from patentsview.org.

[3] The KPSS data with matched patent data is updated until the end of 2020 and available here: https://github.com/KPSS2017/Technological-Innovation-Resource-Allocation-and-Growth-Extended-Data; The DISCERN database includes patents matched to firms (including subsidiaries) until 2015 and is available here: https://zenodo.org/record/4320782

[4] The provided data builds on previous efforts such as the NBER patent citation data file as well as disambiguated inventor-level data of Li et al. (2014).

3.2.2. Institutional Details and Data on Enforcement Changes

A non-compete agreement usually puts limitations on industry, geographic reach (sometimes specified, and ranges between a well defined radius to a state, country or even worldwide), and duration (usually 1-2 years) of an employee. The Appendix lists some examples of NCAs obtained from firms' 10-K or 10-Q. Microvision states in the annual statement that the firm heavily relies on NCAs. In an appendix to a 10-Q, Nuance Communications explicitly mentions that they prohibit employees "from working for an employer which is engaged in activities or offers products that are competitive with the activities and products of the company."

I summarize changes in state-level NCA enforcement in table 3.1. I rely on Ewens and Marx (2018), who provide an extensive discussion on court rulings and legislative changes from 1985-2016.[5] Kini, Williams, and Yin (2021) is the second source of data. They extend a score of NCA enforceability across states originally developed by Garmaise (2011) to the years 1992-2014.

[5]The data is available here: https://github.com/michaelewens/Non-compete-Law-Changes

Table 3.1.: Overview of State-Level Changes in Non-Compete Enforceability
This table provides an overview of changes of enforceability of NCAs. The changes are based on Ewens and Marx (2018) as well as Kini, Williams, and Yin (2021). Ewens and Marx (2018) gather data from Malsberger, Brock, and Pedowitz (2016) and consult lawyers. Kini, Williams, and Yin (2021) extend a score of NCA enforceability across states originally developed by Garmaise (2011) to the years 1992-2014. To do so, they use data provided by the law firm Beck Reed Riden LLP. Those two sources together are a comprehensive list of changes during the years 1985-2016. Panel A includes states that increased the enforceability of NCAs. Panel B includes decreases. Panel C includes states that had several changes in the enforceability of NCAs. Brackets in Panel C indicate the direction of the change, (+) equal to an increase in enforceability.

State	Case	Year
Panel A: Increase of Non-Compete Agreement Enforcement		
AL	Alabama legislature	2016
AR	Arkansas legislature	2016
FL	Florida legislature	1996
GA	Georgia legislature	2011
ID	Idaho legislature	2008
MI	Michigan legislature	1985
OH	Lake Land v. Columber	2004
VT	Summits 7 v. Kelly	2005
VA	Assurance Data Inc. v. Malyevac	2013
Panel B: Decrease of Non-Compete Agreement Enforcement		
MT	Wrigg v. Junkermier	2009
NH	New Hampshire legislature	2011
NV	Golden Rd. Motor Inn. v. Islam	2016
OR	Oregon legislature	2008
SC	Poynter Investments v. Century Builders of Piedmont	2010
UT	Utah legislature	2016
Panel C: Repeated In-/Decreases of Non-Compete Agreement Enforcement		
CO	Luncht's Concrete Pumping v. Horner (+)	2011
CO	see Kini, Williams, and Yin (2021) (-)	2013
IL	Fire Equipment v. Arredondo (+)	2011
IL	Fifield v. Premier Dealership Servs. (-)	2013
KY	Gardner Denver Drum v. Peter Goodier and Tuthill (+)	2006
KY	Creech v. Brown (-)	2014
LA	Shreveport Bossier v. Bond (-)	2001
LA	Louisiana legislature (+)	2003
TX	Light v. Centel Cellular (-)	1994
TX	Baker Petrolite v. Spicer (+)	2006
TX	Mann Frankfort Stein & Lipp Advisors v. Fielding (+)	2009
TX	Marsh v. Cook (+)	2012
WI	Star Direct v. Dal Pra. (+)	2009
WI	Runzheimer International v. Friedlen (-)	2015

What happens when NCAs are more enforceable? Restrictions included in a NCA and what is ultimately enforceable can differ. California does not allow the use of NCAs. Florida is on the other end of the spectrum and enforces NCAs most strictly. Often, NCAs are enforceable conditional on passing a ”reasonableness” test. After a 1996 legislative change, NCAs in Florida need to protect "legitimate business interests" in order to be enforceable. This clarified previous uncertainty and shifted power towards employers.[6]

For some specifications, I use data on firm-level reliance on NCAs. I proceed in similar fashion as Kini, Williams, and Yin (2021). First, I obtain form 10-K and form 10-Q filings from EDGAR. I parse and strip the text of figures, pictures and html tags. I obtain identifiers from historical Compustat from WRDS servers, as well as a historical CIK-CUSIP mapping.[7] Form 10-K and form 10-Q filings commonly include NCAs of senior employees at a firm. I use the information to construct a panel of US corporations with an indicator variable equal to 1 if the corporate employer mentions the use of a NCA either in an executive/board contract or mentions the reliance on NCAs in the annual statement. I do this similar to Acikalin et al. (2022) and screen for instances of ”non-compete agreement”, ”covenant not to compete”, etc. I compute a panel on a firm-year level and construct a dummy variable equal to one if a firm relies on NCAs. This panel is comprehensive from the year 1996 onwards. I compare the frequency of NCA use with the literature. In my sample, 54% of firms rely on NCAs.[8] This is close

[6]There are many other examples on how NCAs become more enforceable. For example, the Ohio Supreme Court decided in 2004 that a sufficient consideration to uphold a NCA was continued employment. Another example is Idaho, which changed to a so-called ”blue pencil” rule where a judge can modify the contract to make it more reasonable whereas in other states one invalid part of a NCA renders the whole agreement void. Interested readers should refer to Marx and Fleming (2012) for history and background literature. Ewens and Marx (2018) provide extensive details on individual court cases and legislative changes

[7]Ekaterina Volkova provides this mapping here: https://sites.google.com/view/evolkova/data-cik-cusip-link

[8]This data is available to download on the authors website.

to previous survey and empirical evidence. To compare, Starr, Prescott, and Bishara (2021) find that almost one fifth of all employees in the US are bound by NCAs. The share of NCAs for technical workers is around 50% (Marx, 2011), 62.5% for CEOs with employment contracts (Kini, Williams, and Yin, 2021), and 70% for corporate executives (Garmaise, 2011).

3.2.3. Sample Construction and Descriptive Statistics

The sample construction starts with all corporate innovation from the two sources mentioned previously. This gives a mapping with a unique identifier for each corporation and the patent number assigned by the USPTO. In principle, data on corporate patents is available from 1926, however the USPTO provides digitized patent information with disambiguated inventor data from 1976 onwards, which marks the start of the sample. In a next step, I merge the inventors of all corporate-owned patents with the disambiguated inventor data. The resulting dataset is a panel at the inventor-year level.

I identify industry employment changes as follows: The inventor files two subsequent patent applications for a different employer with a different industry affiliation. I follow the previous literature (Song, Almeida, and Wu 2003; Marx, Singh, and Fleming 2015) and use the yearly midpoint between two subsequent patents to proxy for the year of employment change.[9] The application year rather than the grant year is used, in order to have a more timely measure of innovation creation[10] and employment changes. I remove inventors from the sample who only patent once in the sample period. All regressions

[9]Patent-based measures of employment histories thus include measurement error. On average, there is a gap of 0.9 years between two subsequent patents filed by the same inventor. The median number of years between two filings is zero. When alternatively limiting the sample to patent filings with at most one year between two subsequent patents, the results become stronger.

[10]This avoids a lag between applying for and being granted a patent, which is 4 years at the median.

include inventor fixed-effects, so these inventors would not provide any meaningful variation on labor market employment.

Innovation is an ideal laboratory for several reasons: First, the universe of corporate patenting in the last 40 years provides tractable employment histories of inventors based on granted patents.[11] In the context of this paper, it also seems plausible that highly skilled human capital such as inventors, are likely affected by NCAs.

Second, patent documents also capture the location (on a city level) of each inventor listed on a patent. This greatly improves measurement for empirical research that uses location-based variation in treatment. Previous studies often proxy for location using the headquarter location of the employer.

Third, corporate innovation data allows to look at two distinct but related dimensions: measures of product and technology similarity. Product markets for employers are readily available as SIC and NAICS industry codes, as well as text-based industry classifications following Hoberg and Phillips (2010) and Hoberg and Phillips (2016). The latter is a measure with desirable econometric properties which can be used to measure the similarity between the old and the new employers of inventors. Patent data provides technology classifications of every patent (e.g. CPC, WIPO, IPC). This is useful as it allows researchers to compute technology similarities between the patents of inventors and their employers.

Fourth, and lastly, patent data provides a useful metric on a patent basis to measure the productivity of an inventor over time. A researcher can thus observe the number of

[11] The caveat here is that non-patented innovation is unobserved and thus overall labor mobility is likely underestimated

patents, the number of citations received[12] (Lerner and Seru 2021), and the economic value of patents (Kogan et al. 2017). The latter measure is available for all patents granted until 2020 and is comprised of a USD value on a patent basis. The measure is calculated using stock market reactions of listed patent assignees on the grant day of a patent. This measure is available before and after an employment change.

Table 3.2 shows descriptive statistics. The timeframe is from 1976-2018. In total, the matched sample includes 436,382 inventor-year observations. This includes data of around 1.8 million patents of roughly 600,000 inventors. The sample includes 6,345 listed firms as employers. An industry move, defined on a SIC 4-digit industry, appears in 4% of observations. I compare this to the previous numbers in the literature such as Melero, Palomeras, and Wehrheim (2017) who show based an patent application data, that inventors move employers (without considering industries) at a rate of 10% per year. The mean number of patents granted is 5.5 and the number of truncation adjusted citation-weighted patents is 9.8.

Table 3.2.: Summary Statistics
The unit of observation is on an inventor-year level. Variable definitions are provided in the Appendix.

Variable	N	Mean	SD	Min	25%	50%	75%	Max
SIC-4 Industry Change	436,382	0.048	0.20	0	0	0	0	1
SIC-3 Industry Change	436,382	0.042	0.19	0	0	0	0	1
NAICS-6 Industry Change	436,382	0.044	0.19	0	0	0	0	1
NAICS-5 Industry Change	436,382	0.042	0.19	0	0	0	0	1
Employer NCA	322,896	0.49	0.50	0	0	0	1	1
ln(1+Economic Value of Patents)	436,382	0.99	1.46	0	0	0	1.98	9.84
ln(1+Citation-Weighted Patents)	436,382	0.36	0.69	0	0	0	0.37	9.78
Inventor Number Patents	436,382	5.55	13.06	0	1	2	5	1,805
Inventor Total Citations	436,382	9.78	94.23	0	0.25	1.80	6.86	94,891

[12]Newer patents mechanically have less time to accumulate citations than older patents. In order to mitigate this problem I follow Hall, Jaffe, and Trajtenberg (2005), Dass, Nanda, and Xiao (2017), and Lerner and Seru (2021). When using citations as a measure of innovation output, I adjust all cumulative citations received until June 2022 and perform a truncation adjustment by adjusting with respect to year and technology class.

3.3. Staggered State-Level Changes in Non-Compete Enforcement

3.3.1. Event Study and Dynamic Effects

I estimate the following event study regression:

$$IndustryChange_{i,t+1} = \sum_{k=-5}^{k=+10} \delta_k \times D_k + \sum_{k=-5}^{k=+10} \beta_k \times D_k \times NCAIncrease_{s,t} + \theta_i + \phi_t + \epsilon_{i,t} \tag{3.1}$$

where D_k are time dummies relative to the NCA enforcement increase, where i represent inventor i, located in state s, in year t. The dependent variable IndustryChange is equal to one if an inventor moves between two firms with different 4-digit SIC industry codes. The variables θ and ϕ are inventor and year fixed-effects, respectively. Year fixed-effects account for year-specific shocks to mobility. Inventor fixed-effects control for time-invariant unobserved factors on the inventor level.

The coefficients of interest are β_k which capture the treatment indicator interacted with 4 pre-treatment dummies and 10 post-treatment dummies. All coefficients, if feasible, are estimated relative to one year before treatment.

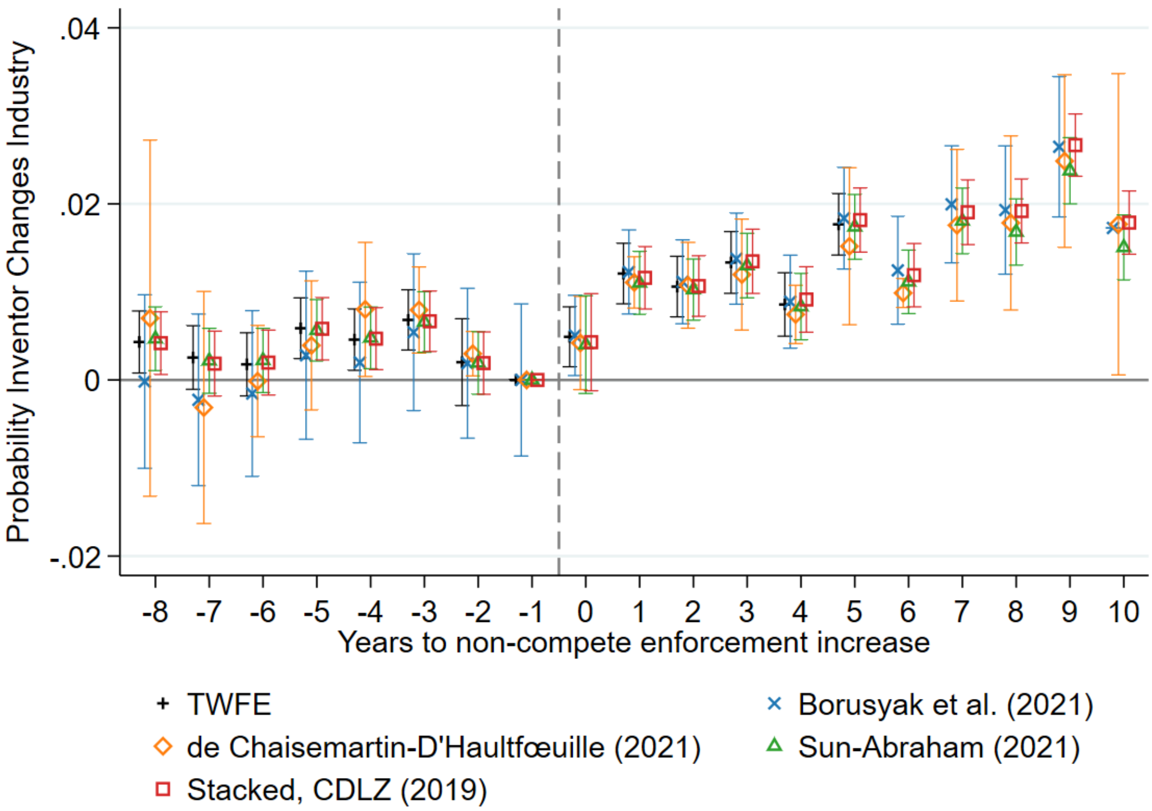
I use nearest neighbor matching to compare treated and control inventors. I match inventors based on year of activity (whether they are currently employed at a firm), lagged number of patents, and lagged total citations. I use these two variables to match inventors of a similar quality. I also include patent technology to guarantee that treatment and control inventors are exposed to similar technological shocks. I match the three nearest neighbors with replacement using the Mahalanobis distance. The analysis includes inventor as well as year fixed effects. I cluster standard errors on the inventor and year level.

A two-way fixed effect estimation of a staggered difference-in-differences design are weighted averages of all possible two-group difference-in-differences estimators (Goodman-Bacon 2021). A potential problem are dynamic treatment effects when we compare early-treated to late-treated inventors (Baker, Larcker, and Wang 2022). I follow recent econometric theory to set up the panel of inventors when using state-level variation in treatment of table 3.1. I compare treated with never-treated inventors. Thus, I compare inventors based in states that experienced increased enforcement of NCAs with clean controls: those inventors who did not experience any changes during the sample period. I use a number of recently proposed estimators such as Borusyak, Jaravel, and Spiess (2021), Chaisemartin and d'Haultfoeuille (2021), Callaway and Sant'Anna (2021), and Sun and Abraham (2021).

Figure 3.1 visualizes the results from Equation 3.1. The probability that an inventor changes industries increases in the first treatment year and we subsequently see a steady increase over time. On average, 1.5 inventors out of 100 move across industries per year, which is a 35% increase in the probability (mean value of SIC 3-digit mobility = 0.043). The alternative estimators are close to the OLS estimates. In the Appendix, I show that there is no effect when looking at decreased NCA enforcement.

Figure 3.1.: Staggered State-Level Increases in Non-Compete Agreement Enforcement: Event Study and Dynamic Effects

This figure reports the result of the difference-in-differences event study of equation 3.1. The sample is on an inventor-year level. The figure plots the coefficients of pre and post time dummies, interacted with a treatment indicator equal to one if the state increases NCA enforcement. The y-axis shows the coefficient on a regression on the variable $IndustryChange$, which is a dummy variable equal to one if the inventor moves to a firm in a different SIC 4-digit industry in that year. The sample compares treated to never-treated inventors. Inventors are matched based on employment year, number of patents, number of citations and patent technology class. I match the three nearest neighbors with replacement using the Mahalanobis distance. Variable definitions are provided in the Appendix. All regressions include Inventor and Year fixed effects. Standard errors are clustered by Inventor and Year. Confidence intervals are at the 5% level.



3.3.2. Non-Compete Agreements and Product Market Similarity

The previous analyses rely on standard, fixed industry classifications such as SIC codes. In the following, I analyze whether the results generalize to a continuous version of industry similarity between two firms. I will rely on the textual based industry scores of Hoberg and Phillips (2016). This provides several improvements, such as 1) the industry definitions are not fixed over time and a continuous measure can vary between two identical firms across years, 2) the measure captures product market proximity regardless of whether two firms are in the same industry or not. Standard classifications can only provide a 0 or 1, which means either two firms are in the same industry or they are not. The regression analyzes the question: Are inventors moving to employers which are further away from their old employers after an increase in NCA enforcement? Formally, I run the following regression:

$$y_{i,t} = \beta \times NCAIncrease_{i,t} + \phi_t + \epsilon_{i,t} \tag{3.2}$$

where $y_{i,t}$ is the product market similarity between the previous and the new employer obtained from Hoberg and Phillips (2016). $NCAIncrease$ is a dummy variable equal to one if the inventor is exposed to an increase in NCA enforcement. The sample is thus composed of all inventor mobility events. An inventors move is included in this regression as long as the inventor is based in the US and moves between two publicly listed firms with available data.

The results are shown in table 3.3. Indeed, inventors exposed to increased NCA enforcement move to firm that are on average around -1.4% less similar in product market similarity. To put this into context, within the universe of all inventors mobility events, the average product market similarity is equal to 6.8%. An increase in NCA enforcement thus leads to inventors moving to a firm that is 21% less similar in the product market compared to other inventor mobility events.

Table 3.3.: Increased NCA Enforceability and Product Market Similarity
This table reports the result of equation 3.2. The dependent variable is the textual similarity measure of Hoberg and Phillips (2016). The measure captures the similarity between the former and the new employer of each inventor mobility event. $NCAIncrease$ is a dummy variable equal to one if the inventor experienced an increase in NCA enforcement. Variable definitions are provided in the Appendix. The regression includes Year fixed effects. Standard errors are clustered by Year. t-statistics are displayed in parenthesis. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively.

Dependent variable:	Product Market Similarity
$NCAIncrease$	-0.014***
	(-6.52)
Observations	126,124
R-squared	0.04
Year FE	YES

3.3.3. Within State-Year: Is the Effect Stronger in the Presence of Non-Compete Agreements?

If NCA enforcement increases indeed lead to increased inter-industry mobility of inventors, then we would expect this effect to be stronger for inventors who are in fact bound to a NCA. Unfortunately individual level NCAs of inventors are unobserved. However, employers might differ on how much they rely on NCAs. I therefore compute a proxy on a firm level as follows: First, I obtain all annual and quarterly (10-K and 10-Q) reports of the employers in the sample from 1996-2018. These filings often include contract information and NCAs of senior employees. I compute a dummy equal to one if a firm relies on NCAs. The assumption is that to some extent, this firm-level dummy is a proxy for the presence of NCAs on an inventor level.

I formally test whether increased enforcement of NCAs leads to more industry mobility especially for those inventors employed at firms that use NCAs. For this purpose, I run a triple difference-in-differences regression as follows:

$$\begin{aligned} IndustryChange_{i,s,j,t+1} = \beta \times NCAIncrease_{s,t} \times Post_{s,t} + \\ \delta \times NCAIncrease_{s,t} \times Post_{s,t} \times EmployerNCA_{j,t} + \theta_i + \phi_t + \epsilon_{i,s,j,t} \end{aligned} \tag{3.3}$$

where $EmployerNCA$ is an indicator variable equal to one if the employer heavily relies on NCAs. The parameter of interest is the triple interaction term $NCAIncrease \times Post \times EmployerNCA$. The variable is equal to one only for inventors in years after an increase in NCA enforcement, and additionally employed at firms who rely on NCAs.

Table 3.4 shows the results. The triple difference-in-differences term is positive and significant throughout. In economic terms, inventors in years following treatment and employed by NCA-relying firms experience an increase in industry mobility of 1.6%. The observed effect seems to be confined to inventors who are likely bound by NCAs. Subject to the constraint that the proxy for NCA on an employer level is imperfect, this is aligned with a causal interpretation of the results.

The regression includes $State \times Year$ fixed effects, as well as Inventor fixed effects, which absorb many of the included interaction terms. The standard errors in this regression are clustered on an inventor level, however different levels of clustering, such as state or state-year do not change the results.

Table 3.4.: Triple difference-in-differences: Inventors Employed at NCA Firms

This table reports the triple-difference-in-differences fixed effect panel regression of equation 3.3. The sample is on an inventor-year level. $IndustryChange_{t+1}$ is a dummy variable equal to one if the inventor moves to a firm in a different industry. $Treat$ is a dummy variable equal to 1 if the state increased the enforceability of NCAs. $FirmNCA$ is a dummy variable equal to one if the firm relies on NCA. This variable is obtained from 10-K and 10-Q filings where firms mention the use of NCA or senior level employee contracts are filed on EDGAR. In column (1) industry is defined on a SIC 4-digit level, in column (2) on a SIC 3-digit level, in column (3) on a NAICS 6-digit level and in (4) on a NAICS 5-digit level. Variable definitions are provided in the Appendix. All regressions include Inventor, as well as $State \times Year$ fixed effects. Standard errors are clustered by $State \times Year$. t-statistics are displayed in parenthesis. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively.

Dependent variable:	$IndustryChange_{t+1}$			
	(1)	(2)	(3)	(4)
$Treat \times Post \times FirmNCA$	0.016***	0.016***	0.013***	0.012***
	(3.64)	(3.82)	(2.70)	(2.61)
Observations	308,517	308,517	308,517	308,517
R-squared	0.13	0.13	0.13	0.13
Industry Definition	SIC 4-digit	SIC 3-digit	NAICS 6-digit	NAICS 5-digit
Inventor FE	YES	YES	YES	YES
$State \times Year$ FE	YES	YES	YES	YES

3.3.4. Does Increased Non-Compete Enforcement *Cause* Industry Mobility?

In order to interpret the results as causal, the critical assumption is that treatment and control inventors are equally likely to change industries in the absence of treatment. As a necessary but not sufficient condition, I visually assess whether treated and control inventors experience parallel pre-trends before the treatment. Reassuringly, the event study in figure 3.1 shows that this is the case.

There is little evidence that the treatment effect is immediate in the very first year. There are several reasons why we should not necessarily expect this: For example, the Florida law change in 1996 was explicitly only applicable to contracts signed after July 1, 1996.[13] This would mean that only employees who start working after this date are exposed to increased NCA enforcement. To increase the chances of legal protection, Ewens and Marx (2018) note that employers commonly require their employees to sign updated employment contracts, which might not lead to immediate responses. This is supported (for the Georgia 2010 case) by Ewens and Marx (2018) who interviewed an employment attorney, who stated: "when the new law went into effect (including our firm), many employers revised their employment and restrictive covenant agreements to take advantage of the law".

Setting the legal point of view aside, there are additional considerations for a delayed response from the point of view of employees. Inventors willing to move might not be well aware of the details of their NCA. They might learn about the increased enforcement of NCAs years after. There is no reason we should expect sudden effects, but rather an increase over time which leads to a new equilibrium in the labor market.

Bishara (2011) extensively analyzes the legal background on the enforceability on non-

[13] However Ewens and Marx (2018) note that continued employment suffices as consideration.

compete agreements. He notes that it can be difficult to predict the consequences for a departing employee when she joins an out-of-state competitor. It is thus often an open question to what extent individual non-compete agreements are in fact enforceable and there is uncertainty involved in the variation I use. The observed effects are thus best seen as the effects of subjective employee behavior rather than clear-cut labor regulatory constraints.

A potential problem for a causal interpretation is whether state legislative changes are correlated with other factors that determine industry mobility. State legislative changes might be problematic if the desired policy change is anticipated. There are two reasons why this is unlikely to be a threat to identification in my setting. First, Jeffers (2017) shows that the state-level shocks are unrelated to macroeconomic conditions and cannot be easily predicted. Given the focus on inter-industry mobility, the positive effect on industry changes of inventors is a plausible unintended consequence of regulatory changes. Overall, the findings are consistent with interview evidence of Marx (2011), where employees admit to taking career detours given that their NCA prohibited them from working in similar industries for the next 1-2 years. Marx (2011) interviewed one speech recognition professional who left the industry after being fired by his co-founder. "Well, if I'm ever gonna leave, what would I do for 2 years if I couldn't do speech recognition?"

3.3.5. Heterogeneity: Outside Opportunities

The results of industry mobility should be stronger if inventors have more outside options to choose from. I empirically test this hypothesis and split the sample at the median into employees which have relatively many industry competitors and those who have relatively few. To do so, I count the number of firms which exceed a certain threshold

(0.1) based on industry similarity scores of Hoberg and Phillips (2016).

The results are shown in table 3.5. As expected, inventors in industries with relatively more outside opportunities are much more likely to move. There is a positive coefficient, however statistically insignificant, for inventors employed in industries with relatively few outside opportunities.

Table 3.5.: Heterogeneity: Within Firm Opportunities
This table reports the results of equation 3.2. The sample is on an inventor-year level. $IndustryChange_{t+1}$ is a dummy variable equal to one if the inventor moves to a firm in a different industry. $NCAIncrease$ is a dummy variable equal to 1 if the state increased the enforceability of NCAs. The sample is split at the median of a proxy for outside opportunities for employees. I compute how many competitors surpass a fixed similarity threshold, which measures the possibilities for inventors to move to other employers. Column (1) includes employers with many many closely related firms. Column (2) includes employers with few closely related firms. Variable definitions are provided in the Appendix. t-statistics are displayed in parenthesis. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively.

Dependent variable:	$IndustryChange_{t+1}$	
	(1)	(2)
Sample:	Many Outside Opportunities	Few Outside Opportunities
$NCAIncrease \times Post$	0.015***	0.004
	(5.62)	(1.26)
Observations	124,050	141,491
R-squared	0.13	0.13
Inventor FE	YES	YES
Year FE	YES	YES

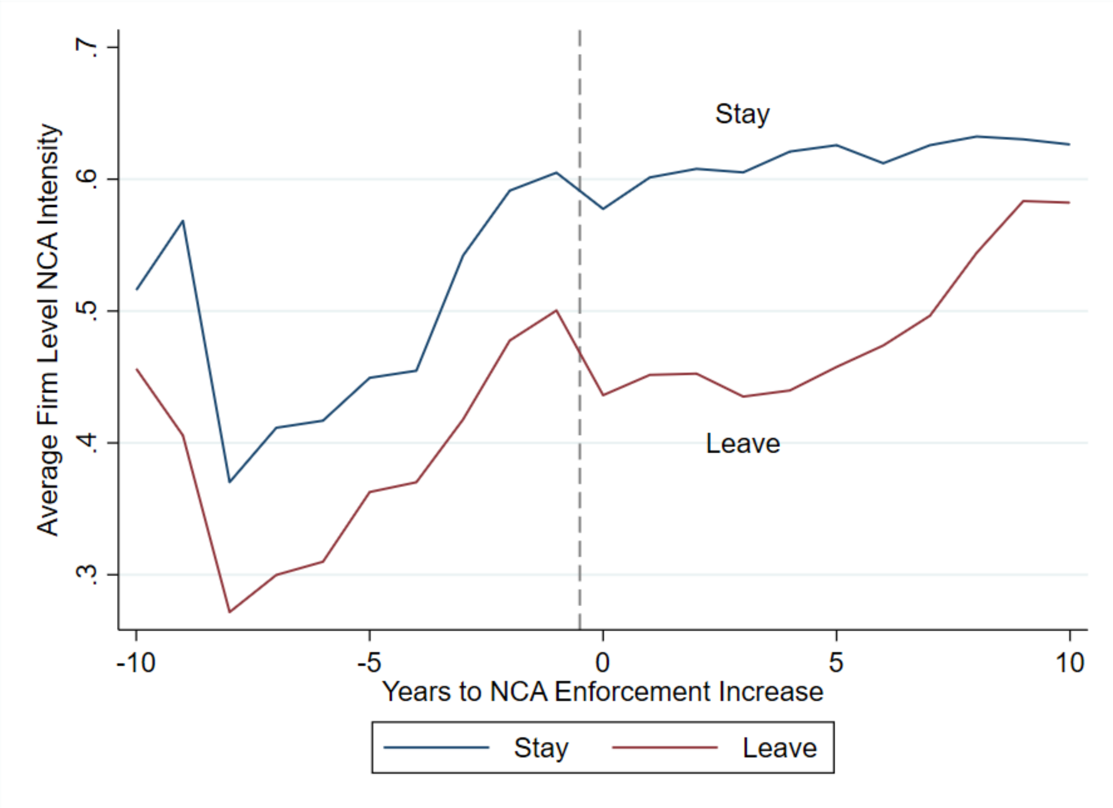
3.3.6. Inventors move to Employers which rely less on NCAs

In the following specification, I analyze whether inventors avoid future employers which rely on NCAs. To do this, I visualize the average presence of whether the employer relies on NCAs. This variable is again computed from firm balance sheet statements (10-Ks) or employment contracts obtained from 10-Q filings.

I analyze all inventors who are affected by an increase in NCA enforcement. However one group decides to move to another industry and others stay. The results are shown in figure 3.2.

A gap emerges after the state increases NCA enforceability. There are very similar pre-trends in the two groups. We do not see any meaningful changes before and after NCA enforcement changes for inventors who stay in their industry. However, inventors who move after an increase in NCA enforcement tend to move to new employers that rely less on NCAs.

Figure 3.2.: Employer NCA Intensity: Stay vs. Leave
This figure visualizes employer NCA intensity on an inventor-year level. Employer NCA intensity is a dummy variable equal to one if the employer explicitly mentions the use of NCAs in 10-Ks or 10-Qs. Time is relative to $NCAIncrease$, which is the year when the state increased NCA enforcement. The graphs are visualizing raw data. Inventors are assigned into two groups: those who move to another more distant product market (leave) and those who do not (stay). A line is drawn at x = -0.5, between -1, the last untreated year and 0, the first treatment year. Variable definitions are provided in the Appendix.



3.4. NCA-Constrained Industry Moves Lead to Lower Productivity

What are the effects on productivity if inventors move across industries in response to NCA enforcement increases? On one hand, it might be beneficial to society if increased inter-industry mobility leads to more idea recombination, and thus higher or more high quality innovation output. On the other hand, inventors might perform worse after a NCA-constrained industry move. For this purpose, I visually compare innovation output of inventors.

In a difference-in-differences style visualization, I compare those inventors who move to more distant product markets (leave) to those who do not (stay). All inventors in this specification are treated, e.g. affected by an increase in NCA enforcement. I compare those inventors who move to those who do not, which means that the difference-in-differences compared two groups of inventors based on a revealed choice. Thus the following analysis is unable to make causal inferences, and should therefore rather be seen as purely descriptive.

I plot annual research productivity of inventors in figure 3.3. We see a significant divergence in the quality of patents produced by affected inventors. The raw data is visualized in an event time framework, relative to an increase in NCA enforcement. Panel A shows the yearly economic value of patents of the inventor. Panel B shows citation-weighted patents.

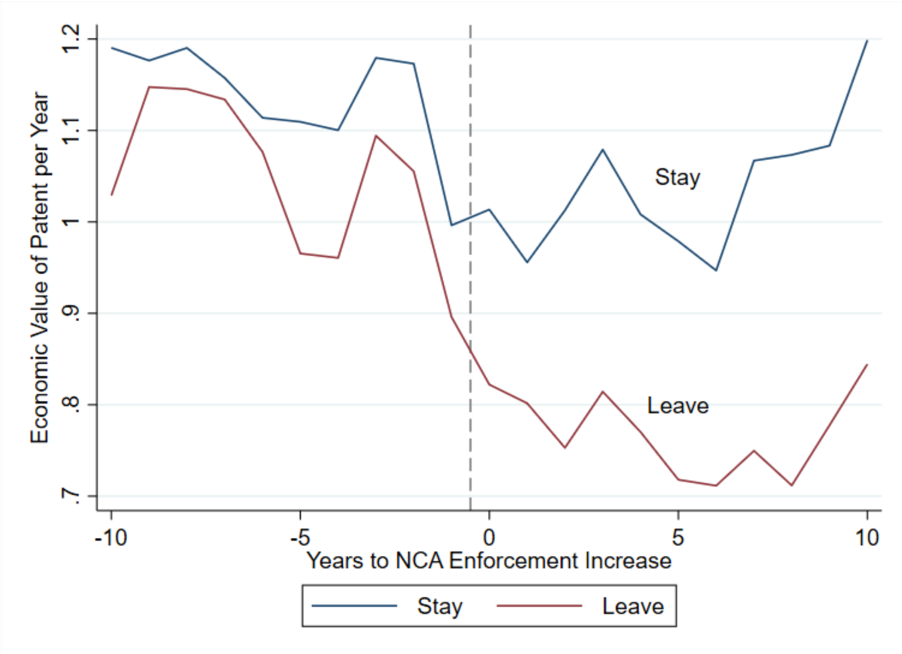
Those inventors who move to more distant product markets subsequently perform worse. Inventors who stay are unaffected and patent with similar quality before and after.

Importantly, there does not seem to be a negative selection into moving to a more distant product market: inventors who move and those who stay are virtually identical and patent with similar quality before an NCA enforcement increase. Only afterwards a performance gap emerges.

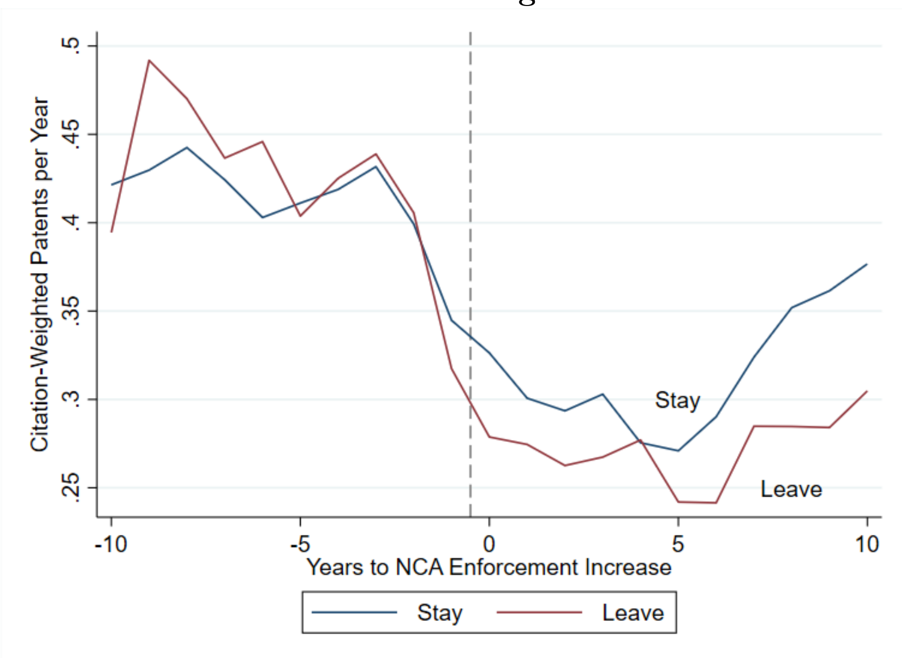
Figure 3.3.: Productivity of Inventors: Stay vs. Leave

This figure visualized innovation output on an inventor-year level. Innovation output is measured by the economic value of patents (stock market reaction to patent grants) in Panel A, and citation-weighted patents in Panel B. Time is relative to $NCAIncrease$, which is the year when the state increased NCA enforcement. The graphs are visualizing raw data. Inventors are assigned into two groups: those who move to another more distant product market (leave) and those who do not (stay). A line is drawn at x = -0.5, between -1, the last untreated year and 0, the first treatment year. Variable definitions are provided in the Appendix.

Panel A: Economic Value of Patents



Panel B: Citation-Weighted Patents



3.5. Channels

In the following section, I analyze potential drivers of productivity effects.

3.5.1. NCA Enforcement leads to Worse Inventor-Firm Matching Quality

The following section differs from the previous in that it draws on a different comparison. I keep the product market dimension constant, e.g. I compare inventors who move across industries to other inventors who also move across industries. The important distinction now is how NCA-constrained industry mobility differs from unconstrained (absent any NCA enforceability changes) industry mobility. I define NCA-constrained as those inventors who move after an increase in NCA enforcement. Unconstrained industry mobility are industry mobility events of inventors in states that did not see increases in NCA enforcement.

For this purpose I analyze new employer-inventor matching characteristics. I analyze whether inventors move to firms that are less similar to them not in a product market dimension, but in a technology dimension. Specifically, I calculate the following measure on technological similarity using patent data between inventor and her new employer:

$$techsimilarity(i,f) = \frac{if^T}{||i||||f||} \tag{3.4}$$

I define two vectors that include the distribution of previous patents across 130 technology subsections. I use the subsection of the Cooperative Patent Classification (CPC) scheme for this purpose, which includes 130 different technology subsections. I use all patents of the inventor up until the year before the industry move and all patents in the previous 5 years of the new employer. The technological similarity is equal to a cosine similarity of the two technology distribution vectors. The measure is bound between zero and one, so it takes a value of zero if no patent section aligns between the employer and the inventor. It is equal to one is the distribution of the two vectors across technology subsections is identical. Technological similarity here is used as a proxy for matching quality between inventor and the firm. If the patent technology subsections of the firm and the patents of the inventors are similar, I assume it is a good match. I then estimate equation 3.2, where y is defined as the technological similarity between inventor i and firm f.

Results are shown in Panel B of table 3.6. The patent technology cosine similarity is reduced by 0.08 for after an increase in NCA enforceability. Given the mean value of 0.4 of technology similarity, this is a reduction of around 20%. This highlights that the matching quality between inventors and employers seems to be much lower in the presence of increased NCA enforcement.

Table 3.6.: Inventor-Employer Matching Quality
This table reports the results of equation 3.2. For Panel A, $EmployerNCA$, a proxy for firm-level use of NCAs, based on information from form 10-Ks and 10-Qs. The variable is equal to one if the firm states that it relies on NCA or whether senior employees sign NCAs. For Panel B, $TechnologyCosineSimilarity$ is the cosine similarity between the distribution of patent technology subsections of the inventor and the new employer. I use all previous patents of the inventor up until one year before the move and the last 5 years of patents for the new employer. Variable definitions are provided in the Appendix. t-statistics are displayed in parenthesis. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively.

Technological Similarity

Dependent variable:	Technology Cosine Similarity
$NCAIncrease$	-0.08***
	(-6.67)
Observations	53,179
R-squared	0.03
Year FE	YES

3.5.2. Non-compete Agreement Enforcement leads to Longer Employment Gaps

NCAs usually have a period of 1-2 years after the end of the employment contract during which employees are not allowed to move to a close competitor. An inventor who wishes to work for another firm faces the following trade-off: Wait until the NCA expires or move to a firm that is further away in the product market. I try to model this trade-off in a regression and hypothesize the following: When NCAs become more enforceable, inventors wait some additional time until they can more easily join a close competitor. This effect should especially be present for within industry moves as they are most likely to be affected by NCAs. I use the following specification:

$$\begin{aligned} \mathit{EmploymentGap}_{i,t} = \beta \times NCAIncrease_i, t + \delta \times Within_i, t + \\ \gamma \times NCAIncrease_i, t \times Within_i, t + \theta_i + \epsilon_{i,t} \end{aligned} \tag{3.5}$$

where $NCAIncrease$ is a dummy variable equal to one if the industry move is after an increase in NCA enforcement. $Within$ is a dummy variable equal to one if the inventor moves to a firm that is in the same SIC 4-digit industry. *EmploymentGap* is the distance in years when an inventor moved between two firms. This is observed in the data by looking at two subsequent patent filing years to different firms by an inventor.

The results are presented in table 3.7. Being constrained by increased NCA enforcement seems to have a general positive impact on employment gaps. This is consistent with the general purpose of NCAs. Moving within the same industry seems to be associated with a reduction of the gap by a little less than one year on average. Most importantly, and consistent with the hypothesis, the interaction of NCA enforcement increase and within industry move is positive and significant. An increase in NCA enforceability especially leads to longer employment gaps for those inventors who move to close industry peers.

Table 3.7.: NCA Enforceability and Employment Gap

This table reports the result of equation 3.5. The dependent variable of interest is employment gap, which is the number of years between two patent filings for each employment move event in the sample. $NCAIncrease$ is a dummy variable equal to one if the inventor moves from a state after an increase in NCA enforcement. $WithinIndustry$ is a dummy variable equal to one if the industry move is within SIC 4-digit industries. Variable definitions are provided in the Appendix. t-statistics are displayed in parenthesis. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively.

Dependent variable:	Employment Gap
$NCAIncrease$	0.89***
	(9.70)
$WithinIndustry$	-0.95***
	(-34.34)
$NCAIncrease \times WithinIndustry$	0.48**
	(2.09)
Observations	263,838
R-squared	0.01
Year FE	YES
State FE	YES

3.6. Industry Mobility and Productivity Across Product and Technology Markets

In the following section, I generalize results on inventor productivity and product as well as technology similarity. I analyze to what extent inventors are more productive depending on how close the new firm is along product and technology dimensions. I introduce a new regression, designed to capture productivity changes after employment changes on the level of individual inventors:

$$Productivity_{i,t} = \beta_i \times Post_i, t + \theta_i + \epsilon_{i,t} \tag{3.6}$$

where $Productivity_{i,t}$ measures the yearly productivity of inventors based on the economic value of patents or citation-weighted patents. The innovation output is firm specific, which means that all patents of the old employer and all patents of the new employer are included in the regression. The dummy variable *Post* is equal to one for years after the inventor has moved to another employer. I estimate the regression for each inventor mobility event, i.e. I run all regressions separately. The coefficient β_i thus captures the extent to which the inventor is more or less productive after moving to another employer. This specification has several desirable properties. First, the inclusion of inventor fixed effects removes the non time-varying quality of the inventor. The specification thus uses patent output of the inventor before and after the move to better tease out productivity differences. Second, the specification is not prone to outliers as each inventor mobility event receives equal weight. Third, the coefficient can be interpreted in an intuitive fashion: How much more/less productive is the inventor after the employment change?

I then use the beta coefficients from these regressions in the following regression:

$$ProductivityCoefficient_{i,f} = \beta_k \times Product_i, f + \delta_k \times Technology_i, f + \theta_i + \epsilon_{i,f} \quad (3.7)$$

where $ProductivityCoefficient_{i,f}$ is defined as the beta coefficient from the inventor productivity regression. It captures to what extent the inventor performs better or worse after moving to another employer. The two variables of interest are product market similarity obtained from Hoberg and Phillips (2016) and the technology similarity calculated from patenting data. I use the last 5 years of patents of the new and the old employer and calculate a cosine similarity based on technology subsections.

The results are shown in table 3.8. Both product market as well as technology similarity are positively correlated with future productivity. This is well aligned with the previous evidence. NCA enforcement can be seen as a constraint primarily on the product market dimension. NCA contract limit employees to freely move to close industry peers. The previous evidence also showed that NCA-constrained employment changes are also associated with lower matching quality. Both of these effects are likely to have negative consequences for future productivity.

Table 3.8.: Inventor Productivity, Technology, and Product Market Similarity

This table reports the result of equation 3.2. The dependent variable of interest is productivity, which captures to what extent the inventor is more productive after changing employers. This variable is measured by economic value of patents and citation-weighted patents following equation 3.6. $TechDistance$ is a variable which captures the patent technology cosine similarity of the inventor and her new employer. $ProductDistance$ captures the extent to which the old employer and the new employer are similar to each other following Hoberg and Phillips (2016). Variable definitions are provided in the Appendix. t-statistics are displayed in parenthesis. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively.

Dependent variable:	Future Productivity (KPSS)	Future Productivity (Citations)
$TechDistance$	0.35*	0.34***
	(1.80)	(2.78)
$ProductDistance$	0.06*	0.32***
	(1.72)	(14.01)
Observations	18,429	18,429
R-squared	0.00	0.01
Year FE	YES	YES

3.7. Firm-Level Productivity Regressions

Does the reallocation of inventors have effects on firms? I answer this question using firm-level productivity growth regressions. I analyze whether the outflow and the inflow of inventors has any effect on future firm growth as measured by productivity growth, output growth, capital growth, and employment growth. I construct these growth regressions following Kogan et al. (2017) and look at whether the in- and outflow of skilled human capital affects growth in the subsequent year. For this I aggregate the yearly out and inflow of inventors on a firm-year level.

The result are shown in table 3.9. Across four different measures of firm productivity, the inflow of inventors is associated with future higher productivity growth. Outflow of inventors is associated with future lower productivity growth.

Table 3.9.: Firm Level Productivity
This table reports firm level productivity regressions following Kogan et al. (2017). The sample is on a firm-year level. The two dependent variables of interest are yearly inventor inflow and outflow which is the natural logarithm of one plus the total number of inflows and outflows respectively. Column (1) is profitability growth, All regressions include the lag of the dependent variable as an additional control. All regressions include Year and SIC 3-digit industry fixed effects. Standard errors are clustered on a Firm as well as on a Year level. t-statistics are displayed in parenthesis. ***, ** and * represents significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)
Dep. Variable:	$ProfitGrowth_{t+1}$	$OutputGrowth_{t+1}$	$CapitalGrowth_{t+1}$	$EmploymentGrowth_{t+1}$
NCA Inventor Inflow	0.04***	0.04***	0.04***	0.04***
	(8.37)	(6.38)	(10.28)	(9.39)
NCA Inventor Outflow	-0.03***	-0.03***	-0.03***	-0.02***
	(-7.18)	(-4.99)	(-8.59)	(-6.99)
Observations	31,765	29,279	33,648	33,419
R-squared	0.58	0.56	0.44	0.46
Year FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES

3.8. Conclusion

Inventors evade their NCAs by moving to new employers in more distant product markets. NCA enforcement increases have a positive causal effect on the probability that an inventor moves across industries. This reallocation has detrimental effects on the productivity of affected inventors. This paper highlights negative consequences of human capital reallocation in response to labor market regulation.

APPENDIX

3.9. Appendix

Variable Definitions This section provides the variable definitions and the sources of the data.

1. *IndustryChange* – Equal to one if an inventor moves from one firm to another with a different industry classification. Obtained from employment histories of inventors from patentsview.org, patents assigned to corporations from Kogan et al., 2017 and Arora, Belenzon, and Sheer, 2021. SIC and NAICS industry codes are obtained from Compustat.

2. *NCA Increase/Decrease* – Equal to one if the state decreased, or increased the enforceability of NCAs. Obtained from Ewens and Marx, 2018 and Kini, Williams, and Yin, 2021.

3. *EmployerNCA* – Equal to one if the firm has mentioned the use of NCAs either in their annual statement or in employment contracts of senior executives. Obtained from 10-K and 10-Q filings downloaded from EDGAR.

4. *Product Market Similarity* – The cosine similarity of the textual product market descriptions between two listed corporations. Obtained from Hoberg and Phillips, 2016 on the Hoberg and Phillips Data Library website:
   https://hobergphillips.tuck.dartmouth.edu/

5. *Patent technology* – The Cooperative Patent Classification (CPC) section was used, which groups patents into 9 different patent sections. Obtained from patentsview.org.

6. *Patent technology subsection* – The Cooperative Patent Classification (CPC) subsection was used, which groups patents into 130 different patent subsections. Obtained from patentsview.org.

7. *Number of patents* – The number of patents of each inventor one year before treatment. Lagged by one year. Obtained from patentsview.org.

8. *Economic Value of Patents, or KPSS* – The economic value of patents, based on stock market reactions to patent grants. Obtained from Kogan et al., 2017, available here:
   https://github.com/KPSS2017

9. *Patent Citations* – The number of received (forward) citations of all patents of an inventor one year before treatment. Citations were truncation adjusted using year and technology fixed effects on a patent basis. See Hall, Jaffe, and Trajtenberg, 2005 and Lerner and Seru, 2021 for details. Obtained from patentsview.org.

10. *Technology Cosine Similarity* – The cosine similarity of the patent technology subsection distributions. The measure includes all previous patents of an inventor and the patents in the last 5 years of the new employer. Obtained from patentsview.org.

11. *Employment Gap* – The difference in years between two subsequent filing years of two patents. The variable is defined when an inventor moves between two firms.

12. *Future Productivity* – Obtained from inventor level regressions. The specification runs separate regressions on each inventor mobility event. The regression includes an inventor fixed-effect as well as a post dummy, which captures the extent to which the inventor is more/less productive after moving to a new employer. Productivity is either measured by the economic value of patents or citation-weighted patents.

13. *Technology Distance* – The cosine similarity of the patent technology subsection distributions (vectors). The measure includes all patents in the last 5 years of the old employer and the new employer. Computed based on data from patentsview.org.

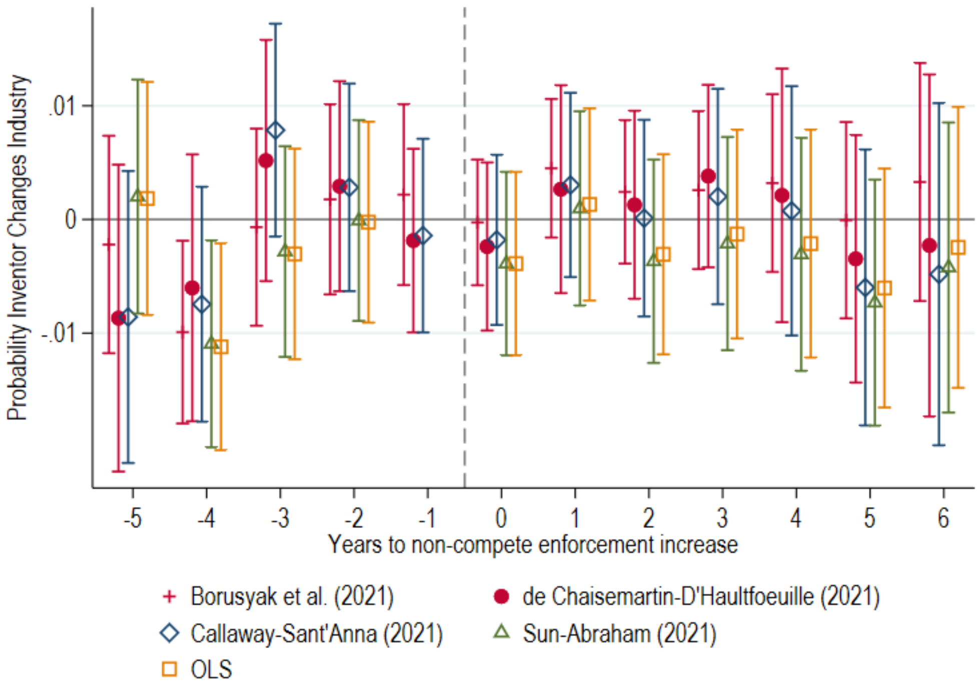
Table A1.: Most Frequent Industry Mobility

This table shows the 5 most common industries ranked according to industry mobility. The table lists the departure industry and the joining industry, a brief description of the industry and the fraction of mobility events compared to the total number of mobility events. Variable definitions are provided in the Appendix.

Rank	Leaving Industry (SIC 3)	Joining Industry (SIC 3)	Fraction
1	Office, Computing, Account. Mach.	Comp. Program., Data Process.	4.4%
2	Office, Computing, Account. Mach.	Electronic Comp. and Accessor.	3.8%
3	Comp. Programming, Data Proce.	Office, Comput., Account. Mach.	2.4%
4	Elec. Components and Accessor.	Comp. Programm., Data Process.	2.3%
5	Communications Equipment	Electronic Comp. and Accessor.	2.1%

Figure A1.: Staggered Difference-in-Differences: NCA Enforcement Decreases

This table reports the result of the staggered difference-in-differences event study of equation 3.1. The sample is on an inventor-year level. The figure plots the coefficient of $NCADecrease$, which is a treatment indicator equal to one for a state that decreases non-compete enforcement. The y-axis shows the effect on the likelihood that an inventor moves across SIC 4-digit industries. The point estimates are normalized to time = -1, the year before treatment. Never-treated inventors are propensity matched based on year, age, number of patents, number of citations and patent technology class. Variable definitions are provided in the Appendix. All regressions include Inventor and Year fixed effects. Standard errors are clustered by Inventor and Year. Confidence intervals are at the top/bottom 5%.



Appendix B: Examples of non-compete agreements

The following are three samples drawn from the sample of innovating firms (those that are assigned patents), of which 54% have references on the use of non-compete agreements. The universe of 10-K and 10-Q filings were obtained from EDGAR.

NUANCE COMMUNICATIONS INC

"In exchange for the severance pay and other consideration under the Severance Agreement to which Executive would not otherwise be entitled, Executive agrees that for a period of one (1) year after the Termination Date, Executive will not, without the express written consent of the Company, in its sole discretion, enter, engage in, participate in, or assist, either as an individual on your own or as a partner, joint venturer, employee, agent, consultant, officer, trustee, director, owner, part-owner, shareholder, or in any other capacity, in the United States of America, directly or indirectly, any other business organization whose activities or products are competitive with the activities or products of the Company then existing or under development. Nothing in this Agreement shall prohibit Executive from working for an employer which is engaged in activities or offers products that are competitive with the activities and products of the Company so long as Executive does not work for or with the department, division, or group in that employer's organization that is engaging in such activities or developing such products. Executive recognizes that these restrictions on competition are reasonable because of the Company's investment in goodwill, its customer lists, and other proprietary information and Executive's knowledge of the Company's business and business plans."

10-Q filing available here:

https://www.sec.gov/Archives/edgar/data/1002517/000100251714000013/nuan12-31x2013ex104.htm

MICROVISION INC

"We also rely on unpatented proprietary technology. To protect our rights in these areas, we require all employees and, where appropriate, contractors, consultants, advisors and collaborators, to enter into confidentiality and non-compete agreements. There can be no assurance, however, that these agreements will provide meaningful protection for our trade secrets, know-how or other proprietary information in the event of any unauthorized use, misappropriation or disclosure of such trade secrets, know-how or other proprietary information."
10-K filing available here:
https://www.sec.gov/Archives/edgar/data/65770/000113626115000080/body10k.htm

LOCKHEED MARTIN CORPORATION

"This Post Employment Conduct Agreement dated [...] (this "PECA"), together with the Release of Claims being entered into contemporaneous with this PECA, is entered into in consideration of the payment ("Severance Payment") to be made to me under the Lockheed Martin Corporation Severance Benefit Plan for Certain Management Employees ("Severance Plan"). By signing below, I agree as follows:
Covenant Not To Compete - Without the express written consent of the [Chief Executive Officer/Senior Vice President, Human Resources] of the Company, during the [two/one]-year period following the date of my termination of employment with the Company ("Termination Date"), I will not, directly or indirectly, be employed by, provide services to, or advise a "Restricted Company" (as defined in Section 6 below), whether as an employee, advisor, director, officer, partner or consultant, or in any other position, function or role that, in any such case, oversees, controls or affects the design, operation, research, manufacture, marketing, sale or distribution of "Competitive Products

or Services" (as defined in Section 6 below) of or by the Restricted Company [...]"

Exhibit of 10-Q filing available here:

https://www.sec.gov/Archives/edgar/data/936468/000119312508156357/dex107.htm

References

Acikalin, Utku, Tolga Caskurlu, Gerard Hoberg, and Gordon M. Phillips (2022). "Intellectual Property Protection Lost: The Impact on Competition and Acquisitions". In: *SSRN Electronic Journal*.

Arora, Ashish, Sharon Belenzon, and Lia Sheer (2021). "Knowledge Spillovers and Corporate Investment in Scientific Research". In: *American Economic Review* 111.3, pp. 871–898.

Babina, Tania (2020). "Destructive Creation at Work: How Financial Distress Spurs Entrepreneurship". In: *The Review of Financial Studies* 33.9, pp. 4061–4101.

Babina, Tania, Paige Parker Ouimet, and Rebecca Zarutskie (2020). "IPOs, Human Capital, and Labor Reallocation". In: *SSRN Electronic Journal*.

Baker, Andrew C., David F. Larcker, and Charles C.Y. Wang (2022). "How much should we trust staggered difference-in-differences estimates?" In: *Journal of Financial Economics* 144.2, pp. 370–395.

Balasubramanian, Natarajan, Jin Woo Chang, Mariko Sakakibara, Jagadeesh Sivadasan, and Evan Starr (2020). "Locked In? The Enforceability of Covenants Not to Compete and the Careers of High-Tech Workers". In: *Journal of Human Resources*, 1218–9931R1.

Bena, Jan, Hernán Ortiz-Molina, and Elena Simintzi (2021). "Shielding firm value: Employment protection and process innovation". In: *Journal of Financial Economics* 146.2, pp. 637–664.

Bishara, Norman (2011). "Fifty Ways to Leave Your Employer: Relative Enforcement of Covenants Not to Compete, Trends, and Implications for Employee Mobility Policy". In: *University of Pennsylvania Journal of Business Law* 13.3, p. 751.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2021). "Revisiting Event Study Designs: Robust and Efficient Estimation". In: *Working Paper*.

Callaway, Brantly and Pedro H.C. Sant'Anna (2021). "Difference-in-Differences with multiple time periods". In: *Journal of Econometrics* 225.2, pp. 200–230.

Chaisemartin, Clément de and Xavier d'Haultfoeuille (2021). "Difference-in-Differences Estimators of Intertemporal Treatment Effects". In: *SSRN Electronic Journal*.

Chen, Yuchen, Xuelin Li, Richard T. Thakor, and Colin Ward (2022). "Appropriated Growth". In: *SSRN Electronic Journal*.

Dass, Nishant, Vikram Nanda, and Steven Chong Xiao (2017). "Truncation bias corrections in patent data: Implications for recent research on innovation". In: *Journal of Corporate Finance* 44, pp. 353–374.

Ewens, Michael and Matt Marx (2018). "Founder Replacement and Startup Performance". In: *The Review of Financial Studies* 31.4, pp. 1532–1565.

Fallick, Bruce, Charles A. Fleischman, and James B. Rebitzer (2006). "Job-Hopping in Silicon Valley: Some Evidence concerning the Microfoundations of a High-Technology Cluster". In: *The Review of Economics and Statistics* 88.3, pp. 472–481.

Garmaise, M. J. (2011). "Ties that Truly Bind: Noncompetition Agreements, Executive Compensation, and Firm Investment". In: *The Journal of Law, Economics, and Organization* 27.2, pp. 376–425.

Goodman-Bacon, Andrew (2021). "Difference-in-differences with variation in treatment timing". In: *Journal of Econometrics* 225.2, pp. 254–277.

Hacamo, Isaac and Kristoph Kleiner (2022). "Forced Entrepreneurs". In: *The Journal of Finance* 77.1, pp. 49–83.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg (2005). "Market Value and Patent Citations". In: *The RAND Journal of Economics* 36.1, pp. 16–38.

He, Zhaozhao (2021). *Motivating Inventors: Non-competes, Innovation Value and Efficiency.*

Hoberg, Gerard and Gordon Phillips (2010). "Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis". In: *The Review of Financial Studies* 23.10, pp. 3773–3811.

— (2016). "Text-Based Network Industries and Endogenous Product Differentiation". In: *Journal of Political Economy* 124.5, pp. 1423–1465.

Hombert, Johan and Adrien Matray (2017). "The Real Effects of Lending Relationships on Innovative Firms and Inventor Mobility". In: *The Review of Financial Studies* 30.7, pp. 2413–2445.

— (2018). *The Long-Term Consequences of the Tech Bubble on Skilled Workers' Earnings.*

Jeffers, Jessica (2017). "The Impact of Restricting Labor Mobility on Corporate Investment and Entrepreneurship". In: *SSRN Electronic Journal*.

Kini, Omesh, Ryan Williams, and Sirui Yin (2021). "CEO Noncompete Agreements, Job Risk, and Compensation". In: *The Review of Financial Studies* 34.10, pp. 4701–4744.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman (2017). "Technological Innovation, Resource Allocation, and Growth". In: *The Quarterly Journal of Economics* 132.2, pp. 665–712.

Lerner, Josh and Amit Seru (2021). "The Use and Misuse of Patent Data: Issues for Finance and Beyond". In: *The Review of Financial Studies*.

Li, Guan-Cheng, Ronald Lai, Alexander D'Amour, David M. Doolin, Ye Sun, Vetle I. Torvik, Amy Z. Yu, and Lee Fleming (2014). "Disambiguation and co-authorship networks of the U.S. patent inventor database (1975–2010)". In: *Research Policy* 43.6, pp. 941–955.

Lipsitz, Michael and Evan Starr (2021). "Low-Wage Workers and the Enforceability of Noncompete Agreements". In: *Management Science* 68.1.

Malsberger, Brian M., Samuel M. Brock, and Arnold H. Pedowitz (2016). *Covenants not to compete: A state-by-state survey*. 10th ed. Bloomberg.

Marx, Matt (2011). "The Firm Strikes Back". In: *American Sociological Review* 76.5, pp. 695–712.

Marx, Matt and Lee Fleming (2012). "Non-compete Agreements: Barriers to Entry ... and Exit?" In: *Innovation Policy and the Economy* 12, pp. 39–64.

Marx, Matt, Jasjit Singh, and Lee Fleming (2015). "Regional disadvantage? Employee non-compete agreements and brain drain". In: *Research Policy* 44.2, pp. 394–404.

Marx, Matt, Deborah Strumsky, and Lee Fleming (2009). "Mobility, Skills, and the Michigan Non-Compete Experiment". In: *Management Science* 55.6, pp. 875–889.

Melero, Eduardo, Neus Palomeras, and David Wehrheim (2017). "The Effect of Patent Protection on Inventor Mobility". In: *SSRN Electronic Journal*.

Samila, Sampsa and Olav Sorenson (2011). "Noncompete Covenants: Incentives to Innovate or Impediments to Growth". In: *Management Science* 57.3, pp. 425–438.

Seru, Amit (2014). "Firm boundaries matter: Evidence from conglomerates and R&D activity". In: *Journal of Financial Economics* 111.2, pp. 381–405.

Shen, Mo (2021). "Skilled Labor Mobility and Firm Value: Evidence from Green Card Allocations". In: *The Review of Financial Studies* 34.10, pp. 4663–4700.

Shi, Liyan (2023). "Optimal Regulation of Noncompete Contracts". In: *Econometrica* 91.2, pp. 425–463.

Song, Jaeyong, Paul Almeida, and Geraldine Wu (2003). "Learning–by–Hiring: When Is Mobility More Likely to Facilitate Interfirm Knowledge Transfer?" In: *Management Science* 49.4, pp. 351–365.

Starr, Evan (2019). "Consider This: Training, Wages, and the Enforceability of Covenants Not to Compete". In: *Industrial and Labor Relations Review* 72 (4), pp. 783–817.

Starr, Evan, J. Prescott, and Norman Bishara (2021). "Noncompete Agreements in the US Labor Force". In: *The Journal of Law and Economics* 64.1, pp. 53–84.

Sun, Liyang and Sarah Abraham (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". In: *Journal of Econometrics* 225.2, pp. 175–199.

Chapter IV

Angels and Demons: The Negative Effect of Employees' Angel Investments on Corporate Innovation

with Santanu Kundu

Angels and Demons: The Negative Effect of Employees' Angel Investments on Corporate Innovation

Santanu Kundu and Clemens Mueller[1]

Abstract

We link data on angel investors in the US to their employment histories. Employers' innovation output decreases when their employees personally invest in early-stage firms. We find evidence for two novel channels: agency conflicts and loss of highly skilled human capital. On a positive note, early-stage firms benefit from financing by angel investors employed at public firms. Angel investors divert time and effort from their employer to their personal investments. We highlight a trade-off between the costs of angel investors for their employers and the benefit for their start-ups.

[1]Comments appreciated. Please contact Santanu Kundu at santanu.kundu@uni-mannheim.de or Clemens Mueller at clemens.mueller@.uni-mannheim.de. We would like to thank Paul Beaumont (discussant), Jan Bena, Johan Cassel (discussant), Cláudia Custódio, Miguel Ferreira, Bernhard Ganglmair, Camille Hebert (discussant), David Heller (discussant), Sabrina Howell (discussant), Gunchang Kim (discussant), Ernst Maug, Daniel Rettl (discussant), Stefan Ruenzi, Johannes Voget, and Chris Yung (discussant) for helpful comments and advice. We also thank seminar participants at the European Finance Association, Paris December Finance, Eastern Finance, Northern Finance, French Finance Association, student-led Workshop on Entrepreneurial Finance and Innovation (WEFI) among others for their helpful comments. We thank Crunchbase for providing us data. The authors acknowledge the support of the high performance cluster computing facilities and the German Research Foundation (DFG) through grant INST 35/1134-1 FUGG. All errors are our own.

4.1. Introduction

Angel investors, individuals who personally invest in start-ups, are an important driver of innovation and success of their portfolio start-ups (Kerr, Lerner, and Schoar, 2014; Lerner et al., 2018). In this paper, we investigate the broader role of such angel investors. We define so-called *angel employees* as angel investors who are simultaneously employed at a publicly listed corporation. In this paper, we ask the following question: Do angel employees help or hinder the innovation output of their employer?

Ex-ante, it is unclear how angel employees impact their employer's innovation. On the one hand, angel employees could help their employer innovate. Angel employees might use their personal investments to acquire knowledge about start-ups' existing and future innovative activities[1] and guide innovative activities at their employer. On the other hand, angel employees could have a detrimental impact on corporate innovation. This is rooted in a standard principal-agent or multitask framework (Jensen and Meckling, 1976; Holmstrom and Milgrom, 1991). Angel employees must trade off exerting effort in the innovative activities of their employer or their personal investments.

To test these conflicting hypotheses, we exploit novel data, which link angel investments of individuals to their employment history. We collect personal equity investments in US early-stage firms between 2001-2018 from Crunchbase which totals more than $21 billion of early stage capital. We then obtain employment histories of angel investors in our sample from LinkedIn and manual searches. We match 1,845 angel employees who have 5,379 investments and are employed at 792 unique publicly listed firms in the US. The company with the most angel employees in our sample is Alphabet, with a total of 196 employees who personally invested in 433 start-ups between 2001 and 2018.

[1]Indeed, previous research has shown that many angel investors obtain a board seat or act in an advisory role (Wallmeroth, Wirtz, and Groh, 2018).

We start with descriptive statistics. The average size of a funding round with angel employee participation is large at 5.7 million USD at the mean and 1.7 at the median. Angel employees are closely related to innovation and many employees have functions such as "tech", "engineer", "innovation" or "product". They are also senior and are thus important decision makers at their employers. 43% are directors and 14% are executives with the remainder managers, or vice presidents. The early-stage firm is usually closely located near the headquarter of the employer. Also, the industry of the start-up is closely related to the industry of the employer.

In our baseline specification, we analyze correlations between the presence of angel employees and future innovation output. Our preferred dependent variable of interest is the economic value of patents, based on stock market reactions to patent grants (Kogan et al., 2017). Our second measure is (forward) citations received. The independent variable of interest is either defined as a dummy when angel employees are present or not or the natural logarithm plus the total number of angel employees. In our baseline regression, the presence of angel employees is associated with 4% lower economic value of patents.

This baseline result is purely correlational and likely suffers from omitted variable bias. For example, a life-cycle based hypothesis implies that employees invest in innovative start-ups if innovation output of their employer is declining. To address this, we exploit within firm-year variation. We use the fact that angel employees and innovation output are geographically dispersed across states within a firm. We aggregate innovation across research departments of a firm in a firm–state–year panel. This allows us to control for observed and unobserved firm–year, firm–state and state–year fixed effects. These fixed effects account for firm-year explanations such as the life cycle stage. We analyze whether innovation declines in a research department where angel employees are co-located. Indeed, innovation output declines by $47 million after four years when angel employees are co-located.

Next, we analyze dynamic effects in a staggered differences-in-differences event time framework. Within a firm, we compare a firm-state with angel employees (treated) to a firm-state without (never treated). Reassuringly, we do not see any evidence of pre-trends. The negative effect is confined to the time *after* employees personally invest in start-ups. This alleviates the concern that reverse causality explains our results.

We complement our results with an instrumental variable regression. To some extent, venture capital competes with angel employees. We use a sudden inflow of venture capital due to the staggered implementation of the PIR, the so-called "prudent man rule". We provide evidence that this competition shock crowds out angel employees. This is because angel employees invest in large funding rounds and both types of investors invest locally. The second stage confirms our baseline and difference-in-differences results. We argue that the exclusion restriction is plausible as previous research has showed that if anything, venture capital *positively* affects innovation output (Kortum and Lerner, 2000). This would bias us against finding a negative effect.

We provide evidence on two novel channels which explain the negative baseline effects. First, agency conflicts and second, employee exit and thus loss of highly skilled human capital. Angel employees trade-off whether to exert effort at their employer or their personal investments. This trade-off is exacerbated by the long-term nature of angel investments and the potential to earn extra-ordinary returns. There are two ways how a agency conflict can manifest: selection and/or treatment. Careful selection of investments might involve time intensive deal scouting. Angel employees might also actively monitor their portfolio start-ups. In doing so, angel employees might help the start-ups in their day-to-day operations and provide advice and expertise. They could also be involved in intensive board meetings for follow-on financing rounds, an acquisition, or going public. We proxy for active monitoring and selection with ex-post startups success, and expect lower innovation for ex-post relatively more successful start-ups. We find this to be the case. The negative effect is more pronounced if the linked start-ups were ex-post relatively successful.

To further analyze agency conflicts, we explore the effect of the passage of the Small Business and Jobs Act (SBJA), which made angel investments tax exempt after 2010. We use the passage of the law as a quasi-exogenous shock that affects the *incentives* of angel employees to select and monitor their start-up investments. Thus, angel employees would be incentivized to spend more time and effort in their invested start-ups rather than at their employer. This allows us to tease out the effect of agency conflicts inherent in angel investments. The negative effect is much more pronounced for angel employees who invested in start-ups eligible for tax exemption after 2010.

As a second channel, we look at exit rates and the loss of human capital. Active angel employees are, during the time of their investment, 2% more likely to exit their firm. This is costly for employers due to high replacement costs especially for highly skilled employees.

We turn the focus to the point of view of the start-ups. We ask the question whether angel employees are beneficial for their portfolio start-ups. In difference-in-differences regressions, we compare start-ups financed by angel employees to start-ups financed by other angel investors. Angel employee participation in a funding round sharply increases follow-on venture capital participation, a higher M&A probability as well as higher innovation output of the early-stage firms. Angel employees are thus either skilled in selecting and/or treating their investments, which our analysis does not allow us to disentangle. This evidence is inconsistent with the hypothesis that angel employees are low quality employees. It is rather consistent with the hypothesis that innovation output of employers particularly suffers when its skilled employees divert their time and effort.

Lastly, we perform some plausibility and robustness exercises. We first show that our baseline effects are primarily due to angel employees who are *directly* related to innovation. We generalize our findings to non patent based innovation outcomes such as new product announcements, trademarks, and scientific publications. Our results also hold when generalizing to only private employers as well as excluding recently IPO firms.

Our analysis is nuanced as we show negative effects for employers and positive effects for early-stage firms. We perform a back-of-the-envelope welfare analysis and compare the innovation output lost on the employer side with the innovation gained on the start-up side. We calculate an aggregate yearly loss of 1,506 patents for public employers and a gain of 1,993 patents for start-ups. This crude calculation indicates that the welfare effects are likely positive.

We primarily contribute to the literature which explores the role of angel investors in the economy. To the best of our knowledge, we are the first to combine employment and investment information of angel investors. 20% of angel investors are angel employees and we show that they can be important corporate decision makers. Previous research has shown that angel investors are beneficial for early-stage firms (e.g., Sudek, Mitteness, and Baucus, 2008; Kerr, Lerner, and Schoar, 2014; Lerner et al., 2018) We add to the literature an important source of heterogeneity of angel investors. It seems to be highly skilled employees with industry expertise who are beneficial for early-stage firms.

4.2. Data

4.2.1. Angel Employees

Our data on angel investments comes from Crunchbase. Crunchbase contains information on more than 450,000 funding rounds across 173 countries. The company gathers information initially through crowd sourcing and validates accuracy through its dedicated data science team. The vast majority of data is collected through its partnerships with more than 3,500 investment firms, an active community of users, and staff who continuously update data.[2] We focus on angel investments and thus only retain participation in funding rounds of individuals. In total, there are 25,999 unique angel investors in the sample, of which 14,772 have investments in early-stage firms headquartered in the US.

4.2.2. Employment Histories

Key to our data collection is matching angel investors to their employer. We obtain biographies either through LinkedIn or through manual searches. We obtain historical employment data from public LinkedIn profiles. Crunchbase provides individual profile links for the majority of angel investors in our sample. We verify these and collect missing links through manual searches. We match employer names from LinkedIn to publicly listed firms using a fuzzy name matching algorithm. For this purpose, we obtain historical names from CRSP. We standardize names and remove legal pre- and suffixes. Then we compute a Levenshtein distance which measures the distance between

[2]Crunchbase has been compared to traditional datasets and is the most extensive database for early-stage start-up funding round information (Retterath and Braun, 2020; Dalle, den Besten, and Menon, 2017; Ling, 2015) A number of recent papers rely on Crunchbase for data on early-stage private financing rounds (Kaplan and Lerner, 2017; Dimmock, Huang, and Weisbenner, 2019; Edwards and Todtenhaupt, 2020).

strings. We manually verify close strings. Second, we complement this data with manual data collection. Well known individuals such as Mark Zuckerberg might be less likely to have a LinkedIn profile page. Omitting such angel employees might introduce sample selection. We mitigate this problem by manually obtaining the employment history of all angel employees with at least three investments in our data. Of the 10,723 unique angel investors, we obtain full employment history (LinkedIn and manual searches) of 9,006 angel investors, a coverage of 84%.

We note that angel investments do not need to be disclosed, so we are likely to capture a lower bound of the angel investor universe. A concern is whether the public disclosure of angel investments suffers from selection bias. Start-ups might have an incentive to strategically disclose prominent investors as they can serve as a credible signal to the market. Strategic disclosure by angel employees on the other hand might be more problematic. If employees of more innovative corporations are less likely to disclose their angel investments, we would overestimate the negative effect of angel employees. However, this is unlikely for two reasons. First, we look at the total number of angel employees of the whole corporation, so strategic disclosure needs to be correlated on a firm level. Second, in order for this to be a problem, there needs to be a correlation between angel investment disclosure and *future* declining corporate innovation. We do not think that this is likely, but this remains a potential concern for our analysis.

4.2.3. Sample Construction and Independent Variables

We display filter steps in table 4.1. In order to restrict ourselves to angel investments, we only keep data on equity or equity-like investments that are tied to individuals. We remove investments tied to venture capital partners and individuals employed in a corporate venture capital unit. We restrict the sample to US early-stage firms in the years 2001 to 2018 due to low data coverage before 2001. After matching angel investors to corporations, our final data set of angel employees is comprised of 1,845 unique angel employees, which work for 792 unique corporate employers. Since angel employees have multiple investments and the size of the funding rounds are large, the total of all unique funding rounds in the final sample covers more than $21 billion early stage financing. This includes many well-known startups and angel employees. More detail on data collection and background information is available in the Appendix.

Table 4.1.: Sample Selection Steps
This table shows the filters applied and the corresponding number of angel investors.

No.	Sample Selection Step	No. of Angels
(1)	Investments tied to individuals in Crunchbase as of January 2022	25,299
(2)	Only investments into US start-ups	14,772
(3)	Only equity-like investments	14,624
(4)	Only investments between 2001-2018	10,723
(5)	Only angel investors with employment histories	9,006
(6)	Angel investors are at some point employed at a public firm	3,812
(7)	Angel investors employed around time of investment	1,845

We make use of the standardized nature of LinkedIn profile information and collect information on location and the individual's role within the organization for all angel employees in our sample. Following evidence from business angel surveys, we assume an average angel investment holding period of five years.[3] Our variable of interest is either

[3]The American Angel (2017) among others say that the target mean and median duration of a typical angel investment is five years. The results are quantitatively and qualitatively similar when assuming that angel employees keep their investments for shorter time periods or forever.

defined as a dummy variable equal to one if there is at least one angel employee on a firm level. Alternatively we take the natural logarithm plus one of the total number of angel employees. The variation in this variable comes from two sources: 1) existing employees of a firm start to invest in early-stage firms and effectively become angel employees and 2) existing angel employees move across firms. We retain both sources of variation, however the vast majority comes from the first source.

4.2.4. Innovation Output

Our main measure of firm innovation is the economic value of patents obtained from Kogan et al. (2017), henceforth referred to as KPSS. The provided data links patent numbers to publicly listed firms and includes the economic value of patents from 2001 to 2018, which in total includes 1,283,974 patents granted to 5,321 firms. The economic value of patents is based on stock market reactions to patent grants. We are primarily interested in understanding whether angel employees provide value to shareholders. As noted in Kline et al. (2019), the KPSS measure is particularly suitable for this purpose as opposed to other measures of innovation. We aggregate our preferred innovation variable on a yearly level and scale by total assets following Kogan et al. (2017). Our second measure of firm innovation are citation-weighted patents. Since younger patents naturally have less time to be cited, we perform a truncation-adjustment and control for year and technology class fixed effects (Lerner and Seru 2021, and Dass, Nanda, and Xiao 2017). In order to limit a possible truncation bias, we follow the suggestions by Lerner and Seru (2021). We obtain citations received until December 31, 2021 directly from the United States Patent Office (USPTO) accessed through Patentsview. Our main regressions only use patents granted until the year 2018, such that each patent has at least three years to be cited. In order to identify innovation creation more timely, we use the application year of the patent. We use three alternative measures of innovation that are non-patent based: the number of trademarks and new product launches. We also look at science publications, obtained from Arora, Belenzon, and Sheer (2021).

Some of our regressions make extensive use of the precise location where innovation is generated. Specifically, we aggregate innovation on a more granular firm-state-year level. In the following, we use IBM as an example to highlight the potential usefulness of such a panel. Patent data includes detailed information on which city where each inventor works. Over our timeframe, we observe patent filings of IBM inventors in 45 different states (plus Washington D.C. and Virgin Islands). Based on this data, IBM only lacks research departments in Mississippi, Montana, North Dakota, South Dakota, and West Virginia, but otherwise they have a presence in all other states. This indicates that innovation creation at a large firm can be geographically quite dispersed. IBM is officially incorporated and headquartered in Armonk, New York. In the year 2016, we analyze the distribution of innovation generation across US states for IBM. New York, as the headquarter location, is the state with the largest share of innovation generation. As a percentage of citations generated, New York however, only makes up around 10% of the total citations generated that year for IBM. California follows with 7.2%, and Texas with 6.8%. IBM, as a large and research-intensive firm, is arguably an extreme example, so we repeat this exercise systematically. The average firm in our sample generates 64% at the mean and 76% at the median of innovation in the headquarter state. The average non headquarter state generates 6% of innovation at the mean and 0% at the median. The previous literature often attributes all innovation to the headquarter state. In later analyses, we will make use of where the innovation output is created.

We also match patents to startups in our sample. We perform a fuzzy name matching algorithm and exclude punctuation, capitalization and legal pre- and suffixes. We only keep matched firms in the two databases if they are located in the same state. In total, we match 418,973 patents to 12,236 start-ups. From this data, we compute truncation-adjusted citation-weighted patents on a startup-year level.

4.2.5. Other Control Variables

We obtain additional firm level control variables from CRSP and Compustat. We follow Fang, Tian, and Tice (2014) and control for the following 15 variables: log of market capitalization, research and development expenses, Tobin's Q, profitability, asset tangibility, the log of firm age, the Herfindahl index defined over yearly sales in the 4-digit SIC code, Herfindahl index squared to capture non-linear effects, stock liquidity proxied by the daily mean bid-ask spread, capital expenditures, leverage, financial constraints, past patent stock, and the log of the number of employees. We also control for the presence of a corporate venture capital program following Ma (2020)[4]. All variables and sources are listed and described in the Appendix. To mitigate the impact of outliers, we winsorize all continuous variables at the 1% level.

4.2.6. Descriptive Statistics

Table 4.2, Panel C presents the descriptive statistics of the variables used in our study. The economic value of patents refers to the innovation output of a firm as measured by stock market reactions to patent grants applied in the next year. Our sample statistics are quantitatively similar to previous studies (Fang, Tian, and Tice, 2014). The patent

[4]More corporations have active angels than an active corporate venture capital program. There is hardly any overlap between the two ways of investing in startups within the corporation. It is very rare that an employee invests in a startup and the corporate venture capital program of the employer invests in the same startup.

distribution is highly skewed. The mean economic value of patents in our sample is 3% of the book value of a firm. Our main variable of interest, the number of angel employees, is also highly skewed. The vast majority of firms do not have angel employees. We directly address econometric concerns due to the skewed distribution in the upcoming sections, e.g. in later analyses we confine the analysis to within-firm with angel employees.

Table 4.2.: Industries with Angel Employees and Summary Statistics
Panel A shows the top five SIC industries that employ the most angel employees. For each industry we list two example firms. Panel B and C shows summary statistics on the startup and the firm level, respectively. Variable definitions are provided in Section A of the Appendix.

Panel A: Industries and Employers with Most Angel Employees

Rank	SIC	Description	Example firms
1	7370	Services-Computer Programming, Etc.	Alphabet, Facebook
2	7372	Services-Prepackaged Software	Microsoft, Adobe
3	5961	Retail-Catalog & Mail-Order Houses	Amazon, Wayfair
4	2836	Biological Products	Moderna, Unity Biotech
5	7374	Services-Computer Processing, Data Preparation	Square, Paypal Holdings

Panel B: Summary Statistics on the Angel Employee - Startup level

Variable	N	Mean	SD	Min	25%	50%	75%	Max
Funding Round Size (USD M)	5,379	5,70	29.01	0.01	0.67	1.70	4.00	120.00
Startup-Corporation Distance	3,491	1,185	1,567	0	31	59	2,420	4,389
Same State Dummy	3,491	0.59	0.49	0.00	0.00	1.00	1.00	1.00
Startup-Firm Product Sim.	3,551	0.04	0.04	0.00	0.01	0.02	0.04	0.50
Board of directors dummy	5,379	0.43	0.50	0.00	0.00	0.00	1.00	1.00
Executive dummy	5,379	0.14	0.35	0.00	0.00	0.00	0.00	1.00
Innovation-related Angel	5,379	0.61	0.49	0.00	0.00	1.00	1.00	1.00
Non-Innovation-related Angel	5,379	0.06	0.49	0.00	0.00	0.00	1.00	1.00

Panel C: Summary Statistics on the Employer level

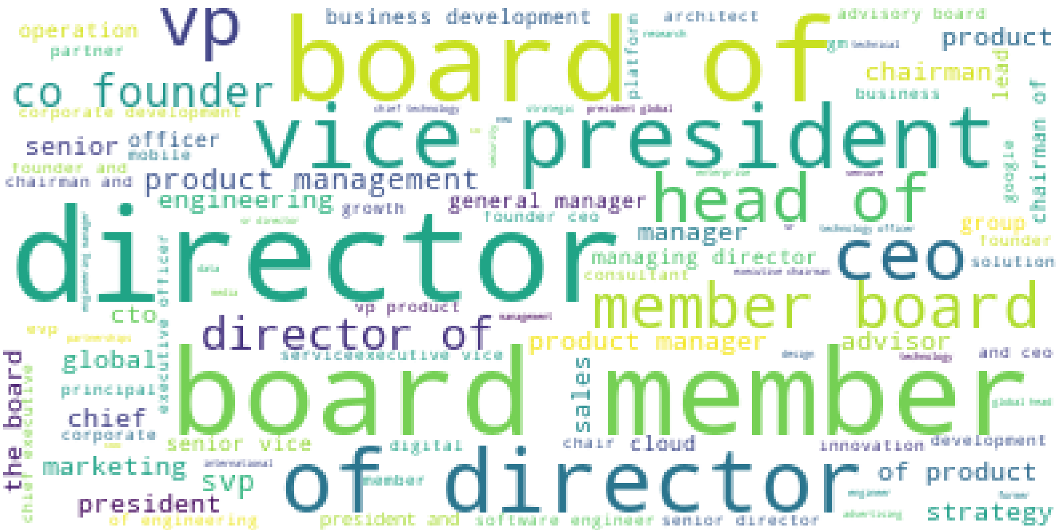
Variable	N	Mean	SD	Min	25%	50%	75%	Max
Firm-Year Panel								
$EconomicValueofPatents_{t+1}$	70,408	0.03	0.12	0.00	0.00	0.00	0.00	1.2
$Citation-WeightedPatents_{t+1}$	70,408	0.42	1.15	0.00	0.00	0.00	0.00	9.21
Angel Employee Dummy	70,408	0.03	0.18	0.00	0.00	0.00	0.00	1.00
Firm-State-Year Panel								
$EconomicValueofPatents_{t+1}$	330,956	27.63	142.17	0.00	0.00	0.00	0.18	1,522
$Citation-WeightedPatents_{t+1}$	330,956	0.26	0.76	0.00	0.00	0.00	0.00	9.11
Angel Employee Dummy	330,956	0.01	0.08	0.00	0.00	0.00	0.00	1.00

Our setting involves personal angel investments of individuals who are simultaneously employed at a public corporation. We are not aware of existing research that has collected this data, therefore, we first provide some descriptive statistics on several dimensions of our sample, to make the presented evidence more accessible to interested readers, encourage more research, and also to motivate the choice of some of our specifications.

Angel employees are not a rare occurrence. 42% of angel investors with observable employment history are or were employed at a listed corporation. This is consistent with survey evidence such as The American Angel (2017). In this survey, 55% of angel investors are or were previously executives at for-profit companies and 46% are or were members of the board. When we look at a narrow time around the angel investment (up to 5 years after), 20% of angel investors are active angel employees. In our data, 792 publicly listed firms have angel employees. Firms in information technology and related industries have the most (see Panel A in table 4.2). The company with most angel employees in our sample is Alphabet, to which we can link a total of 196 employees who personally invested across 433 start-ups between 2001 and 2018.

To provide a sense of what role angel employees play at their employer, we visualize their function in figure 4.1. 43% of angel employees are members of the board of directors. 14% are executives (of which 35% are CEOs) and we classify the remainder as others. When we look more closely into the third category, almost all belong to upper management. Most angel employees report their title as: presidents, vice president, group lead, and other senior managerial roles. Another observation is that many angel employees are in innovation related roles. We often see titles such as product manager, developer, researcher, etc. on the self-provided LinkedIn profile. In a later analysis we separate angel employees in whether they are in a innovation related and non-innovation position. We use the profile information obtained from LinkedIn and based on the title of each employee, we tag employees with the words "product", "innov", "research", "engineer", "tech", among other keywords as innovation-related and angel employees with titles such as "finance", "legal", "accounting", "audit", "operation", "banking", among others as non-innovation related angels. We classify 61% as innovation related, 6% as non-innovation related. The remainder group is classified as neither and is largely composed of directors and executives. This highlights that angel employees are either senior level employees or are also closely related to innovation.

Figure 4.1.: Role of Angel Employees
This figure visualizes the roles of angel employees in our data. Roles are defined as the position the angel employees list in their LinkedIn profile. The size of the font is weighted by counts, i.e. more frequently mentioned roles are displayed more prominently.



We next look at the size of funding rounds in our sample. As shown in the first row of Panel B of table 4.2, the median (mean) funding round in our final data set is \$1.7M (\$5.7M). This amount is significantly larger than a typical angel financing round as previously reported by surveys (The American Angel, 2017). Unfortunately, we do not observe the individual amount each angel employee invests, but only the total amount of each financing round. Many rounds include both angel investors and venture capital investors. The large amounts however make it unlikely that these investments are only token investments, but rather that there are large potential returns, and thus incentives to help the portfolio early-stage firm succeed.

In later analyses we exploit the fact that most investments are local. For this purpose, we compute the distance by using the headquarter location of the employer and the location of the start-ups, when available. We infer the distance from the city level and use the latitude and longitude of the city midpoint. As shown in the second row of Panel B of table 4.2, the median (mean) distance of an angel employee investment is 59 (1,185) miles. We compute a dummy equal to one if the angel and startup are located in the same state. The dummy is equal to one for 59% of angel employees' investments.

We also look at the industry similarity between employers and start-ups in which angel employees invest. Early-stage investments are characterized by high information uncertainty. Angel employees have industry expertise and can leverage this to select high-quality early-stage firms. Crunchbase does not provide standard industry classifications such as SIC or NAICS codes. We therefore compute a textual product market closeness variable between start-ups and corporations similar in spirit to Hoberg and Phillips (2010). We obtain a textual description of corporations from section 1 or 1A from the 10K of corporations from EDGAR. Crunchbase provides a textual description of most start-ups in our sample. We weigh unique words in both vectors by their occurrence and compute a cosine similarity. Linked start-ups have an average cosine similarity of 3.5%. To interpret this number, we compare it to the similarity of randomly matched pairs. We draw 3,000 random startup-corporation matches and receive a cosine similarity of 1.1%. Actual matches are therefore more than three times closer than a randomly matched pair.

To sum up, a large portion of angel investors are angel employees. Angel employees are often board members, executives or other senior employees and are closely related to innovation. They tend to make investments that are large in value, local and the business of the early-stage firms are closely related to the industry of their employers.

4.3. Empirical Results

4.3.1. Baseline Panel Regression

To investigate the effect of angel employees on innovation output, we estimate the following panel regression:

$$Innovation_{i,t+1} = \beta \times AngelEmployeeDummy_{i,t} + \gamma \times \mathbf{X}_{i,t} + \theta_i + \phi_t + \epsilon_{i,t} \tag{4.1}$$

where i represents firm i in year t. We measure *Innovation* as the innovation output of patents filed in the next year. We use two main measures of innovation output: The total yearly economic value of patents scaled by total assets following Kogan et al. (2017), and truncation-adjusted citation-weighted patents. In our baseline specification, we define our independent variable of interest, *Angel Employee Dummy*, as equal to one if a firm employs at least one angel employee in a year.[5] Alternatively, we define the independent variable as the natural logarithm plus one of the total number of angel employees. The vector $\boldsymbol{X}$ represents 15 standard control variables, as previously described. The variables θ and ϕ are firm and year fixed-effects, respectively. Year fixed-effects account for year-specific shocks to innovation. Firm fixed-effects control for non time-varying unobserved factors on the firm level. We cluster standard errors on a firm level to correct for auto-

[5] The results are robust to using the raw number, as well as scaled versions (e.g., by the board size) of the total number of angel employees at a firm.

correlation of innovation at a given firm over time following Fang, Tian, and Tice (2014).

Table 4.3.: Baseline Regression: Angel Employees and Innovation Output
This table reports the fixed effect regression of equation 4.1. The unit of observation is on a firm-year level. The dependent variable in Panel A is the economic value of patents scaled by assets (Kogan et al., 2017) in the next year ($Innovation_{t+1}$). In Panel B, we alternatively use truncation-adjusted citation-weighted patents. The independent variable, *Angel Employee Dummy*, is equal to one if there is at least one angel employee at the firm. Alternatively it is the natural logarithm of the number of angel employees on a firm level. Columns (3) and (4) use a inverse hyperbolic sine transformation of the dependent variable. Column (5) only looks at firms which patent. The regression includes 15 standard control variables. Variable definitions are provided in section A of the Appendix. The regression includes Firm and Year fixed effects. Standard errors are clustered by firm. ***, ** and * represents significance at the 1%, 5% and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)	(3)	(4)	(5)
Panel A – Dependent variable: Economic Value of Patents					
	$KPSS_{t+1}$	$KPSS_{t+1}$	$KPSS_{t+1}$	$KPSS_{t+1}$	$KPSS_{t+1}$
Angel Employee Dummy	-0.03***		-0.03***		-0.03***
	(-4.43)		(-4.45)		(-3.03)
ln(1+Angel Employees)		-0.05***		-0.05***	
		(-5.44)		(-5.52)	
Panel B – Dependent variable: Citation-Weighted Patents					
	CIT_{t+1}	CIT_{t+1}	CIT_{t+1}	CIT_{t+1}	CIT_{t+1}
Angel Employee Dummy	-0.19***		-0.09***		-0.15**
	(-4.23)		(-3.93)		(-2.57)
ln(1+Angel Employees)		-0.27***		-0.10***	
		(-5.22)		(-4.42)	
Observations	70,408	70,408	70,408	70,408	31,915
Controls	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

The results are presented in table 4.3. In Panel A, column (1) of the first row, the presence of at least one angel employee is associated with a decrease in the economic value of patents by 3% of the book value. In column (2) we alternatively use the natural logarithm of one plus the total number of angel employees. Due to the skewed nature of our variables, in columns (3) and (4) we repeat the previous regressions, however we replace the dependent variable with an inverse hyperbolic sine transformation.[6] The results are unchanged.

In column (5), we restrict the sample to only those firms that patented during our sample period. The results remain qualitatively similar. Panel B of the table repeats the analysis but for a different dependent variable: truncation-adjusted citation-weighted patents.

4.3.2. Within Firm-year: Angel Employees and Innovation Output Across States

One reason that prevents us from drawing causal conclusions from our baseline results is a firm life-cycle based explanation. If a firm matures and faces lower future growth opportunities, employees might personally invest in start-ups to diversify themselves. This would lead to a spurious correlation of angel employees with lower future innovation. In order to alleviate such concerns, we introduce firm-year fixed effects to the regression. This controls for observed and unobserved firm-year level omitted variables such as firm life-cycle stage, yearly capital expenses, or the annual research budget. Adding this stringent set of fixed effects is only possible if we use *within* firm-year variation of angel employees and innovation output. For this, we use the location of angel employees as provided by LinkedIn data. For innovation output, we use USPTO data which provides the precise location of each inventor listed on a patent. Similar to Foley and Kerr (2013),

[6]The inverse hyperbolic sine transformation is defined as $log(y_i + (y_i^2 + 1)^{1/2})$ and has the additional desirable property to allow zeros to be included without adding a constant term.

we assign the economic value (or forward citations received) of each patent proportionally to the U.S. states the inventors are located in. We thus aggregate innovation output across each firm per state per year. We ask the following question: Does innovation for a firm decline precisely in the state where an angel employee is co-located? In order to answer this question, we run the following regression:

$$Innovation_{i,s,t+k} = \beta \times AngelEmployeeDummy_{i,s,t} + \theta_{i,t} + \phi_{i,s} + \psi_{s,t} + \epsilon_{i,s,t} \tag{4.2}$$

where the unit of observation in this regression is the innovation output aggregated on a firm-state-year level. $\theta_{i,t}$ captures firm-year fixed effects as discussed previously. $\phi_{i,s}$ captures firm-state fixed effects, and $\psi_{s,t}$ captures state-year fixed effects. We include firm-state dummies to control for the fact that a given firm might have an R&D department in Washington because of local technology expertise in this state. State-year dummies capture state-wide economic shocks. $AngelEmployeeDummy_{i,s,t}$ now is a dummy variable equal to one if there is at least one angel employee in a given firm-year-state observation and zero otherwise. The dependent variable is either the total economic value of patents or total truncation adjusted citation weighted patents, again on a firm-state-year level. The coefficient of interest in this regression is β. We cluster standard errors on a firm-state level. Given this strict set of fixed effects, the relevant variation to identify the effect on innovation output comes from the time varying presence of angel employees for a firm in a given state.

Table 4.4.: Within Firm-Year: Angel Employees and Innovation Output Across States

This table reports the result of the fixed effect regression of equation 4.2. The unit of observation is on a firm-state-year level. The dependent variable in Panel A is the economic value of patents (Kogan et al., 2017) over the subsequent k years ($Innovation_{t+k}$), where $k = [1, 5]$, respectively. In Panel B, we alternatively use truncation-adjusted citation-weighted patents. The independent variable, *Angel Employee Dummy*, is equal to one if the firm employs at least one angel employee in a given year in a given state in the US. Variable definitions are provided in section A of the Appendix. The regression includes Firm-Year, Firm-State, and State-Year fixed effects. Standard errors are clustered by Firm-State. ***, ** and * represents significance at the 1%, 5% and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)	(3)	(4)	(5)
Panel A – Dependent variable: Economic Value of Patents ($KPSS$)					
	$KPSS_{t+1}$	$KPSS_{t+2}$	$KPSS_{t+3}$	$KPSS_{t+4}$	$KPSS_{t+5}$
Angel Employee Dummy	9.63	-12.54	-31.88***	-46.52***	-61.87***
	(1.03)	(-1.32)	(-3.23)	(-4.71)	(-5.82)
Panel B – Dependent variable: Citation-Weighted Patents (CIT)					
	CIT_{t+1}	CIT_{t+2}	CIT_{t+3}	CIT_{t+4}	CIT_{t+5}
Angel Employee Dummy	-0.12	-1.20***	-2.26***	-3.07***	-3.45***
	(-0.27)	(-2.67)	(-4.60)	(-5.99)	(-6.48)
Observations	330,956	311,488	292,020	272,552	253,084
Firm-Year FE	YES	YES	YES	YES	YES
Firm-State FE	YES	YES	YES	YES	YES
State-Year FE	YES	YES	YES	YES	YES

Results are shown in table 4.4. Since the total assets of a firm is absorbed by firm-year fixed effects, the economic value of patents can be directly interpreted in million USD. In Panel A, angel employee presence does not immediately lead to a decline in innovation output. However, the presence of at least one angel employee is associated with -32 million USD lower economic value of patents after three years. The effect persists over time.

To put the economic effect into context, we relate the \$62 million figure after 5 years to the sample mean. Angel employees are most prevalent among large firms that are highly innovative. The average economic value of patents in a firm-state observation conditional on angel employee and patenting activity is \$400 million per year. Relative to this number, the economic effect is thus a decrease of around 15% when angel employees are co-located.

4.3.3. Within Firm-year: Event Study and Dynamic Effects

An important concern for a causal interpretation of our results is reverse causality. Reverse causality makes intuitive sense if innovation output is already on a downwards path and this causes employees to invest in early-stage firms. To address this concern, we estimate the following difference-in-differences regression in an event time framework:

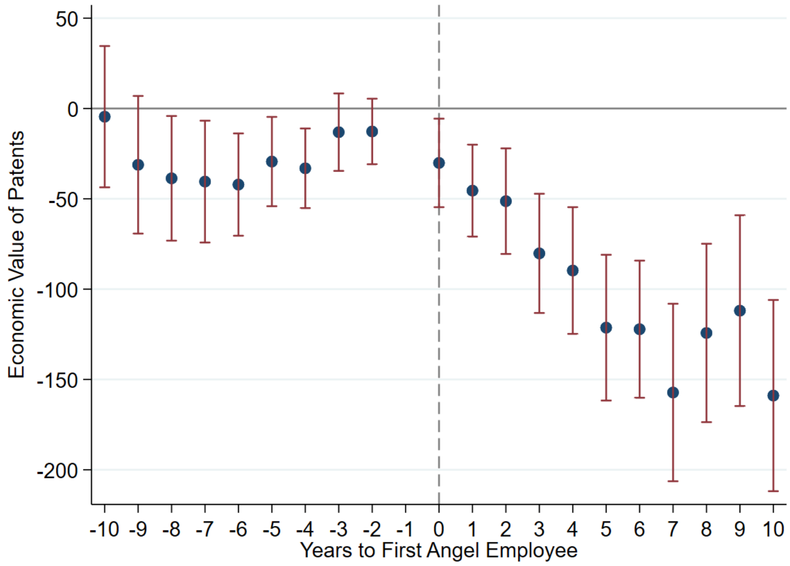
$$Innovation_{i,s,c} = \phi \sum_{c=-10}^{c=+9} D_c \times \sum_{c=-10}^{c=+9} \beta_c D_c \times AngelEmployeeDummy_{i,s} + \phi_{i,s} + \epsilon_{i,s,c} \quad (4.3)$$

where *Angel Employee Dummy* is a dummy variable equal to one if there is at least one angel employee at the firm in the state. This dummy variable is interacted with time dummies relative to the first angel employee in the firm. Control observations are the remaining firm-state observations that do not have a co-located angel employee. This specification automatically controls for firm-year fixed effects as we directly compare states with angel employees to those without in the same firm-year. We additionally include Firm-State fixed effects, as well as time fixed effects. We again cluster standard errors on a firm-state level.

The regression is essentially a staggered difference-in-differences where all events are normalized relative to event time. We thus compare treated firm-states, those with an angel employee presence, with never-treated firm-states. We do so in order to mitigate a negative weighting problem common to many staggered difference-in-differences settings (Baker, Larcker, and Wang 2022).

Figure 4.2.: Event Study: Effect of Angel Employees on Corporate Innovation This figure reports the result of the fixed effect panel regression of equation 4.3 in a firm-state-year event time panel. The dependent variable in Panel A is the economic value of patents (Kogan et al., 2017). In Panel B, we alternatively use truncation-adjusted citation-weighted patents. We visualize the coefficient on time dummies interacted with the dummy variable, *Angel Employee Dummy*, which is equal to one if there is at least one angel employee at the firm in the state. Time is relative to the first angel employee on a firm level. The coefficients are normalized to the year before the first angel employee. Variable definitions are provided in section A in the Appendix. The regression includes Firm-State and Firm-Year fixed effects. Standard errors are clustered by Firm-State. Confidence intervals are at the 5% level.

Panel A: Economic Value of Patents



Panel B: Citation-Weighted Patents

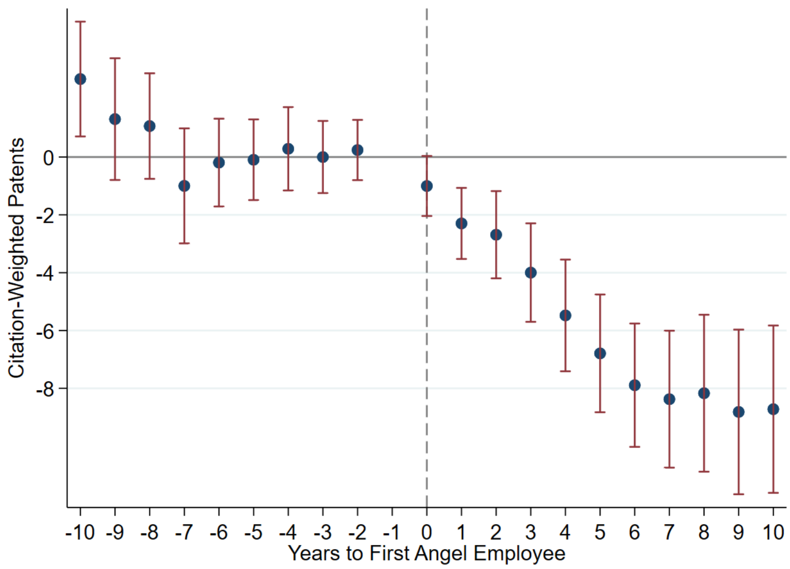


Figure 4.2 visually presents the results of the difference-in-differences event study. The presence of at least one angel employee in a state decreases the economic value of innovation. There is a small effect in the treatment year, however the coefficient turns negative and statistically significant afterwards, increasing over time. There are no pre-trends. The negative effect is prominent only *after* the presence of angel employees. To some extent this reduces concerns that reverse causality drives the effect. We repeat the event study using citation weighted patents in Panel B with similar results.

One concern here is that the results are driven purely by the headquarter state of the firm. We do not know the headquarter state for all firms in the sample, however we repeat the analysis and drop the most important state for each firm as measured by innovation output. Indeed, and as expected for senior level employees, the majority of angel employees are located in the biggest state. We visualize the result of the analysis in figure A1. The results are similar however the economic significance is, as expected, significantly lower in economic magnitudes. The negative effect of angel employees is still present and thus not purely driven by senior employees in the headquarter state.

4.3.4. Instrumental Variable Regression: Competition from Venture Capital

Our previous results address some endogeneity concerns, however we are unable to rule out omitted variable bias for such an individual choice like conducting angel investments. In an ideal experiment, we would want to randomize angel employees across US firms. As this is not feasible, we use an instrumental variable approach to complement our previous analyses. We base the instrument on competition in the early-stage financing market. Recent literature has shown, theoretically (Hellmann and Thiele 2015) and empirically (Hellmann, Schure, and Vo 2021), that venture capital and angel financing can be substitutes. Our descriptive statistics also show that investments by angel employees tend to be large. Thus, to some extent, they might compete with traditional venture capital funds. Our instrumental variable regressions exploit an arguably exogenous variation provided by increased competition in the early-stage financing market.

We exploit the staggered passage of the so-called PIR, the "prudent man rule" (González-Uribe 2020). These regulatory-induced changes lead to an inflow into venture capital of state pension funds. Specifically, state pension funds were mandated to apply modern portfolio theory and invest as a prudent investor should, and thus include local private equity in a diversified portfolio. González-Uribe (2020) shows an economically significant 54% increase in capital commitment after a state passes the PIR. We use this staggered adoption in a 2SLS regression, where the inflow of venture capital serves as a plausibly exogenous instrument which crowds out angel employees. There are two necessary conditions: relevance and the exclusion restriction.

The relevance condition is likely satisfied for several reasons. First, due to their large size, state pension funds are one of the largest investors in private equity (González-Uribe 2020). Second, the literature and summary statistics show that angel investments tend to be local. There is also a strong local bias for pension funds (Hochberg and Rauh 2013). Lastly, when more venture capital enters the market and provides financing to early-stage firms, it is unlikely that future financing rounds will involve angel investors. We test the relevance condition formally in a first stage regression. Consistent with a crowding-out hypothesis, the staggered passage of the PIR negatively predicts angel employees. The F-statistic is equal to 52 and decreases for longer time periods. The F statistic is thus higher than the conventional threshold of 10, but short of recently proposed thresholds such as Lee et al. (2022).

The exclusion restriction is a necessary assumption that our instrument does not directly affect future innovation of the employers of angel employees. We argue that this assumption is plausible for several reasons: First, the source of funds in the PE market comes from pension funds which are often constrained or have a home bias to invest locally (Cumming and Dai, 2010; Hochberg and Rauh, 2013; González-Uribe, 2020). This variation in fund raising is likely to be unrelated to corporate innovation. Second, Hirukawa and Ueda (2011) show that VC has no effect on innovation, whereas Kortum and Lerner (2000) show that VC investments create significant *positive* spillovers. More recently, Howell et al. (2020) show that VC investments tend to be pro-cyclical rather than counter-cyclical. Such evidence biases us against finding a negative effect of our instrument on corporate innovation.

Table 4.5 presents the results of the regression. The point estimates are statistically significant and negative across the specifications and panels. The evidence largely confirms our previous results.

Table 4.5.: Instrumental Variable Regression

This table reports the result of a instrumental variable regression similar to equation 4.2. The dependent variable is innovation output over the subsequent k years ($Innovation_{t+k}$), where $k = [1, 5]$, respectively. In Panel A it is the economic value of patents (Kogan et al., 2017). In Panel B, we alternatively use truncation-adjusted citation-weighted patents. The independent variable, *Angel Employee Dummy*, is equal to one if there is at least one angel employee at the firm in the state. Alternatively it is the natural logarithm of the number of angel employees on a firm level. Variable definitions are provided section A in the Appendix. The first stage F statistic is reported. The regression includes Firm-Year and Firm-State fixed effects. Standard errors are clustered by Firm-State. ***, ** and * represents significance at the 1%, 5% and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)	(3)	(4)	(5)
Panel A – Dependent variable: Economic Value of Patents ($KPSS$)					
	$KPSS_{t+1}$	$KPSS_{t+2}$	$KPSS_{t+3}$	$KPSS_{t+4}$	$KPSS_{t+5}$
Angel Employee Dummy	-500.05**	-464.63*	-570.16**	-780.35**	-734.76*
	(-2.12)	(-1.89)	(-2.08)	(-2.24)	(-1.66)
ln(1+Angel Employees)	-465.88**	-431.43*	-527.67**	-715.42**	-675.18*
	(-2.11)	(-1.89)	(-2.10)	(-2.29)	(-1.72)
Panel B – Dependent variable: Citation-Weighted Patents (CIT)					
	CIT_{t+1}	CIT_{t+2}	CIT_{t+3}	CIT_{t+4}	CIT_{t+5}
Angel Employee Dummy	-47.24***	-48.06***	-47.67***	-65.56***	-80.12***
	(-3.57)	(-3.46)	(-3.14)	(-3.21)	(-2.76)
ln(1+Angel Employees)	-44.01***	-44.62***	-44.12***	-60.11***	-73.63***
	(-3.58)	(-3.49)	(-3.20)	(-3.36)	(-2.97)
Observations	330,956	311,488	292,020	272,552	253,084
Firm-State FE	YES	YES	YES	YES	YES
Firm-Year FE	YES	YES	YES	YES	YES
1st Stage F-Stat	52.1	46.9	40.4	30.5	20.2

The economic magnitudes of our instrumental variable analyses are large. There are several reasons for this. As discussed before, angel investment activity is primarily unobserved and hence measured with error. Given the data available to researchers, we are likely to capture a lower bound of angel employee activity. It is likely that angel investments within the firm are followed by even more investments due to internal spillover effects of investment decisions similar to Ouimet and Tate (2020). The instrumental variable regressions may additionally pick up this unobserved angel investment activity and account for measurement error. We also note that, assuming unbiasedness, the OLS regression estimates an average treatment effect, whereas the IV estimates a *local* average treatment effect. For our setting, it is plausible that the subset of employees who are dis-incentivized to invest in early-stage firms (compliers) due to competition from VC are the ones that are likely to have a higher negative impact. The subset of employees who are likely to invest regardless of competition in the early-stage financing market (never-takers) might have a weaker localized effect.

Our magnitudes are of economic significance but smaller compared to the effect of CEO characteristics on innovation. One example is the importance of independent boards in Balsmeier, Fleming, and Manso (2017). Another example is the effect of inventor CEOs in Islam and Zein (2020), who find that firms with inventor CEOs have 66% more patent applications, 117% more citations, and 122% higher economic value of patents. Given that angel employees are likely not as influential as CEOs, it is not surprising that our estimates are smaller in magnitude.

4.4. Channels

We present evidence on two novel channels. First, we provide a string of evidence that suggests agency conflicts are a channel for the negative relationship between angel employees and future firm innovation. The second channel we highlight is that angel employees are more likely to leave the firm and the loss of highly skilled human capital is costly for employers.

4.4.1. Agency conflicts

Angel employees' attention is a limited resource and agents strategically allocate time and effort to their tasks. Essentially, an angel employee faces the trade-off between exerting effort at her corporate employer and her personal investments.

Ex-Post Successful Start-ups

Angel investments are characterized by high risk and potentially high reward. Angel investors often receive so-called *homerun* returns of more than 100% of their initial investment (Wiltbank et al. 2009). Such a risky endeavor might incentivize angel employees to spend significant time selecting and/or monitoring their own investments rather than exerting effort at their corporate employer. In the following, we hypothesize that relatively more successful investments, i.e., which did not fail, should lead to a stronger negative effect. This can be due to several reasons. The investment duration in a non-failed start-up is likely longer compared to a start-up that fails. For ongoing investments, angels might be engaged with their start-ups to help them succeed. Additionally, if some angels obtain board seats, there will be time intensive board meetings for follow-on financing rounds, or if the start-up is ultimately acquired, or goes public. Finally, for relatively more successful start-ups, angel employees might be particularly engaged and exert effort to select such investment opportunities. Overall, we expect the negative relationship to be stronger for firms associated with relatively more successful (non failed) start-ups.

We incorporate ex-post information about a start-up's status to test whether results are driven by ex-post non failed start-ups. We mark start-ups as failed if they are flagged as defunct or did not receive additional funding in the last 5 years. We take the number of failed and non-failed start-ups of all employees' investments for each corporation in a given year. We then run the baseline specification of equation 4.1. We replace the *Angel Employees* variable with *Non-Failed Start-ups* and *Failed Start-ups* which is the natural logarithm of the number of non-failed and failed start-ups per firm-state per year, respectively.

Table 4.6.: Effect of Relatively Successful Start-ups
This table reports the result of the fixed effect regression of equation 4.2. The unit of observation is on a firm-state-year level. The dependent variable in Panel A is the economic value of patents scaled by assets (Kogan et al., 2017) over the subsequent k years ($Innovation_{t+k}$), where $k = [1, 5]$, respectively. In Panel B, we alternatively use truncation-adjusted citation-weighted patents. The independent variable of interest is split into two parts, depending on whether an angel employee is linked to a failed or non-failed startup. We categorize a start-up as failed when the startup is either defunct of did not receive financing in the last 5 years. Variable definitions are provided in section A of the Appendix. The regression includes Firm-Year, Firm-State, and State-Year fixed effects. Standard errors are clustered by Firm-State. ***, ** and * represents significance at the 1%, 5% and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)	(3)	(4)	(5)
Panel A – Dependent variable: Economic Value of Patents ($KPSS$)					
	$KPSS_{t+1}$	$KPSS_{t+2}$	$KPSS_{t+3}$	$KPSS_{t+4}$	$KPSS_{t+5}$
Non-Failed Start-ups	-10.57	-42.09***	-62.26***	-95.27***	-108.08***
	(-0.77)	(-2.67)	(-3.84)	(-5.01)	(-4.78)
Failed Start-ups	42.68***	15.20	-23.90**	-40.07***	-69.92***
	(4.09)	(1.40)	(-2.13)	(-3.58)	(-5.53)
Panel B – Dependent variable: Citation-Weighted Patents (CIT)					
	CIT_{t+1}	CIT_{t+2}	CIT_{t+3}	CIT_{t+4}	CIT_{t+5}
Non-Failed Start-ups	-1.57**	-2.93***	-4.73***	-5.87***	-5.97***
	(-2.44)	(-4.12)	(-6.12)	(-6.88)	(-5.71)
Failed Start-ups	1.84***	0.07	-1.58***	-2.97***	-4.25***
	(3.76)	(0.14)	(-3.18)	(-5.59)	(-7.60)
Observations	330,956	311,488	292,020	272,552	253,084
Firm-Year FE	YES	YES	YES	YES	YES
Firm-State FE	YES	YES	YES	YES	YES
State-Year FE	YES	YES	YES	YES	YES

Table 4.6 documents our results. As before in Panel A, we document the results using the economic value of patents as our dependent variable. In Panel B, we use truncation adjusted citation weighted patents as our dependent variable. In both panels, we consistently observe that links to non-failed start-ups are associated with lower future firm innovation. The effect for non-failed start-ups is more pronounced and is statistically different from the coefficient on failed start-ups. The coefficient for failed start-ups is still negative in the long run.

Hence, the results show that the negative effects are most pronounced when the early-stage firms of angel employees are relatively successful. The results are also consistent with the explanation that highly skilled senior employees are affected by agency conflicts.

Incentives to Invest: Evidence from the Small Business and Jobs Act 2010

To identify the effect of angel employees' *incentives* to engage with their portfolio start-ups, we exploit the passage of The Small Business and Jobs Act 2010 (SBJA). This regulatory change presents a plausibly exogenous shock to angel employees' incentives to be more involved with their invested start-ups. The SBJA allows investors to exclude 100% of the eligible gain from qualified small business stock (QSBS) upon sale or exchange from September 27, 2010 onwards (Edwards and Todtenhaupt 2020). To qualify as a QSBS, the firm must be listed as a domestic C-corporation and have less than $50M in total assets. Exemption from capital gains taxes is granted if an angel investor holds her investment for at least 5 years. Some industries are excluded, however, almost all start-ups in our sample are in treated industries.[7]

This regulatory change provides us with a unique setting to test some predictions using our data. In principle, if angel investments are tax-exempt, conditional on being an

[7]Explicitly excluded are, for example, financial services, accounting, law, farming, hotels, among others.

angel investor, there are more incentives to divert time and effort as future capital gains are higher. We argue that if agency conflicts indeed drive the observed negative effect, capital gains tax exemption of angel investments should lead to a stronger agency conflict and hence a more negative effect. We note that our objective is not to randomly allocate angel employees across firms. Rather, we attempt to disentangle the *incentives* of angel employees to engage with their portfolio firms and study the effect of higher angel incentives on employer innovation output. In order to test our hypothesis, we run a difference-in-difference regression, shown in Equation 4.4.

$$Innovation_{i,s,t+k} = \beta_1 \times Treated_{i,s,t} + \beta_2 \times Treated_{i,s,t} \times Post_t + \theta_{i,t} + \phi_{i,s} + \psi_{s,t} + \epsilon_{i,s,t} \quad (4.4)$$

where we identify treated firm-state-years as any firm-state-year with the presence of angel employees. $Treated_{i,s,t}$ takes the value of 1 if there is at least one angel employee in a firm-state-year and zero otherwise. *Post* is a dummy equal to one in the years after 2010.

The results of the analysis are shown in table 4.7. Panel A presents the results with the economic value of patents as the dependent variable. Our results are driven by angel employees after the SBJA 2010 came into effect. The coefficient on the double interaction term is negative and statistically significant across specifications. In panel B, we repeat the same analysis with citation-weighted patents and reach similar conclusions.

Table 4.7.: Evidence from the SBJA Capital Gains Exemption
This table reports the result of the fixed effect regression of equation 4.4. The unit of observation is on a firm-state-year level. The dependent variable in Panel A is the economic value of patents scaled by assets (Kogan et al., 2017) over the subsequent k years ($Innovation_{t+k}$), where $k = [1, 5]$, respectively. In Panel B, we alternatively use truncation-adjusted citation-weighted patents. We interact the variable, *Angel Employee Dummy*, with Post, which is equal to one for all years after the passage of the SBJA. Variable definitions are provided in section A of the Appendix. The regression includes Firm-Year, Firm-State, and State-Year fixed effects. Standard errors are clustered by Firm-State. ***, ** and * represents significance at the 1%, 5% and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)	(3)	(4)	(5)
Panel A – Dependent variable: Economic Value of Patents ($KPSS$)					
	$KPSS_{t+1}$	$KPSS_{t+2}$	$KPSS_{t+3}$	$KPSS_{t+4}$	$KPSS_{t+5}$
Angel Employee Dummy	51.20***	48.79***	42.28***	32.75***	16.24*
	(3.80)	(4.93)	(4.48)	(3.30)	(1.80)
Angel Employee Dummy × Post	-53.15***	-81.27***	-103.75***	-121.17***	-138.01***
	(-4.35)	(-6.27)	(-8.52)	(-9.22)	(-8.01)
Panel B – Dependent variable: Citation-Weighted Patents (CIT)					
	CIT_{t+1}	CIT_{t+2}	CIT_{t+3}	CIT_{t+4}	CIT_{t+5}
Angel Employee Dummy	3.29***	2.72***	1.85***	0.80	0.33
	(6.68)	(6.93)	(3.15)	(1.37)	(0.87)
Angel Employee Dummy × Post	-4.35***	-5.19***	-5.75***	-5.91***	-6.68***
	(-7.91)	(-8.39)	(-6.07)	(-4.32)	(-4.72)
Observations	330,956	311,488	292,020	272,552	253,084
Firm-Year FE	YES	YES	YES	YES	YES
Firm-State FE	YES	YES	YES	YES	YES
State-Year FE	YES	YES	YES	YES	YES

The results in this section are consistent with an agency conflict based explanation behind our main results. Angel employees seem to divert more effort from their employers when there are regulatory-induced higher ex-ante *incentives* to benefit from their investments.

4.4.2. Angel Employee Exit: Loss of Human Capital

In the following we provide evidence that angel employees are more likely to leave their employer. We argue that the loss of these individuals might be costly for their employers.

$$Exit_{i,t} = \beta \times ActiveAngelEmployee_{i,t} + \theta_i + \phi_{ft} + \epsilon_{i,t} \tag{4.5}$$

where $Exit_{i,t}$ is a dummy variable equal to one if the employee leaves the firm in a given year. *Active Angel Employee* is a dummy variable equal to one if the employee is around this time an active angel investor. This specification allows to analyze an within-individual comparison as we can compare an angel employee during the time of investment to before or after. Thus, θ is an employee fixed effect. ϕ is a firm-year fixed effect.

We thus assess exit rates of (angel) employees over time in table 4.8. Active angel employees are significantly more likely to exit their firm. We see a 2% increased exit rate. Due to high employee replacement costs especially of highly skilled employees, this loss of human capital is likely detrimental for employers.

Table 4.8.: Loss of Human Capital: Exit of Angel Employees
This table reports the fixed effect regression of equation 4.5. The unit of observation is on a firm-year level. The independent variable, $Exit_{t+1}$ is a variable equal to one if the employee has exited the firm. The dependent variable *Active Angel Employee* is equal to one if the employee is currently an active angel employee, defined as investing in early-stage firms in the last 5 years. Variable definitions are provided in the Appendix. The regression includes Firm-Year and Employee fixed effects. Standard errors are clustered by Firm-Year. ***, ** and * represents significance at the 1%, 5% and 10% level, respectively. t-statistics are displayed in parenthesis.

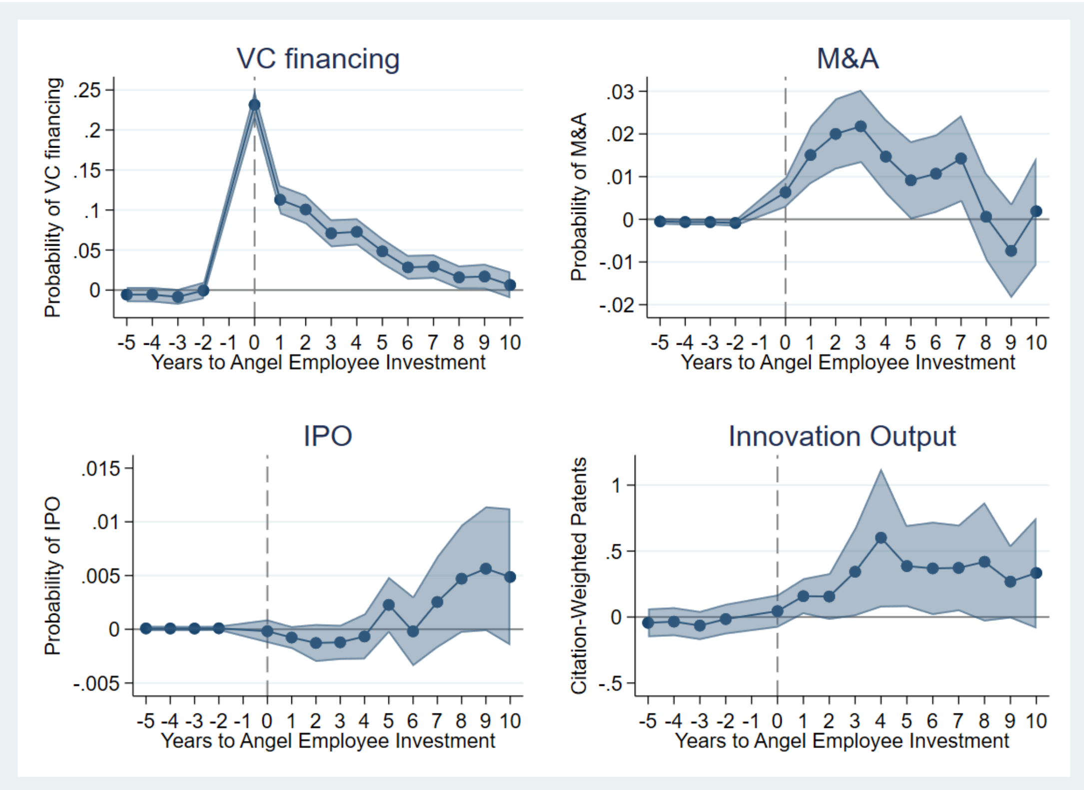
	(1)
Dependent variable:	$Exit_{t+1}$
Active Angel Employee	0.02***
	(6.07)
Observations	245,304
Firm-Year FE	YES
Employee FE	YES

4.5. Angel Employees are Beneficial for Start-ups

We now turn to the start-up perspective and ask the following question: Is it beneficial for an early stage firm to be financed by an angel employee? Angel employees might be very skilled individuals and might possess valuable industry expertise and networks. They might also either carefully select or treat their investment. Selection would be before the investment and means that angel employees contact entrepreneurs, attend presentations and analyze pitch documents. Treatment would be after the investment and means that angel employees attend meetings, establish connections and give advice to the early-stage firm. To analyze this, we compare the future success of start-ups financed by angel employees to those financed by other angel investors in the Crunchbase sample.

We regress the presence of angel employees on the probability of start-up success and future innovation of the startup. We run a difference-in-differences regression to look at the effect of angel employees' presence. We measure start-up success with four separate variables: 1) a dummy equal to one if the early-stage firm receives venture capital financing, 2) a dummy variable equal to one if the early-stage firm was ultimately acquired (M&A), 3) a dummy variable equal to one if the early-stage firm went public (IPO), and 4) future innovation output of the early-stage firm as measured by citation-weighted patents. We look at the effect in an event time framework normalized to one year before investment. We include start-up and time fixed effects and cluster standard errors on a start-up level.

Figure 4.3.: Effect of Angel Employees on Start-up Success
These figures report the result of difference-in-difference regressions on the relationship between angel employee investment and start-up success. The dependent variables, from top left to bottom right, are: a VC financing dummy, an M&A dummy, an IPO dummy, and citation-weighted patents on a startup-year level. We visualize the coefficient on time dummies interacted with a treatment dummy, defined as equal to one if the startup is financed by an angel employee from our sample. Control startups are those financed by other angel investors from the Crunchbase universe. The coefficients are normalized to the year before the first angel investment. All regressions include *Startup* as well as *Time* fixed effects. Standard errors are clustered by start-up. ***, ** and * represents significance at the 1%, 5% and 10% level, respectively. *t*-statistics are displayed in parenthesis.



Compared to the universe of angel-backed start-ups, the presence of angel employees increases future success of start-ups. We first note that receiving financing from angel employees significantly increases the chances of receiving venture capital financing. We also see an increased probability that the start-up will be acquired. We see significantly positive effects until 7 years after angel employee investment and the probability is increased by between 1% and 2% annually. For IPOs, there does not seem to be any economically sizable effect. Lastly, angel employee participation is related to higher future innovation output. The economic effects are equal to around 0.4 more citation-weighted patents, an effect that prevails in the long run.

These results suggest that angel employees have an economically meaningful positive effect on their portfolio firms. This highlights an important source of heterogeneity among angel investors: Those angel investors with industry expertise tend to be the ones who have a positive effect on future success. We stress that we are unable to separate between treatment and selection in this analysis. Angel employees might be able to carefully select promising early-stage firms on (unobservable) characteristics. They might also provide valuable advice to help them succeed.

Additionally, these results also shed some light on the underlying mechanism behind our main results. These results are inconsistent with a "low quality employee" hypothesis, specifically, that the negative effects on innovation output of the employer is due to low quality employees. These results are rather consistent with the view that a firm's future innovation suffers particularly because its highly skilled workforce is diverting time and effort from their corporate employer to their personal investments.

Overall, we present both sides of the coin. On the one hand, angel employees are detrimental for corporate employers' innovation. On the other hand, start-ups seem to benefit from angel employees' participation. This raises the question of welfare implications, which we address later on in the paper.

4.6. Robustness

4.6.1. Evidence from Angel Roles: Innovation-related Angels

In the following, we introduce one source of heterogeneity: the role of angel employees at their employers. We analyze whether the observed negative effect on innovation output is more pronounced when the angel employees are closely related to innovation. To do so, we split the total number of angel employees into those that are likely to work in innovation-related functions and those that are not. As described earlier, based on the textual title information of each employee, we tag employees with the words "product", "innov", "research", among other keywords as innovation-related and angel employees with titles such as "finance" or "legal" as non-innovation related angels. We then run the baseline regression with the key independent variable split into two parts: two indicator variables measuring innovation-related and non-related angel employees separately.
Panel A of table 4.9 shows that the results are largely driven by the sub-sample of angel employees who are innovation related. The economic magnitudes for innovation-related angel employees are higher compared to the baseline specification. Non-innovation related angel employees do not have an effect. There is, however limited power to detect significant effects for non-innovation related angel employees. The majority of employees in the sample have titles such as "engineer", "tech", "product" which we classify as innovation related. The remainder of the sample with titles such as "legal" or "finance" only makes up a small fraction of angel employees.

Table 4.9.: Innovation-related Angel Employees
This table reports the result of the fixed effect regression of equation 4.2. The unit of observation is on a firm-state-year level. The dependent variable in Panel A is the economic value of patents scaled by assets (Kogan et al., 2017) over the subsequent k years ($Innovation_{t+k}$), where $k = [1, 5]$, respectively. In Panel B, we alternatively use truncation-adjusted citation-weighted patents. The dependent variable of interest is split into two parts, depending on whether the angel employee is working in an innovation-related role or not depending on some keywords such as "product manager", "technology", "researcher", etc. Variable definitions are provided in section A of the Appendix. The regression includes Firm-Year, Firm-State, and State-Year fixed effects. Standard errors are clustered by Firm-State. ***, ** and * represents significance at the 1%, 5% and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)	(3)	(4)	(5)
Panel A – Dependent variable: Economic Value of Patents ($KPSS$)					
	$KPSS_{t+1}$	$KPSS_{t+2}$	$KPSS_{t+3}$	$KPSS_{t+4}$	$KPSS_{t+5}$
Innovation related Angels	9.32	-19.61	-58.82***	-79.59***	-97.05***
	(0.79)	(-1.60)	(-4.55)	(-6.04)	(-6.19)
Non-Innovation related Angels	63.21*	72.11*	51.89	48.92	30.02
	(1.68)	(1.95)	(1.38)	(1.39)	(0.67)
Panel B – Dependent variable: Citation-Weighted Patents (CIT)					
	CIT_{t+1}	CIT_{t+2}	CIT_{t+3}	CIT_{t+4}	CIT_{t+5}
Innovation related Angels	0.48	-1.95***	-3.71***	-4.09***	-4.87***
	(0.88)	(-3.56)	(-6.31)	(-7.12)	(-6.90)
Non-Innovation related Angels	0.41	0.24	0.41	-1.10	-0.02
	(0.24)	(0.13)	(0.25)	(-0.73)	(-0.01)
Observations	330,956	311,488	292,020	272,552	253,084
Firm-Year FE	YES	YES	YES	YES	YES
Firm-State FE	YES	YES	YES	YES	YES
State-Year FE	YES	YES	YES	YES	YES

4.6.2. Non-Patent-Based Measures of Innovation

Another possible concern is the use of patents as a measure for firm innovation. After successful innovation, a corporation faces the challenge to either patent or keep the invention secret (trade secret). Since our dependent variable only captures disclosed patents, if the most valuable inventions are not disclosed and protected due to low imitation costs, then this would lead to a measurement error in our estimates. It can also be the case that firms do not file patents, but are still innovative (Koh et al., 2021). Therefore, we obtain data on three non-patent based outcome variables: trademarks, new product launches, and scientific publications. Firms have strong incentives to file trademarks and launch new products. Compared to patents, there is less substitution with trade secrets.

If innovation output is reduced, then one would expect to find fewer trademarks, new product launches, and scientific publications. Indeed, that is what we find. Angel employees are associated with fewer new product launches over the subsequent 1 to 5 years in Panel A of table 4.10. The point estimates suggest that a firm-year with at least one angel employee is associated with approximately 10% fewer product launches after three years ($= exp^{-0.11}$). Similar conclusions can be drawn from trademarks in Panel B, and scientific publications in Panel C of table 4.10.

Table 4.10.: Non-Patent Based Innovation Output
This table reports the fixed effect regression of equation 4.1 in a firm-year panel. The dependent variable in Panel A, (*NPA*) is the natural logarithm of one plus the number of new product announcements of the firm. In Panel B, the dependent variable is the natural logarithm of one plus the total number of trademarks (TM) of the firm. In Panel C, the dependent variable is the natural logarithm of one plus the total number of scientific publications ($PUBS$) of the firm. The dependent variable Angel Employee Dummy is equal to one if the firm employs at least one angel employee. The regression includes 15 standard control variables. Variable definitions are provided in the Appendix. The regression includes Firm and Year fixed effects. Standard errors are clustered by firm. ***, ** and * represents significance at the 1%, 5% and 10% level, respectively. t-statistics are in parenthesis.

Panel A: New Product Announcements					
	NPA_{t+1}	NPA_{t+2}	NPA_{t+3}	NPA_{t+4}	NPA_{t+5}
Angel Employee Dummy	-0.05	-0.08**	-0.11***	-0.07	-0.10*
	(-1.56)	(-2.23)	(-4.05)	(-1.67)	(-1.94)
Panel B: Trademarks					
	TM_{t+1}	TM_{t+2}	TM_{t+3}	TM_{t+4}	TM_{t+5}
Angel Employee Dummy	-0.07**	-0.11**	-0.17***	-0.12**	-0.11***
	(-2.23)	(-2.85)	(-3.97)	(-2.31)	(-2.00)
Panel C: Scientific Publications					
	$PUBS_{t+1}$	$PUBS_{t+2}$	$PUBS_{t+3}$	$PUBS_{t+4}$	$PUBS_{t+5}$
Angel Employee Dummy	-0.02	-0.18***	-0.32***	-0.33***	-0.41***
	(-1.31)	(-2.64)	(-3.64)	(-3.12)	(-3.19)
Other Controls	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

4.6.3. Private Firms

Our analysis so far has focused on the effect of angel employees on publicly listed corporations. We generalize the negative relationship and alternatively only consider private employers. We match patent data to angel employees and to their private employers.

Table 4.11.: Angel Employees at Private Firms
This table reports the fixed effect regression of equation 4.1 in a firm-year panel. The dependent variable is the economic value of patents scaled by assets (Kogan et al., 2017) over the subsequent k years ($Innovation_{t+k}$), where $k = [1, 5]$, respectively. The dependent variable Angel Employee Dummy is equal to one if the firm employs at least one angel employee. The sample is composed of all private firms in the US obtained from ORBIS. We limit ourselves to firms with turnover of at least 10M$. Due to limited data availability of private firms, the regression does not include control variables. Variable definitions are provided in the Appendix. The regression includes Firm and Year fixed effects. Standard errors are clustered by firm. ***, ** and * represents significance at the 1%, 5% and 10% level, respectively. t-statistics are displayed in parenthesis.

	CIT_{t+1}	CIT_{t+2}	CIT_{t+3}	CIT_{t+4}	CIT_{t+5}
Angel Employee Dummy	0.01	-0.04*	-0.08***	-0.08***	-0.13***
	(0.23)	(-1.95)	(-3.57)	(-3.55)	(-3.84)
Observations	2,349,209	2,338,687	2,323,400	2,146,491	1,970,683
Firm FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

The results are shown in table 4.11. The effect of angel employees is negative and statistically significant after year 2. By and large, the results are comparable to the results for publicly listed firms. Our observed negative effect thus generalizes to private firms.

4.6.4. Excluding Recent IPO Years

Recent IPOs can provide employees with more liquidity to conduct angel investments. At the same time, going public could have a detrimental impact on the innovation of firms (Bernstein, 2015). Hence, such a mechanical association could bias our inferences. To address such concerns, in table 4.12, we re–run our baseline regressions by excluding firm–years that are immediately within two years of the IPO–year for a firm.

The results do not change. In fact, some of the point estimates become larger compared with those in table 4.3 suggesting that an endogenous association between IPOs and employee liquidity is unlikely to explain our results.

Table 4.12.: Baseline Regressions excluding recent IPOs: Angel Employees and Innovation Output

This table reports the fixed effect regression of equation 4.1. The unit of observation is on a firm-year level. We drop the first two years of observations after a firm went public. The dependent variable in Panel A is the economic value of patents scaled by assets (Kogan et al., 2017) in the next year ($Innovation_{t+1}$). In Panel B, we alternatively use truncation-adjusted citation-weighted patents. The independent variable, *Angel Employee Dummy*, is equal to one if there is at least one angel employee at the firm. Alternatively it is the natural logarithm of the number of angel employees on a firm level. Columns (3) and (4) use a inverse hyperbolic sine transformation of the dependent variable. Column (5) only looks at firms which patent. The regression includes 15 standard control variables. Variable definitions are provided in section A of the Appendix. The regression includes Firm and Year fixed effects. Standard errors are clustered by firm. ***, ** and * represents significance at the 1%, 5% and 10% level, respectively. t-statistics are displayed in parenthesis.

	(1)	(2)	(3)	(4)	(5)
Panel A – Dependent variable: Economic Value of Patents					
	$KPSS_{t+1}$	$KPSS_{t+1}$	$KPSS_{t+1}$	$KPSS_{t+1}$	$KPSS_{t+1}$
Angel Employee Dummy	-0.04***		-0.04***		-0.04***
	(-4.61)		(-4.62)		(-3.59)
ln(1+Angel Employees)		-0.07***		-0.07***	
		(-5.77)		(-5.83)	
Panel B – Dependent variable: Citation-Weighted Patents					
	CIT_{t+1}	CIT_{t+1}	CIT_{t+1}	CIT_{t+1}	CIT_{t+1}
Angel Employee Dummy	-0.28***		-0.11***		-0.25***
	(-5.05)		(-4.39)		(-3.59)
ln(1+Angel Employees)		-0.45***		-0.16***	
		(-6.72)		(-5.56)	
Observations	56,823	56,823	56,823	56,823	26,283
Controls	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

4.6.5. Outsourcing

We look at interactions between the employer and the start-ups themselves. An alternative interpretation to our findings is that corporations might actively push their employees to invest in early-stage firms. We call this interpretation *outsourcing*. It would be problematic if some innovation is outsourced and not situated within the boundary of the employer. A string of evidence shows that this hypothesis is unlikely. First, there are few common investments between angel employees and corporate venture capital units. We do, however for future financing rounds of the startups, observe participation by other venture capital funds. This indicates that the linked start-ups have financing needs that are not met by the employees' corporations. Second, we analyze exchanges between corporations and angel employees' start-ups. We find very few citations from the corporation to the start-up and the other way around. Angel employees' investments seem to be mostly personal and no knowledge (as measured by citations) flows between start-up and corporation. We do observe a small number of M&A transactions and see negative announcement effects on average. This could potentially be consistent with our evidence if employees are cashing out and the acquisition is the manifestation of an agency conflict similar to Benson and Ziedonis (2010). However the sample size and the number of transactions is too small to make substantiated claims.

4.6.6. Welfare Analysis

Our analysis indicates that angel employees have a negative effect on the innovation output of their employers, but a positive one on their portfolio startups. This raises the question of welfare effects. We perform a back-of-the-envelope calculation to assess overall welfare of these two effects. We attempt an apples-to-apples comparison by only considering citation-weighted patents, a measure available on both sides. Ultimately, we compare the reduced innovation output on the employer level to the increased innovation output on the start-up level.

From column (4) in table 4.4, i.e., four years after angel employee presence, we obtain a coefficient of -3. Thus the presence of angel employees reduces citation-weighted patents by -3 truncation-adjusted citation-weighted patents. In total, we observe 502 *first–time* presence of angel employees in the firm-state panel. Hence, taking the 3 reduction in citation-weighted patents per angel employee, we compute a total loss of around 1,506 truncation-adjusted citation-weighted patents. This is the yearly aggregate innovation loss for the economy due to angel employees.

We perform a similar exercise at the start-up level. From the regressions in figure 4.3 we obtain a coefficient of 0.4 in the long run for citation-weighted patents. This is the effect of angel employee participation on citation-weighted patents on a startup level. The start-ups in our sample patent at a much lower rate compared to publicly listed firms. A benefit from the startup point of view is scalability: We have 1,845 angel employees in our sample and an angel employee on average invests in 2.7 startups. We thus compute an increase of $(0.4 \times 1,845 \times 2.7 =)$ 1992.6 citation-weighted patents. This number is a bit higher compared to the innovation lost at publicly traded firms.

This back-of-the-envelope estimate suggests that the welfare effects when only considering citation-weighted patents are positive. We emphasize the crude nature of these calculations as we cannot consider other negative or positive effects of angel employees in these calculations. Furthermore, we only consider innovation while leaving out aspects such as IPOs, M&As, employment, industry competition, and, other non-patent based measures of innovation. Our analysis highlights the need for more research in this direction to better understand welfare implications of more complex interactions in the economy.

4.7. Conclusion

Using novel data linking angel investors to their corporate employers, we find that what we refer to as - angel employees - negatively impact the innovation output of their corporate employer. The negative effect is stronger when incentives of angel investments are higher. We shed light on two novel channels: agency conflicts and the loss of human capital. On the positive side, angel employees seem to have a positive impact on start-up success. Taken together, angel employees trade off time and effort between their employer and their personal start-up investments.

APPENDIX

4.8. Appendix

Variable Definitions This section provides the variable definitions. All variables are measured at an annual frequency. All continuous variables are winsorized at the 1% and 99% level.

This section provides the variable definitions. All variables are measured at an annual frequency. All continuous variables are winsorized at the 1% and 99% level.

1. *Angel Employee* – Either a dummy equal to one if there is at least one angel employee. Alternatively the natural logarithm plus one of the total number of angel employees on a firm level. "Angel Employee" is an individual who is an angel investor and around the time of investment employed at a publicly traded corporation. We assume a holding period of 5 years and aggregate the number of unique individuals on a firm basis. We obtain this variable by combining information on angel financing from Crunchbase and employment information from LinkedIn (plus some manual searches).

2. *Innovation Output* – Either the economic value of patents aggregated on a firm-year level or on a firm-state-year level. On a firm-year level, the variable is scaled by total assets following Kogan et al., 2017. Alternatively, we use truncation-adjusted citation weighted patents as in Hall, Jaffe, and Trajtenberg, 2005. Patents linked to firms is obtained from the website of Noah Stoffman. All other patent data is directly from the United States Patent and Trademark Office (USPTO).

3. *Size* – Natural logarithm of the market value of the firm. The information is obtained from Compustat.

4. *R&D Expenditures* – Total R&D expense scaled by book value of assets. The information is obtained from Compustat.

5. *Tobin's Q* – Book value of assets (AT) + market capitalization (MC) - common equity value (CEQ) - balance sheet deferred taxes, if available (TXDB) / total assets (AT). The information is obtained from Compustat.

6. *Profitability* – Operating income scaled by book value of assets. The information is obtained from Compustat.

7. *Tangibility* – Property, plant and equipment scaled by book value of assets. The information is obtained from Compustat.

8. *Age* – Natural logarithm of the number of years the firm appears in Compustat.

9. *Herfindahl-Index (Squared)* – Industry competition as measured by the Herfindahl index (squared) defined over yearly sales in a 4-digit SIC code. The information is obtained from Compustat.

10. *Liquidity* – Stock liquidity measured as the daily mean bid-ask spread. The information is obtained from CRSP.

11. *Capital Expenditures* – Capital Expenditure scaled by the book value of the firm. The information is obtained from Compustat.

12. *Leverage* – Leverage ratio of the firm's total debt scaled by book value of assets. The information is obtained from Compustat.

13. *Financial Constraints* – Dummy variable indicating Financial Constraints if a firm is flagged as falling in the top tercile of the distribution of financial constraints every year by either of the measures proposed by Kaplan and Zingales, 1997, Whited and Wu, 2006 and Hadlock and Pierce, 2010. The information is obtained from Compustat.

14. *Patent Stock* – Total number of patents assigned to a firm in the last 20 years (equivalent to patent expiry period). The information is obtained from the website of Noah Stoffman.

15. *Number of Employees* – Natural logarithm of the total number of employees. The information is obtained from Compustat.

16. *Corporate Venture Capital* – A dummy variable equal to one if the firm has an active corporate venture capital program. The variable was constructed following Ma, 2020. The information is obtained from Refinitiv (formerly VentureXperts by Thomson Reuters).

17. *staggered PIR* – We obtain this data from González-Uribe, 2020.

18. *Failed Start-ups* – The number of Start-ups that are either defunct or did not receive any financing in the last 5 years. The information is obtained from Crunchbase.

19. *Board/Executive Angel Dummy* – We tag angel employees as board members if they mention "director" or "board member" in their title. We tag them as executives if they mention "executive" or any C-suite abbreviation. The information is based on textual information from the job title on LinkedIn.

20. *Innovation /Non-innovation related Angels* – We tag employees with the words: "product", "innov", "research", "tech", "engineer", among other keywords as innovation-related and angel employees with titles such as "finance", "legal", "accounting", "audit", "operation", "banking", among others into non-innovation related angels. The information is based on textual information from the job title on LinkedIn.

21. *Trademarks* – The log of one plus the total amount of trademarks applied in a given year. We obtain trademarks linked to gvkeys from Heath and Mace, 2020.

22. *Product Announcements* – The log of one plus the total amount of new product launches in a given year. We follow the methodology of Chu et al., 2020 and proxy for new product launches by screening the key developments (Compustat) database for the following keywords: "unveil", "launch", and "new product". We obtain the data from Compustat.

23. *Scientific Publications* – The log of one plus the total amount of scientific publications. We obtain the data from Arora, Belenzon, and Sheer, 2020. We use version 7 (December 2020) available here: https://zenodo.org/record/4320782

24. *Funding Round Size* – This variable is equal to the size of the funding round in million USD. We obtain this data from Crunchbase.

25. *Startup Corporation Distance* – This is the We compute this data by combining Crunchbase location information (when available) with headquarter location info from firms 10-Ks.

26. *Startup-Corporation Industry Similarity* – This is the cosine similarity of textual descriptions of startups with that of their employers similar to Hoberg and Phillips, 2010. We compute this data by combining Crunchbase textual descriptions (when available) with product descriptions from firms 10-Ks.

27. *VC financing/M&A/IPO* – These are indicator variables equal to one if the firm raises venture capital, is acquired, or goes public. We obtain this data from Crunchbase.

28. *Exit* – This variable is equal to one if the employee leaves the firm. We obtain this data from LinkedIn.

Data Description

In the parts below, we provide more details on how we obtained the data used in this paper. We start with a more detailed description of the Crunchbase dataset and then explain how we obtained the employment histories from LinkedIn.

Crunchbase

Crunchbase was the starting point for our data collection. We obtained the data through a private API and used a bulk download on January 1st 2022. The relational database provides information on staged funding rounds, e.g. which company raised how many funds, who participated, and when the investment took place. We first merge the funding round data with information on investments, e.g. which investors participated in which funding round. This provides an overview of who invested in which funding round. Most of the investments are venture capital investments, so the next step is to obtain personal (angel) investments. We do this by merging the dataset with the people database. The people database covers more than 870,000 individuals who have connection to the start-up world. Most individuals in the database are founders, so they are not material to our research. We only keep investments that are tied to individuals. The next step is to limit the dataset to US individuals investing in US firms. Additionally, we manually verified our angel investors. E.g. we eliminated individuals tied to venture capital firms and individuals tied to a corporate venture capital arm of a firm.

Crunchbase also provides information on employment histories in the so-called jobs database. We can therefore see which individual worked in which firm. We initially used this data for preliminary results, but decided that the coverage was not sufficient. We therefore looked for an alternate database which provides more comprehensive employment histories.

Employment History

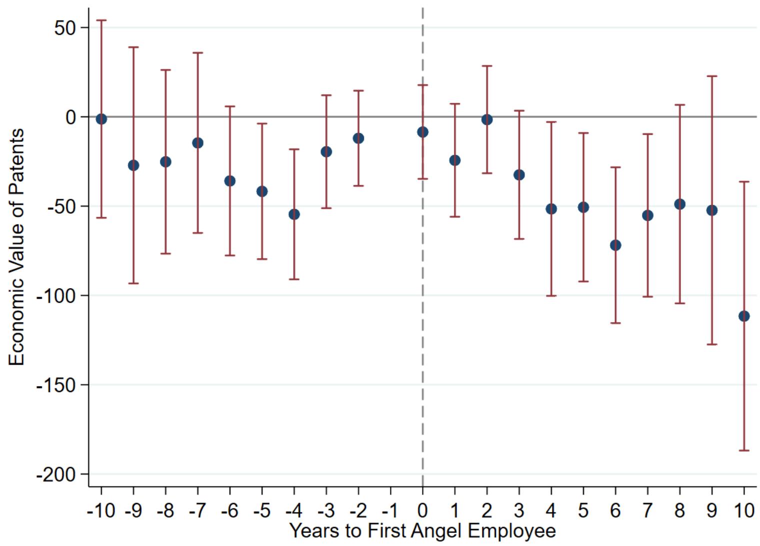
Crunchbase already provided links to individual LinkedIn profiles to the vast majority of investors in our sample. We manually verified whether these links were in fact correct and compared the employment history listed in Crunchbase with the history listed in LinkedIn. For the subset of individuals with missing LinkedIn links, we were able to collect the link for roughly 66% of the remaining subsample. We again verified whether we map the correct individual by comparing employment histories. As mentioned in the paper, we were left with a small set of individuals (who sometimes had many investments) without a LinkedIn profile. This could result in a substantial selection bias if high-level employees are less likely to have a LinkedIn profile. We thus ranked the sample by number of investments and manually obtained employment histories for all individuals with at least 3 angel investments. We were able to find employment information for 98% of all angel investments in our sample.

We also performed a number of cleaning exercises. One can in principle provide any information on LinkedIn. The information is self reported and not independently audited. We remove jobs when the job title refers to being an investor in the firm. For instance, many individuals claim to work for Tesla and state their position as "investor" or "shareholder". We remove these jobs from our data, as it is unlikely that these individuals are decision makers at that firm. Also, many start-up founders are stating their firm name as follows: FIRM NAME (acquired by ACQUIRER). We cleaned the employer name such to make sure that we do not falsely match an investor to a listed ACQUIRER. We performed a string search to look for instances of "M&A", "acquired", "acquisition", etc. to eliminate these instances.

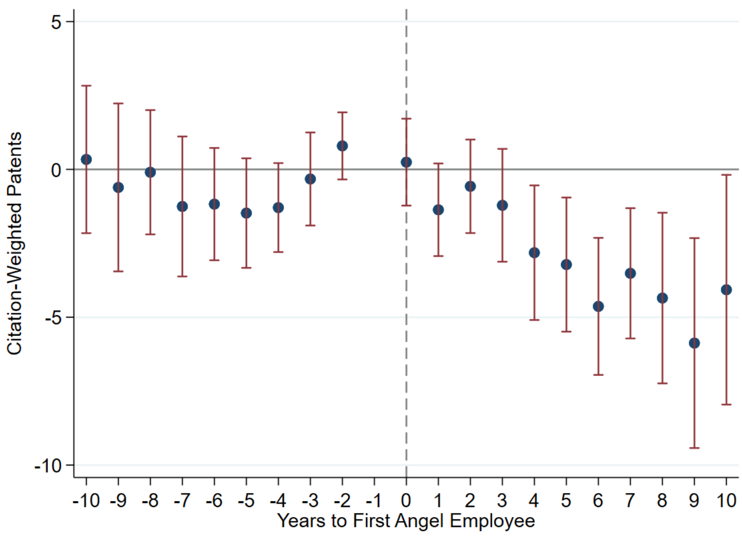
Figure A1.: Event Study: Effect of Angel Employees on Corporate Innovation - Without Biggest State

This figure reports the result of the fixed effect panel regression of equation 4.3 in a firm-state-year event time panel. The dependent variable in Panel A is the economic value of patents (Kogan et al., 2017) over time. In Panel B, we alternatively use truncation-adjusted citation-weighted patents. We visualize the coefficient on time dummies interacted with the dummy variable, *Angel Employee Dummy*, which is equal to one if there is at least one angel employee at the firm in the state. Time is relative to the first angel employee on a firm level. The coefficients are normalized to the year before the first angel employee. The sample here excludes the headquarter state which is proxied by the omission of the biggest state as measured by innovation output. Variable definitions are provided in section A of the Appendix. The regression includes Firm-State and Firm-Year fixed effects. Standard errors are clustered by Firm-State. Confidence intervals are at the 5% level.

Panel A: Economic Value of Patents



Panel B: Citation-Weighted Patents



References

Arora, Ashish, Sharon Belenzon, and Lia Sheer (2020). “DISCERN: Duke Innovation and Scientific Enterprises Research Network [Data set]”. In.

— (2021). “Knowledge Spillovers and Corporate Investment in Scientific Research”. In: *American Economic Review* 111.3, pp. 871–898.

Baker, Andrew C., David F. Larcker, and Charles C.Y. Wang (2022). “How much should we trust staggered difference-in-differences estimates?” In: *Journal of Financial Economics* 144.2, pp. 370–395.

Balsmeier, Benjamin, Lee Fleming, and Gustavo Manso (2017). “Independent boards and innovation”. In: *Journal of Financial Economics* 123, pp. 536–557.

Benson, David and Rosemarie H. Ziedonis (2010). “Corporate venture capital and the returns to acquiring portfolio companies”. In: *Journal of Financial Economics* 98.3, pp. 478–499.

Bernstein, Shai (2015). “Does Going Public Affect Innovation?” In: *Journal of Finance* 70.4, pp. 1365–1403.

Chu, Jenny, Yuan He, Kai Wai Hui, and Reuven Lehavy (2020). *New Product Announcements, Innovation Disclosure, and Future Firm Performance.*

Cumming, Douglas and Na Dai (2010). “Local bias in venture capital investments”. In: *Journal of Empirical Finance* 17.3, pp. 362–380.

Dalle, Jean-Michel, Matthijs den Besten, and Carlo Menon (2017). *Using Crunchbase for economic and managerial research.*

Dass, Nishant, Vikram Nanda, and Steven Chong Xiao (2017). "Truncation bias corrections in patent data: Implications for recent research on innovation". In: *Journal of Corporate Finance* 44, pp. 353–374.

Dimmock, Stephen, Jiekun Huang, and Scott Weisbenner (2019). *Give Me Your Tired, Your Poor, Your High-Skilled Labor: H-1B Lottery Outcomes and Entrepreneurial Success.*

Edwards, Alexander and Maximilian Todtenhaupt (2020). "Capital gains taxation and funding for start-ups". In: *Journal of Financial Economics* 138.2, pp. 549–571.

Fang, Vivian W., Xuan Tian, and Sheri Tice (2014). "Does Stock Liquidity Enhance or Impede Firm Innovation?" In: *The Journal of Finance* 69.5, pp. 2085–2125.

Foley, C. Fritz and William Kerr (2013). "Ethnic Innovation and U.S. Multinational Firm Activity". In: *Management Science* 59.7, pp. 1529–1544.

González-Uribe, Juanita (2020). "Exchanges of innovation resources inside venture capital portfolios". In: *Journal of Financial Economics* 135.1, pp. 144–168.

Hadlock, Charles J. and Joshua R. Pierce (2010). "New Evidence on Measuring Financial Constraints: Moving Beyond the KZ Index". In: *Review of Financial Studies* 23.5, pp. 1909–1940.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg (2005). "Market Value and Patent Citations". In: *The RAND Journal of Economics* 36.1, pp. 16–38.

Heath, Davidson and Christopher Mace (2020). "The Strategic Effects of Trademark Protection". In: *Review of Financial Studies* 33.4, pp. 1848–1877.

Hellmann, Thomas, Paul Schure, and Dan H. Vo (2021). "Angels and venture capitalists: Substitutes or complements?" In: *Journal of Financial Economics* 141.2, pp. 454–478.

Hellmann, Thomas and Veikko Thiele (2015). "Friends or foes? The interrelationship between angel and venture capital markets". In: *Journal of Financial Economics* 115.3, pp. 639–653.

Hirukawa, Masayuki and Masako Ueda (2011). "Venture Capital and Innovation: Which is First?" In: *Pacific Economic Review* 16.4, pp. 421–465.

Hoberg, Gerard and Gordon Phillips (2010). "Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis". In: *The Review of Financial Studies* 23.10, pp. 3773–3811.

Hochberg, Yael V. and Joshua D. Rauh (2013). "Local Overweighting and Underperformance: Evidence from Limited Partner Private Equity Investments". In: *The Review of Financial Studies* 26.2, pp. 403–451.

Holmstrom, Bengt and Paul Milgrom (1991). "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design". In: *Journal of Law, Economics & Organization* 7, pp. 24–52.

Howell, Sabrina, Josh Lerner, Ramana Nanda, and Richard Townsend (2020). "Financial Distancing: How Venture Capital Follows the Economy Down and Curtails Innovation". In: *NBER Working Paper* 27150.

Islam, Emdad and Jason Zein (2020). "Inventor CEOs". In: *Journal of Financial Economics* 135.2, pp. 505–527.

Jensen, Michael C. and William H. Meckling (1976). "Theory of the firm: Managerial behavior, agency costs and ownership structure". In: *Journal of Financial Economics* 3.4, pp. 305–360.

Kaplan, S. N. and L. Zingales (1997). "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?" In: *The Quarterly Journal of Economics* 112.1, pp. 169–215.

Kaplan, Steven N. and Josh Lerner (2017). *Venture Capital Data: Opportunities and Challenges: Chap. 10 in Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*. Vol. 75. University of Chicago Press.

Kerr, William, Josh Lerner, and Antoinette Schoar (2014). "The Consequences of Entrepreneurial Finance: A Regression Discontinuity Analysis". In: *Review of Financial Studies* 27.1, pp. 20–55.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar (2019). "Who Profits from Patents? Rent-Sharing at Innovative Firms". In: *The Quarterly Journal of Economics* 134.3, pp. 1343–1404.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman (2017). "Technological Innovation, Resource Allocation, and Growth". In: *The Quarterly Journal of Economics* 132.2, pp. 665–712.

Koh, Ping-Sheng, David M. Reeb, Elvira Sojli, Wing Wah Tham, and Wendun Wang (2021). "Deleting Unreported Innovation". In: *Journal of Financial and Quantitative Analysis, forthcoming.*

Kortum, Samuel and Josh Lerner (2000). "Assessing the Contribution of Venture Capital to Innovation". In: *The RAND Journal of Economics* 31.4, p. 674.

Lee, David S., Justin McCrary, Marcelo J. Moreira, and Jack Porter (2022). "Valid t-Ratio Inference for IV". In: *American Economic Review* 112.10, pp. 3260–90.

Lerner, Josh, Antoinette Schoar, Stanislav Sokolinski, and Karen Wilson (2018). "The globalization of angel investments: Evidence across countries". In: *Journal of Financial Economics* 127.1, pp. 1–20.

Lerner, Josh and Amit Seru (2021). "The Use and Misuse of Patent Data: Issues for Finance and Beyond". In: *The Review of Financial Studies.*

Ling, Yun (2015). "The Impact of Venture Capital on the Life Cycles of Startups". In: *SSRN Electronic Journal.*

Ma, Song (2020). "The Life Cycle of Corporate Venture Capital". In: *The Review of Financial Studies* 33.1, pp. 358–394.

Ouimet, Paige and Geoffrey Tate (2020). "Learning from Coworkers: Peer Effects on Individual Investment Decisions". In: *The Journal of Finance* 75.1, pp. 133–172.

Retterath, Andre and Reiner Braun (2020). "Benchmarking Venture Capital Databases". In: *SSRN Electronic Journal.*

Sudek, Richard, Cheryl R. Mitteness, and Melissa S. Baucus (2008). “Betting on the horse or the Jockey: The impact of expertise on angel investing”. In: *Academy of Management 2008 Annual Meeting*.

The American Angel (2017). *The first in-depth report on the demographics and investing activity of individual American angel investors.*

Wallmeroth, Johannes, Peter Wirtz, and Alexander Peter Groh (2018). “Venture Capital, Angel Financing, and Crowdfunding of Entrepreneurial Ventures: A Literature Review”. In: *Foundations and Trends in Entrepreneurship* 14.1, pp. 1–129.

Whited, Toni M. and Guojun Wu (2006). “Financial Constraints Risk”. In: *Review of Financial Studies* 19.2, pp. 531–559.

Wiltbank, Robert, Stuart Read, Nicholas Dew, and Saras D. Sarasvathy (2009). “Prediction and control under uncertainty: Outcomes in angel investing”. In: *Journal of Business Venturing* 24.2, pp. 116–133.

Curriculum Vitae

Clemens Mueller

Professional Experience

2022 - 2023	PhD Researcher Chair of Corporate Governance, Prof. Dr. Alexandra Niessen-Ruenzi University of Mannheim, Mannheim, Germany
2020 - 2022	PhD Researcher Chair of Risk Theory and Insurance, Prof. Dr. Peter Albrecht University of Mannheim, Mannheim, Germany
2018 - 2020	PhD Researcher Chair of Finance, Prof. Dr. Erik Theissen University of Mannheim, Mannheim, Germany
2017 - 2017	Research Assistant, House of Finance Goethe University Frankfurt, Frankfurt, Germany
2017 - 2017	Research Assistant, Panel on Household Finance German Central Bank, Frankfurt, Germany
2015 - 2017	Teaching Assistant Chair of Corporate Finance, Prof. Ernst Maug, PhD University of Mannheim, Mannheim, Germany
2015 - 2016	Research Assistant Chair of Finance, Prof. Dr. Erik Theissen University of Mannheim, Mannheim, Germany

Education

2023 - 2023	Visiting PhD Finance HEC Paris, Paris, France
2017 - 2023	PhD Program GESS CDSB, University of Mannheim, Mannheim, Germany
2016 - 2016	Visiting MBA Guanghua School of Management, Peking University, Beijing, China
2014 - 2017	Master of Science in Management University of Mannheim, Mannheim, Germany
2012 - 2013	Visiting Student, Erasmus Program Krakow University of Economics, Kraków, Poland
2010 - 2014	Bachelor of Arts in Business Administration HTW Saar, Saarbrücken, Germany