Essays in Applied Microeconomics

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2025

vorgelegt
von

Larissa Fuchs, M.Sc.

aus

Landau in der Pfalz

Referentin:	Prof. Dr. Pia Pinger
Koreferentin:	Dr. Barbara Boelmann
Tag der Promotion:	21.02.2025

Für

Oma Hedwig & Opa Alfred

und

Charlotte & Vanessa.

Acknowledgements

As I conclude this chapter of my academic journey, I am filled with gratitude for the countless individuals who have supported, inspired, and guided me.

First and foremost, I would like to thank my PhD supervisors, Dr. Pia Pinger and Dr. Barbara Boelmann, for their support, guidance, and encouragement throughout my journey. I have learned so much from you, and I continue to be impressed by the thoughtfulness of your feedback. Thank you also for fostering a work environment defined by curiosity, trust, friendliness – it has been both inspiring and uplifting.

To my co-authors, I am grateful for the intellectually stimulating collaborations that have challenged and enriched my academic experience.

My fellow PhD students and colleagues at the University of Cologne have made this journey unforgettable, creating a supportive community I deeply value. I am particularly grateful to Anna Person, my steadfast office companion, whose gift for finding humor in challenging moments made our workspace a joy to inhabit.

I extend my sincere thanks to the firm that collaborated with us on our research. Their openness allowed us to make meaningful progress in the important area of diversity. Your collaboration has reinforced my belief in the power of academic research to effect real change.

Huyền Nguyễn, you have been more than a co-author; you've been a mentor, an inspiration, and a role model. Our chance meeting during my internship at the German Central Bank has profoundly shaped my path. Your friendship is a gift, and your brilliance as an economist continues to inspire me.

Jessica Palka, your friendship and collaboration have broadened my understanding of economics and its impact. Your extraordinary ability to see beyond disciplinary boundaries, to understand complex systems in their full context, and to draw upon a wisdom that transcends economics has both enriched our work and deeply inspired me personally.

To Silke Zeller, whose mentorship and guidance went far beyond the classroom: thank you for planting the seed of self-belief. Our friendship, enduring since 11th grade, speaks volumes about the influence of an exceptional mentor.

To my family, I am grateful for the love you have given me. I am especially thankful to my uncle Dominic, whose presence and support have enriched my life in countless ways; to my aunt Regina, whose love and care have created a bond that goes far beyond aunt and niece; to my cousin Daniel, who is like a big brother to me – supportive, protective, and a true example of loyalty; and to my sister Chittri, whose bond with me defies description.

To my friends, who have become my chosen family: Alexandra, Babs, Caro, Charlotte, Constanze, Disi, Huyền, Jess, Johannes R. Kopka, Lilian, Lukas, Susi, Talia, Tina G., Tina H., and Vanessa – your steady presence is transformative. Through your support, I've learned that vulnerability is not weakness, but strength, and that true friendship is both grounding and liberating. Your warmth and energy have made every step of this journey lighter and more meaningful. I could not have done it without you – your unwavering support, wisdom, and encouragement have been my foundation. You are my confidants, cheerleaders, and a constant source of strength, and I extend my deepest love and gratitude to each of you. Your belief in me has often exceeded my own, helping me grow in ways I had not imagined. I thank each of you for the color, depth, and love you bring to my world.

As I close this chapter and look toward the future, I carry with me the collective wisdom, support, and love of this incredible environment. For that, I am deeply grateful.

Contents

List of Tables

List of Figures

List of Abbreviations

ACPR	Autorité de Contrôle Prudentiel et de Résolution
bps	Basis Points
ECB	European Central Bank
ERI	Environmental Risk Index
ESG	Environmental, Social, and Governance
EU	European Union
GPA	Grade Point Average
GVKEY	Global Company Key
ISIN	International Securities Identification Number
NGFS	Network for Greening the Financial System
NRW	North Rhine-Westphalia
RCT	Randomized Controlled Trial
ROA	Return on Assets
sd	Standard Deviation
SIC	Standard Industrial Classification
STEM	Science, Technology, Engineering, and Math
UN	United Nations
UNEP	United Nations Environment Programme

1 Introduction

Economics offers powerful tools to address urgent societal challenges. Aligned with the UN's Sustainable Development Goals, this dissertation explores three critical domains by analyzing mechanisms across talent allocation, education, and economic practices, offering insights that can drive meaningful change.

Chapter 2 addresses selection into the labor market by examining how content in job advertisement affects talent pools and young professionals' beliefs about the work environment. By analyzing variations of certain job amenities, the research shows how subtle textual cues can influence workforce diversity, particularly in terms of geographical representation and gender composition, with direct implications for more inclusive recruitment strategies. Prior to this selection stage, Chapter 3 explores selection *post* high school by investigating how relative performance between STEM and non-STEM subjects influences educational choices and gender disparities in higher education. The research reveals how performance differences affect STEM enrollment, with significant implications for understanding gender representation in technical fields. Chapter 4 investigates how climate stress tests can transform banking practices, enhancing understanding of transition risk and promoting sustainable lending.

The dissertation's main contributions are threefold. First, given the particular challenge of skilled labor and diversity in the tech sector, this dissertation investigates the role of job advertisements as a tool for companies to attract talent. Through a randomized controlled trial (RCT) conducted at one of Europe's largest tech firms, it offers causal insights into the impact of two job attributes – flexibility and career advancement – which, until now, have only been tested in controlled laboratory settings (e.g., Wiswall and Zafar 2018). This study assesses their effect on both the quality and diversity of the applicant pool, and how they shape young professionals' beliefs about job characteristics (Chapter 2). Second, addressing the core issue of skilled labor shortage and limited diversity at its source, this dissertation also analyzes the role of relative performance in STEM and non-STEM subjects during high school, and its differential impact on gender in further decision-making and sorting into STEM fields. It quantifies the extent to which these differences in performance contribute to the gender gap in STEM enrollment in higher education (Chapter 3). Last, the dissertation presents evidence on the efforts of bank supervisors to mitigate climate change and its impacts on the banks' lending practices and their borrowers' transition to a carbon-neutral economy (Chapter 4). Through an analysis of lending decisions, it uncovers previously undocumented mechanisms through which supervisory actions related to climate change influence bank borrowers in their endeavors to transition their businesses towards carbon neutrality.

In the remainder of this section, I provide a summary of the main chapters of this dissertation and discuss their contributions to the existing literature in more detail. For co-authored chapters, I outline my contributions to each chapter.

Chapter 2: How to Attract Talents? Field-Experimental Evidence on Emphasizing Flexibility and Career Opportunities in Job Advertisements. This study investigates the causal impact of job-advertisement content on applicant pools in the labor market for STEM professionals. We implement a field experiment at a major European technology firm, randomizing the emphasis on job flexibility and career-advancement opportunities in job postings. Our findings reveal significant treatment effects for entry-level positions, with heterogeneous impacts across gender and geographical location. Specifically, advertisements highlighting job flexibility increase applications from both genders, while those emphasizing career advancement primarily attract male applicants. These effects are entirely driven by applicants residing outside of the federal state in which the firm is located. To complement our field experiment, we conduct a survey among STEM students to examine how job-advertisement content shapes beliefs about the firm's work environment and job characteristics. Our results indicate that emphasizing career advancement leads to higher anticipated career benefits but lower expected work-life balance.

This paper makes several significant contributions to the literature on labor-market dynamics and hiring practices. Our primary contribution lies in extending our understanding of how strategic framing in job advertisements influences not only the quantity of applications, but also the quality, diversity, and geographical distribution of the applicant pool. We provide causal evidence on how specific informational content in job advertisements affects the size and composition of applicant pools, focusing on commonly featured job amenities – flexibility and career advancement. This approach extends the existing literature, which has primarily focused on large-scale regulatory changes (e.g., Kuhn and Shen (2023) in China and Card, Colella and Lalive (*forthcoming*) in Austria) or interventions aimed at reducing gender imbalances (Dal Bó, Finan and Rossi 2013, Ashraf et al. 2020, Flory et al. 2021, Del Carpio and Guadalupe 2022, Del Carpio and Fujiwara 2023, and Delfino 2024). We offer a nuanced examination of how subtle differences in advertised job amenities affect hiring outcomes and applicant-pool characteristics in terms of region of residence and quality. By leveraging detailed CV data and recruiter ratings, we provide insights into the types of individuals who respond to specific job amenities. This approach underlines the interest of firms in applicant quality over quantity and complements findings on potential trade-offs in applicant selection, such as those of Del Carpio and Guadalupe (2022). Further, we present novel evidence on how highlighting job amenities in advertisements shapes potential applicants' beliefs about job characteristics and the working environment. This contribution bridges a gap in the employer-branding literature (e.g., Lievens and Slaughter 2016) by explicitly examining the mechanism through which job advertisements influence applicant perceptions and decision-making. Our findings shed light on the formation of jobseekers' beliefs about potential employers, offering valuable insights for both labor economics and management of human resources. These contributions collectively enhance our understanding of application, sorting, and hiring decisions, building upon and extending the work of scholars such as Wiswall and Zafar (2018), Coffman, Collis and Kulkarni (2024), and Vattuone (2024).

The chapter is based on the working paper by Fuchs et al. (2024*b*). It is joint work with Matthias Heinz, Pia Pinger, and Max Thon. I contributed to the project by developing ideas, collecting and analyzing data, searching for literature and writing the drafts. The other authors contributed with differing shares to the idea development, data collection, data cleaning, data analysis, literature search, and the writing of the drafts. All authors finished the final draft together.

Chapter 3: Relative Grades and Gender Differences in STEM Enrollment. This study examines how relative performance in STEM and non-STEM subjects during high school influences the gender gap in STEM enrollment, based on novel administrative and survey data from Germany. We show that while males display a higher relative STEM performance than females, this advantage primarily stems from females' stronger achievement in non-STEM subjects. Our analysis reveals gender-specific responses to relative STEM performance: a one-standard-deviation increase in grade-based STEM advantage raises males' likelihood of pursuing a STEM degree by approximately 19 percentage points, but yields only half this effect for females. A decomposition analysis demonstrates that if major preferences aligned with male patterns, 26% of the STEM gender gap could be explained by differences in grade-based STEM advantage. However, in scenarios where preferences mirror female patterns, relative grades play a minimal role. These findings suggest that STEM performance differences have limited influence on females' educational choices. Thus, while relative STEM performance significantly contributes to the observed gender gap in STEM enrollment, this relationship is predominantly driven by male behavior.

Our study contributes to the literature in four key ways. First, we confirm the importance of STEM advantage for educational choices, analyzing actual decisions rather than intentions (e.g., Breda and Napp 2019; Goulas, Griselda and Megalokonomou 2022). Second, we provide the first analysis of grade- and rank-based performance indicators' relative importance across genders in the German setting, where grades are crucial for university enrollment. Third, there is evidence that students have imperfect knowledge of their own ability (Zafar 2011; Stinebrickner and Stinebrickner 2012, 2014; Bobba and Frisancho 2016) and are uncertain about the returns to education (Jensen 2010; Attanasio and Kaufmann 2014; Wiswall and Zafar 2015). In our analysis, we show that female students seemingly place too little weight on their relative advantage when making education decisions. Our decomposition exercise allows us to differentiate the effects of gender preferences from performance differences. This enables us to quantify how observed performance differences in STEM and non-STEM fields contribute to the overall gender STEM gap, in a context where females may face barriers. Building on prior literature (e.g., Delaney and Devereux 2019; Card and Payne 2021; Riegle-Crumb et al. 2012), we provide an explanation for the paradox of women selecting lower-wage non-STEM fields despite equal or superior academic performance across disciplines. We extend research on how ability cues influence decision-making (e.g., Stinebrickner and Stinebrickner 2012; Murphy and Weinhardt 2020; Elsner, Isphording and Zölitz 2021; Bond et al. 2018; Li and Xia 2024; Tan 2023) and gender differences in grade responsiveness. Prior work shows females' persistence in subjects correlates with strong performance (Owen 2010), yet they exit male-dominated and STEM fields more readily after poor performance than males (Kugler, Tinsley and Ukhaneva 2021; Rask and Tiefenthaler 2008). While existing studies focus on absolute grades, we demonstrate that females respond less to relative performance differences across subjects. This diminished sensitivity of females to comparative advantages may require stronger signals to encourage female STEM participation.

The chapter is joint work with Pia Pinger and Philipp Seegers. I contributed to the project by developing ideas, collecting and analyzing data, searching for literature and writing the drafts. The other authors contributed with differing shares to the idea development, data collection, data cleaning, data analysis, literature search, and the writing of the drafts. All authors finished the final draft together.

Chapter 4: Climate Stress tests, Bank Lending, and the Transition to the Carbon-Neutral Economy. This study investigates the impact of climate-related supervisory activities, specifically climate stress tests, on the lending behavior of banks and their transition to the carbon-neutral economy. By using data from the French prudential regulatory agency's climate pilot exercise and combining it with information on the borrowers' carbon emissions, we compare outcomes between banks participating in the exercise and those that do not. The findings reveal that participating banks increase lending to borrowers with higher transition risk while raising loan rates, in contrast to non-participating banks that reduce credit supply. Participating banks also demonstrate an increased collection of new information about climate risks and boost lending for green purposes. Borrowers of participating banks show improved environmental performance, including a higher likelihood of implementing emission policies, setting carbon targets, and using renewable energy. However, no significant reductions in direct carbon emissions or changes in relationships with environmentally unfriendly suppliers were observed. The study concludes that climate stress tests serve a dual purpose: They not only identify vulnerable spots in the financial system, but also reduce information asymmetries between banks and borrowers, ultimately supporting the transition to a net-zero economy and encouraging more climate-resilient business practices.

Our research contributes to multiple strands of literature. First, while numerous studies examine how supervisory resources, standards, intensity, and enforcement actions affect banks and their borrowers (e.g., Eisenbach, Lucca and Townsend 2016; Hirtle, Kovner and Plosser 2020; Goldsmith-Pinkham, Hirtle and Lucca 2016; Ivanov, Kruttli and Watugala 2024; Kiser, Prager and Scott 2012; Bassett, Lee and Spiller 2015; Agarwal et al. 2014; Rezende and Wu 2014; Delis and Staikouras 2011; Danisewicz et al. 2018), we add to this by showing how supervisory efforts to address climate change generate new information, enabling banks better to assess the borrowers' transition risks and to adjust their lending decisions accordingly. Second, we extend the literature on stress tests. While Morgan, Peristiani and Savino (2014) and Flannery, Hirtle and Kovner (2017) highlight the valuable information generated by stress tests, and others show their effects on lending behavior (e.g., Acharya, Berger and Roman 2018; Cortés et al. 2020;

Gropp et al. 2019; Kok et al. 2023), our study uniquely establishes a direct connection between climate stress tests and borrowers' actions to make their business models more resilient without triggering capital surcharges. This addresses the gap noted by Acharya et al. (2023) in the need for more research on climate stress tests. Third, we advance understanding of how the lending behavior of banks responds to climate change. While existing research documents reductions in credit supply and increased securitization following climate-risk signals (e.g., Chava 2014; Delis et al. 2024; Anginer et al. 2023; Mueller and Sfrappini 2022; Mueller, Nguyen and Nguyen 2022; Kacperczyk and Peydró 2022; Bruno and Lombini 2023; Nguyen et al. 2022; Correa et al. 2022; Meisenzahl 2023), our findings reveal that banks participating in the climate pilot exercise actually increase lending. This supports the view that climate stress tests help banks better to understand and manage climate risks, influencing their business strategies and lending practices. Finally, we contribute to the limited literature on financial constraints and decarbonization. Unlike Accetturo et al. (2022), who identify credit availability as a barrier to green investments, our results show that the participation of banks in climate stress tests increases credit availability, underscoring the real effects of supervisory efforts on climate-related lending and business-model adjustments.

The chapter is based on the working paper Fuchs et al. (2024*a*). It is joint work with Huyền Nguyễn, Trang Nguyễn, and Klaus Schaeck. I contributed to the project by developing ideas, collecting and analyzing data, searching for literature and writing the drafts. The other authors contributed with differing shares to the idea development, data collection, data cleaning, data analysis, literature search, and the writing of the drafts. All authors finished the final draft together.

2 How to Attract Talents? Field-Experimental Evidence on Emphasizing Flexibility and Career Opportunities in Job Advertisements

joint with Matthias Heinz, Pia Pinger and Max Thon*

Abstract

Job advertisements are a key instrument for companies to attract talent. We conduct a field experiment in which we randomize the content of job advertisements for STEM jobs in one of the largest European technology firms. Specifically, we study how highlighting job flexibility and career advancement in job advertisements causally affects the firm's pool of applicants. We find large treatment effects of entry-, but not for senior-level positions in the firm: highlighting job flexibility increases the total number of female and male applicants, while emphasizing career advancement only raises applications by men. Both effects are entirely driven by applicants residing outside of the federal state in which the firm is located. In a survey experiment among STEM students,

*We thank Anna Bindler, Katherine Coffman, Alexia Delfino, Oliver Gürtler, Paul Heidhues, Mitchell Hoffman, Boon Han Koh, Astrid Kunze, Rafael Lalive, Johannes Münster, Eva Oeß, Clémentine Van Effenterre, Johannes Wohlfahrt, Lise Vesterlund, and seminar participants at the Bocconi Gender Lab, Bern University, Paderborn University as well as participants at the WEE2024 conference in Copenhagen for helpful comments. We further thank our student assistants Clara Barrocu, Milena Hüffel, Leonard Loechelt, Rosanna Simonis, Lea Striek, and Robert Szwed for their great research assistance. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1 – 390838866. Financial support through the German Research Foundation Collaborative Research Centre TR 224 is gratefully acknowledged. The study is approved by the Ethics Committee of the University of Cologne (Reference: 220022MT). The experiment was pre-registered at the Registration portal of the American Economic Association with the RCT IDs AEARCTR-0010433.

we find that the content of job advertisements shapes young professionals' beliefs about the work environment at the firm. In particular, highlighting career advancement leads to a shift in beliefs towards better career benefits, but lower work-life balance.

Keywords: Beliefs, Hiring, Field Experiments, Survey Experiment, Job Advertisements, Gender
JEL classification: M51, M52, D22

2.1 Introduction

How do workers choose which jobs to apply to? Earnings are an important factor in this decision, but workers typically also consider many other job characteristics. These include the job's location, job flexibility, development opportunities, and co-worker characteristics. The decision to apply will then depend on (i) a worker's preferences for these workplace characteristics and (ii) her beliefs about whether a job comes with certain amenities. Preferences for job characteristics are fixed in the short term, but may vary greatly across individuals (Ashraf et al. 2020), in particular between women and men (Flory, Leibbrandt and List 2015; Wiswall and Zafar 2018). Some individuals may, for example, prefer to work in a particularly dynamic or challenging environment while others value flexibility. Beliefs about job characteristics, on the other hand, depend on the type of information companies provide, e.g., by highlighting certain characteristics in their job advertisements.[1] In job advertisements, firms not only inform about the existence of a vacancy, but also send signals about the job's characteristics and the working environment at the firm (Del Carpio and Guadalupe 2022; Delfino 2024; Card, Colella and Lalive *forthcoming*). This may help firms to attract talented workers, a key strategic and scarce resource in today's knowledge-driven economy. Although attracting the best talents is crucial for long-term success (Coff 1997; Bapna et al. 2013; Del Carpio and Guadalupe 2022), many firms report skilled labor shortages.[2] Moreover, if highlighting certain job characteristics leads to a better fit between worker's preferences and job characteristics, this can also improve the matching process and long-run worker satisfaction. The type of information that is highlighted in job advertisements is thus of key importance for firms, workers, and the worker-firm match alike.

We study how job characteristics highlighted in job ads affect the size and composition (i.e., quality, gender, region of residence) of the pool of applicants, and job seekers' beliefs. We run a RCT within one unit of one of Europe's largest technology firms, which employs around 3,000 workers. In our RCT, we randomized job characteristics for all STEM vacancies posted by the firm for a period of 12 months. Specifically, we posted the same job ad three times, with a sequence of treatments randomized at 10-day intervals: In one instance, we emphasized the high level of job flexibility at the firm (*flexibility* treatment); in another instance, we highlighted career advancement, in form of good personal and wage growth opportunities (*career* treatment), and in one instance without highlighting either characteristic (*control* treatment).

[1] In 2018, job boards accounted for half of all job applications and ultimately contributed to 30 percent of successful hires (Jobvite 2019*a*,*b*).

[2] See, for instance, Marjenko, Müller and Sauer (2021) or ManpowerGroup (2024).

We focus on flexibility and career advancement, as these attributes significantly influence job attractiveness (Mas and Pallais 2017; Wiswall and Zafar 2018) and emerged as distinctive features of the unit's jobs through pre-RCT interviews with firm managers, workers, and worker representatives.

Our empirical investigation is grounded in a conceptual framework that informs the empirical analysis and elucidates the expected effects and their underlying mechanisms. In the framework potential applicants receive utility from job-specific ability and job characteristics, such as job flexibility and career opportunities. Highlighting specific job characteristics (the treatment) is then interpreted as a signal that shifts beliefs about the attractiveness of the job (the mechanism) inducing a change in the likelihood to apply (the outcome). Based on this framework, we derive a number of empirical predictions that can be tested using our data. First, both treatments should increase the total number of applications, but the effect should be larger for entry-level jobs than for senior-level jobs. Second, the *flexibility* (*career*) treatment should increase the number of female (male) applicants relatively more than that of male (female) applicants. Third, both treatments should cause a positive belief shift about the expected level of job flexibility and career opportunities. Last, if job preferences correlate with worker productivity or background characteristics (Nekoei 2023), we also expect changes to the quality and composition of the applicant pool, which we study in an exploratory manner.

Our main findings can be summarized as follows. First, both the *flexibility* and the *career* treatment weakly increase the number of applications. Among inexperienced workers the effect is sizeable, amounting to an increase in applications of 44 percent for the *flexibility* and of 35 percent for the *career* treatment, respectively. Moreover, the *flexibility* treatment is relatively more attractive to women compared to the *career* treatment, while no significant differences are observed for men between the two treatments. To investigate the effect on the quality of the additional applicants, we rely on CV information and ratings of the firm's recruiters. We find suggestive evidence that the *career* treatment induces a more positive selection compared to the *flexibility* treatment. New applicants mainly come from Germany, but not from an area close to the firms' location, suggesting that the treatments allow the firm to source talent from a wider regional labor market.

To assess potential mechanisms that relate to the role of preferences and belief-updating, we complement our field experiment with an online survey experiment among more than 2,000 STEM students, using the subject pools from 12 German and Austrian economic research laboratories. The online surveys were conducted in parallel to the field experiment and showed individuals a job advertisement that fit their experience level and that was posted by the firm (almost) at the same point in time. We randomized the treatments between subjects and invited participants such that their educational background matched the requirements of the particular job advertisement. We find that both treatments shift beliefs in the experienced direction. In terms of a composite flexibility score capturing work-life balance, the *flexibility* treatment increases expectations regarding flexible working conditions by 0.132 standard deviations. The *career* treatment increases expectations regarding career benefits (in terms of wage and career progression), again measured by a composite score, by 0.162 standard deviations. Moreover, we find evidence of belief trade-offs between workplace characteristics. While the *career* treatment increased beliefs about career advancement, it simultaneously *lowered* expectations about workplace flexibility by 0.094 standard deviations. Regarding, preferences, we confirm previous evidence that females exhibit a higher preference for flexibility than males, while finding weak evidence of slightly higher preferences for career growth among males compared to females. No differences in preferences were observed, however, with respect to age.

The contribution of this paper is threefold. First, our findings demonstrate that experimentally-induced highlighting of content provided in job advertisements can affect the size and composition of the applicant pool. This evidence complements a literature exploiting large-scale regulatory changes to show that a removal of gender preferences in job ads has led to an increase in applications from the previously non-preferred gender in China (Kuhn and Shen 2023) and to more gender-neutral hiring outcomes in Austria (Card, Colella and Lalive *forthcoming*). It also relates to several interventions aiming to reduce gender imbalances especially in training programs or public-sector jobs, by avoiding stereotypical language, signaling interest in employee diversity, or by highlighting past employee performance (Dal Bó, Finan and Rossi 2013; Ashraf et al. 2020; Flory et al. 2021; Del Carpio and Guadalupe 2022; Del Carpio and Fujiwara 2023; Delfino 2024). In terms of evidence, our paper differs from these studies in that we focus on highlighting flexibility and career advancement – job amenities that are commonly part of job advertisements. In that respect, our second contribution becomes important, namely that we can investigate how subtle differences in job amenities can affect hiring outcomes and the composition of the applicant pool in terms of region of residence and quality. We do this by utilizing detailed CV data and recruiter ratings. After all, a firm's primary interest may not be in the number of applications overall, but in average or top applicant quality. Besides, this provides evidence on which types of individuals respond to a certain type of job amenity offered, thus revealing information about underlying preferences. This relates to the evidence provided in Del Carpio and Guadalupe (2022), who has shown that a treatment reducing gender stereotypes adversely affects selection. Last, we provide first evidence of how information about highlighting job amenities in advertisements affects the beliefs of potential applicants regarding both expected job characteristics and the working environment. Such changes in beliefs, albeit not explicitly, are the focus in the employer-branding literature (Lievens and Slaughter 2016). As regards all three contributions, our paper also relates to studies investigating application, sorting, and hiring decisions more generally, in particular as regards preferences of both employers and employees. Research shows that preferences differ across different types of employees, most prominently men and women (Wiswall and Zafar 2018; Ashraf et al. 2020; Coffman, Collis and Kulkarni 2024; Vattuone 2024). Firms also differ in their preferences for certain candidates, as becomes evident when companies react to signals and subtle cues on CVs when selecting candidates (Heinz and Schumacher 2017; Hoffman, Kahn and Li 2018). If firms knew about the preferences of their preferred "types" of workers they could make strategic use of that knowledge and provide – as well as highlight – those amenities. If successful, such firm strategies could improve the matching process, increase firm productivity, and lead to long-term stable employment.

The structure of the paper is as follows. In Section 2.2, we present the background of our study by providing a description of our study firm and the motivation for our treatments. In Section 2.3, we present the conceptual framework guiding our empirical analysis. Section 2.4 presents the experimental design. Section 2.5 presents the main empirical results of the field experiment. Section 2.6 discusses potential mechanisms focusing on the results of our complementary survey experiment. In Section 2.7, we present a series of robustness checks to validate our findings. Lastly, Section 2.8 concludes.

2.2 Background and Motivation

2.2.1 The Firm

We conducted a field experiment in collaboration with one of Europe's largest technology firms. This leading company operates in the semiconductor market and generated a total revenue of roughly 16 billion euro in the business year 2023 with a total workforce of roughly 60,000 workers. The semiconductor industry experienced strong growth in demand in the past and is expected to grow further according to industry experts (see, e.g., Burkacki, Lehmann and Dragon 2022). For our project, we collaborate with one plant of the company situated in Germany. This particular plant experienced strong growth in the last years as well. From 2012 to 2022, the workforce at the plant increased by roughly 50 percent, from approximately 2,000 employees to 3,000. The majority of employees have a STEM background, specifically in fields such as engineering, manufacturing, construction, computer science, mathematics, and physics.

The proportion of female STEM workers in the company is roughly equivalent to that of female STEM graduates in Germany.[3] In leadership positions, 5-10% of the employees are women. The personnel turnover rate among workers is relatively low.[4] Due to the strong growth, the firm is constantly hiring.

2.2.2 Motivation of our Intervention

An essential step of the cooperation with the firm was to gain a comprehensive understanding of the firm's recruiting strategy, its main challenges, and its strategic goals. To do so, we engaged in discussions with key stakeholders, including top managers from the HR department, the head of diversity, recently hired employees as well as those hired a long time ago (especially women), the head of the workers' council, and management executives. We learned that the firm faces two challenges. First, the overall number of applications is low. On average, for each advertised position, the company receives only 12 applications. Second, the share of female applicants is also low. On average, only 12.8% of applications are from female applicants. This is problematic, as the firm's publicly announced goal is to increase the share of female workers from middle-management onwards to 20%.[5]

The main objective of the cooperation was to find ways to overcome both challenges and, in particular, to increase the total number of applications. As job advertisements are still among the most important instruments to attract applicants, changes to them are nearly costless, and current research provides evidence about the important role their content plays for application decisions (see, e.g., Del Carpio and Guadalupe 2022; Delfino 2024), we quickly consented that we want to investigate how changes in job advertisements may help to attract more applicants.

[3] As reported by the OECD, in 2021 the share of female graduates in the field of STEM for a bachelor's degree or equivalent level amounts to 16%, and it is 22% for a master's degree or equivalent level.

[4] We have no data on personnel turnover, but HR officials told us that it is around 1%.

[5] Before our intervention, the firm already had a number of initiatives in place to increase the total number of applications, in particular from women. They engage in cooperation with local universities, went to regional job fairs and fairs at big universities, and increased active talent-sourcing. However, the recruiting challenges remained.

We conducted a number of in-depth interviews about the recruiting processes and challenges carried out among different groups of workers within the firm. During these interviews, when asked about the distinctive characteristics of jobs within the plant, almost all workers mentioned two aspects. First, the plant offers a lot of flexibility to maintain work-life balance. In particular, the plant offers workers the opportunity to work full-time or part-time, and jobs that are shared by two employees are fairly common. The local municipality offers a sufficient number of day-care spots with moderate care fees.[6] Employees generally state that the culture of the plant is 'family-friendly'; for example, workers argue that it is 'socially accepted' in the firm to leave early when kids are sick or to work only at certain times. The HR department also argued that it is common to design individual solutions for new employees with care-giving responsibilities.[7] Second, because of the growth in the sector overall, wages grew substantially in the past. With expected future growth, it is likely that wages and career opportunities (e.g., there are constantly new leadership positions available) will keep growing. Indeed, firm growth and wage growth within firms are highly correlated (Fox 2009; Brown and Medoff 1989; Groshen 1991; Idson and Oi 1999).

From standard economic theory, many individuals take career advancement into account when deciding to apply for a job. Prior research has shown that the degree of job flexibility significantly influences job choice, particularly for women (Wiswall and Zafar 2018; Mas and Pallais 2020).

Given the importance of career advancement and flexibility for application decisions in general and their overwhelming presence at the firm, we agreed to test the effect of highlighting these job characteristics in the company's job advertisements.[8] Before we started with the research project, we presented the project outline to the work council of the firm, who provided their agreement and support.

[6]In Germany, the demand for day-care spots for young children is much higher than the supply; the estimated gap for children aged one and younger is 24 percent (Alt et al. 2017). Thus, daycare is a major challenge for many young professionals.

[7]Job security at the plant is fairly high. However, this is not a unique selling proposition; rather, it reflects the broader German labor market environment, where strong employment protection laws and works councils make terminations difficult, particularly in large companies.

[8]Before our intervention, the firm did not highlight (e.g., in job ads or on the homepage) the large opportunities for flexibility and career advancement, but only mentioned it in very small text at the bottom of the page. When we asked the HR department in our study firm why flexibility and career advancement were not highlighted in the job ads, they told us that the reason for this are the multinational firm's standard centralized HR policies and standardized IT-systems.

2.2.3 Details of the Hiring Process

The hiring process consists of three steps and is managed by one person from the HR department, the 'talent attraction manager', who mainly takes care of the administrative process, as well as a 'hiring manager', who is usually the head of the department for which the position is advertised. Both the screening process and final hiring decisions are made jointly by the hiring manager and talent attraction manager.

Step one is an initial screening and evaluation by the hiring manager and the talent attraction manager. This evaluation is either an A, B, C, or 'No rating'. An A rating is given to candidates who are highly promising and meet the outlined criteria of the ideal candidate by 70-100 percent. A B rating is assigned to candidates who meet the criteria by 50-70 percent. A C rating is for applicants who lack most of the required qualifications or possess characteristics that make them unsuitable for the position, with a fit of less than 50 percent. The 'No rating' category typically includes candidates who are screened out early in the hiring process because they do not meet the minimum requirements for the position. Step two of the process consists of an interview, usually conducted with the hiring manager and the talent attraction manager. In the third and final step, they decide after the interview whether to extend a job offer. If both approve, salary and contract negotiations begin with the candidate. Upon successful agreement, the candidate is hired.

2.3 Conceptual Framework

In this section, we discuss a conceptual framework that illustrates how highlighting job flexibility or career advancement in job ads affects belief-updating and the expected job utility of potential applicants. The idea is to provide an intuition for how a change in job ads might affect workers' application behavior through a change in expected utility from job flexibility (*flexibility* treatment) and career advancement (*career* treatment).[9] The goal is to derive empirical predictions about the size and characteristics of the applicant pool, treatment effect heterogeneities, and changes in worker beliefs, which guide our empirical analysis. The framework is formalized in Section 2.9.1 of Appendix A. In the following, we describe its main implications and related predictions.

[9] For a related framework based on a similar idea, see Delfino (2024).

Consider the following framework, which reflects upon relevant characteristics for an application decision. There are two types of individuals, either with or without previous work experience. Each individual considers applying to a job advertised by one firm (i.e., our study firm). An applicant applies to the job if the expected utility derived from the job is larger than the (fixed) utility from an outside/alternative offer. Potential applicants derive utility from the (fixed) wage payment, the individual returns to ability, the expected level of flexibility, and career-advancement opportunities provided by the firm. Individuals are uncertain about job flexibility and career advancement, but hold a belief about both. Additionally, we allow for beliefs about these two job characteristics to be correlated. This implies that some applicants may believe that these two characteristics are not related (i.e., no trade-off), while some others might think that career growth comes at the cost of flexibility (i.e., a negative trade-off) or that career growth is not possible without flexibility (i.e., a positive trade-off).[10]

Moreover, we assume that the distributions of prior beliefs differ between experienced and inexperienced applicants. Longer experience in the labor market likely leads to better networks, and thus greater knowledge of the industry and its firms.[11] In our framework, this translates into the assumption that experienced applicants hold a more precise and weakly more positive belief about the exact level of provided flexibility and career-advancement opportunities.[12] Indeed, the true level of flexibility and career-advancement opportunities provided by the firm is assumed to be higher than experienced and inexperienced applicants expect.

We interpret the different treatments, namely the highlighting of *flexibility* and *career* in the job ads, as a way for the company to signal flexibility and career-advancement opportunities. These informational treatments induce applicants to update their beliefs.

The firm's signaling of flexibility and career-growth opportunities leads to positive belief-updating among potential applicants. More positive beliefs in turn lead to a higher expected job utility among applicants and to an increase in the likelihood to apply for the job. For this to hold, it is merely necessary that applicants' beliefs about the trade-off between flexibility and career-advancement opportunities are not too negative.

[10]Arguably, there are other job characteristics that might matter and enter the utility function. As these are not part of our study, we abstract from those.

[11]The economic literature notes, for instance, that more experienced workers receive information through better co-worker networks (Glitz, 2017).

[12]All results derived from the model still hold even if the prior belief of experienced workers is slightly more negative than that of inexperienced workers, as long as it is not too far away and the prior of the experienced workers is sufficiently more precise.

Next, we discuss possible effect heterogeneities. As experienced applicants hold more precise and positive beliefs about the provided level of flexibility and career-advancement opportunities at the firms, their expected utility gain is smaller relative to that of inexperienced applicants. As utility gains lead to more applications, we expect that both treatments lead to a relatively larger increase in applications among inexperienced candidates compared to experienced candidates.

Additionally, it is conceivable that the preferences for flexibility and career differ between female and male applicants. In fact, Wiswall and Zafar (2018) find that females have a relatively higher willingness to pay for jobs with more flexibility than males and that males have a relatively higher willingness to pay for jobs with a higher potential for career-advancement opportunities than females. In line with these findings, we assume that women have a stronger relative preference for flexibility and males have a stronger relative preference for career advancement. This translates to larger expected utility gains for women when they see a job ad highlighting flexibility, and larger gains for men when they see a job ad highlighting career-advancement opportunities. Subsequently, the increase in the number of applications should be larger for female (male) applicants for the *flexibility* (*career*) treatment.

All of the discussed effects rely on belief-updating of potential applicants upon observing the treatments. Thus, a necessary requirement is that the *flexibility* treatment leads to a positive shift in beliefs about the provided workplace flexibility, while the *career* treatment induces a positive shift in beliefs about career advancement.

The above discussions yield several empirical predictions: We predict that 1) both treatments increase the number of applications, 2) that the increase is larger for entry-level than for senior-level positions, 3) that the *flexibility* (*career*) treatment leads to a stronger increase in applications for women (men) than men (women), and 4) that both treatments lead to positive shift in beliefs about the provided job flexibility and career-advancement opportunities.

Note that empirical predictions 2) and 3) both address heterogeneities in treatment effects, but that the mechanism underlying the heterogeneity is conceptually different. For empirical prediction 2) differences between experienced and inexperienced workers arise due to differences in belief-updating, while for empirical prediction 3) we expect effect heterogeneities, due to gender differences in preferences. As 5), we thus predict differences in belief updating between inexperienced and experienced workers (but no difference in preferences) and differences in preferences between males and females (but no differences in belief-updating), as differential underlying mechanisms of potential effect heterogeneities.

Last, the framework does not yield clear predictions about the expected change in applicant quality, which depends on the correlation between preferences for workplace flexibility, career-growth opportunities and job-specific ability. We will investigate this in an exploratory manner.

In the next section, we discuss the experimental design in more detail.

2.4 Experimental Design

2.4.1 Job Ads

The job advertisements have a uniform structure and are presented on the homepage of the company as well as on different job boards, the main ones being Indeed, LinkedIn, and one local job board.[13] Most of the applications, however, are received by the company via their own homepage. The purpose of the job advertisements is to inform potential applicants about the vacancy and to convince potential and ideally suitable applicants to apply.

Figure 2.1 shows a a fictitious sample of a job ad of the study firm. The content is generated via OpenAI (2024) based on input of real job ads of the study firm. The wording, font, and color are manually changed to guarantee the firm's anonymity. At the top, the company presents varying pictures of employees at work. The job titles are usually very short consisting of a maximum of three terms. Below the job title, the ads provide a so-called 'teaser text'. This text provides a superficial description of the advertised job. The *Job description* section provides a summary of the job and outlines the specific tasks in bullet points. The *Your profile* section summarizes the requirements the applicant should ideally fulfill. The *At a glance* section lists the general conditions of the job (e.g., location, the desired start date, contract type). The benefits section shows the provided employee benefits of the study firm. Each icon symbolizes one particular benefit.[14] The benefits for the job, in order of the symbols, are as follows: 1) Coaching & mentoring, 2) A wide range of training opportunities and career development, 3) International assignments, 4) Various career paths, 5) Flexible working conditions, 6) Option for part-time work, 7) Paid holiday, 8) Childcare support, 9) Social counseling

[13]Mentioning the name of this job board would threaten the anonymity of the study firm.
[14]A text description naming the benefit appears when the mouse cursor hovers over the icon.

& work doctor, 10) Health promotion programs, 11) On-site canteen, 12) Home-office options, 13) Corporate pension benefits, 14) Performance bonus, and 15) Accessibility. The contact opportunities section shows the name and e-mail address of the responsible talent attraction manager who can be contacted in case of further questions.

Figure 2.1: Sample Job Ad

Note: This figure presents a fictitious sample of a job ad from the study firm. While manually created, the content is generated using OpenAI (2024) based on real job ads. Wording, font, and color are altered to maintain anonymity. At the top, the company features images of employees at work, followed by concise job titles. Below the title is a 'teaser text' offering a brief job overview, followed by a detailed job description in bullet points. The 'Your profile' section outlines the ideal candidate's requirements, while the 'At a glance' section lists general job conditions. The benefits section, represented by icons, highlights employee offerings. Hovering over each icon reveals a description. The benefits, in order, are: 1) Coaching & mentoring, 2) Training opportunities and career development, 3) International assignments, 4) Career paths, 5) Flexible working conditions, 6) Part-time work options, 7) Paid holiday, 8) Childcare support, 9) Social counseling & work doctor, 10) Health programs, 11) On-site canteen, 12) Home-office options, 13) Pension benefits, 14) Performance bonus, and 15) Accessibility.

Our treatments consist of two particular statements, one of which (or none) is randomly shown just below the 'teaser text'. The exact treatment texts are presented in Figure 2.2. Figure 2.2a shows the *flexibility* treatment, which reads: "FLEXIBILITY is very important to us! Together we look for individual solutions, so that your job does not get in the way of your personal life". It thus highlights the opportunity of flexibility in a very general way. This is necessary, as the potential and detailed conditions for flexibility vary with the particular job. The aim of the treatment is to signal that the firm guarantees to provide an above-average level of flexibility conditional on the requirements of the job. Figure 2.2b shows the *career* treatment, which reads: "GROWTH is very important to us! With us, you do not only grow personally, but also your salary". It signals that the firm provides a job with wage growth and career opportunities, as well as opportunities for personal growth. Similar to the *flexibility* treatment, the conditions for career advancement vary by job, as opportunities for career progression and pay raises depend on the specific tasks and department. The goal of the treatment is to signal that the firm is committed to providing above-average career-advancement opportunities.

(a) *Flexibility* treatment

Ready to lead the future of power semiconductor innovation? As a Product Development Engineer, you'll transform groundbreaking ideas into high-volume production realities. Join our team and elevate your career by shaping the next generation of advanced technology. FLEXIBILITY is very important to us! Together we look for individual solutions, so that your job does not get in the way of your personal life.

(b) *Career* treatment

Ready to lead the future of power semiconductor innovation? As a Product Development Engineer, you'll transform groundbreaking ideas into high-volume production realities. Join our team and elevate your career by shaping the next generation of advanced technology. GROWTH is very important to us! With us, you do not only grow personally, but also your salary.

Figure 2.2: Treatments

2.4.2 Randomization

In the past, the majority of job applications were received within the first 30 days of the job being online. Due to the limited number and considerable heterogeneity of jobs advertised by the firm within one year, we randomize the treatment within job ads, each over a period of 10 days. Thus, our randomization procedure is as follows: Once a department reports a vacant position to the HR department and the job posting is approved, a random draw determines the treatment – either the *control*, *flexibility*, or *career* teaser text. The job ad is then posted in this version for 10 days. After 10 days, a random draw decides which of the two remaining treatments is posted. This means that from day 11 to 20 the same job ad is posted with a teaser text corresponding to one of the two remaining treatments. Finally, after 20 days, so from day 21 to 30, the same job ad is posted with a teaser text corresponding to the remaining treatment.[15] Each job ad is thus posted sequentially under each treatment condition.

The randomization was conducted by an external intermediary person, who was hired as an external employee by the company. We provided the randomization schedule to this person. As a "firewall" measure, this person was not involved in any other tasks of the HR department, nor in any of the research. Recruiters were not informed about the chosen treatments for the different time periods of the jobs.[16] The field experiment took place between October 2022 and October 2023. It only included job ads requiring a STEM background. Throughout our treatment period, we randomized a total of 105 job ads.

[15]Some job ads are posted longer than 30 days until the position is filled. As outlined in our pre-registration, we do not include applicant data collected after the 30-day period.

[16]As a safeguard for the field experiment, one of our research assistants checked every day that the 'correct' job ad was posted online on each platform. The research assistant documented the treatments every day, without being informed about the scheduled treatment. The research assistant detected three inconsistencies in terms of a missing treatment switch when scheduled. This explains the slight imbalance of three daily observations in Table 2.1 presenting the summary statistics.

2.4.3 Data

Our main analysis draws on firm data about a total number of 1,084 applications, applicant characteristics, and applicant ratings. The sample comprises all applicants who applied to job ads in our experiment between October 2022 and October 2023, with a maximum of two applications each.[17] The data comprise the date of application, the applicants' gender, their place of residence (if available), as well as their performance in the hiring process (i.e., recruiter ratings, interview invitation, and hiring outcome). Besides, we draw on anonymized data from the applicants' CVs.[18]

Table 2.1 summarizes our data. It provides information on the *daily* number of applications by gender, by quality (in terms of recruiter ratings and interview invitation), and by region of residence. To assess whether the treatment led to applications from a wider pool, we categorize applicants as either living in Germany, but not in the federal state of the firm (Germany w/o state), living in the federal state where the firm is located (State), and applicants from abroad (Abroad).

Our main outcome variable is the total number of daily applications per job advertisement, overall and by gender. A focus of our analysis is the investigation of heterogeneous effects across entry-level and senior-level jobs.

[17]We exclude applicants who submitted more than two applications, representing 4.8% of our sample. These are classified as mass applicants by the firm. It is plausible to assume that these application decisions are not driven by our treatments. Some applicants even sent up to 20 applications during our experimental time period.

[18]As part of the field experiment, we collect sensitive and personal data from applicants. To align with data-protection standards, we implemented several processes aimed at GDPR compliance. Central to our approach is the establishment of an anonymous intermediary person, who is hired as an external employee of the firm and prepares the data in a sufficiently anonymized way so that we can use it for our analyses. The most important guideline overall was to ensure that we never handle personal data that could lead to individual identification.

Table 2.1: Summary Statistics: Daily Application Data

Variables (daily)	Control Mean	Control SD	Flexibility Mean	Flexibility SD	Career Mean	Career SD
A. Applications by gender						
Total	0.374	0.906	0.422	1.750	0.374	0.824
Male	0.301	0.745	0.336	1.349	0.302	0.666
Female	0.074	0.305	0.087	0.491	0.071	0.323
B. Applications by quality						
A rating	0.038	0.206	0.027	0.161	0.045	0.229
B rating	0.074	0.297	0.057	0.248	0.075	0.294
C rating	0.087	0.320	0.088	0.328	0.084	0.322
Screened out	0.175	0.692	0.250	1.683	0.170	0.571
Invited for interview	0.075	0.288	0.059	0.244	0.082	0.294
C. Applications by region of residence						
Germany w/o state	0.134	0.496	0.185	0.951	0.154	0.446
State	0.117	0.348	0.106	0.433	0.104	0.365
Abroad	0.109	0.381	0.114	0.486	0.100	0.344
Observations	1,047		1,051		1,052	

Note: This table shows the mean and standard deviations of daily applications received by gender, quality, and region of residence. 'Control' refers to the control treatment, 'Flexibility' refers to the *flexibility* treatment, and 'Career' refers to the *career* treatment.

2.5 Main Empirical Analysis – Field Experiment

In this section, we present our estimation strategy and the main results. Our aim is to investigate how both treatments affect the total number of applications. We split the analysis between job ads for entry-level and senior-level positions. Furthermore, we analyze how the treatments affect the composition of the applicant pool. To do so, we rely primarily on recruiter ratings to assess applicant quality and on CV data to investigate changes in the applicants' region of residence.

2.5.1 Descriptive Results

We first provide descriptive evidence on the relationship between the presented job ad and the number of applications received per day. Figure 2.3 shows the average number of daily applications for entry-level positions in total (2.3a), by gender (2.3c, 2.3e) as well as the average number of daily applications for senior-level positions in total (2.3b) and by gender (2.3d, 2.3f). We observe that both treatments boost the number of applications for entry-level positions in Figure 2.3a. The effects are sizeable, amounting to 0.119 additional applications per day for the *flexibility* treatment and to 0.0973 additional applications per day for the *career* treatment. Figures 2.3c and 2.3e present treatment effects separately by gender. We find that both treatments increase the number of male applicants to entry-level positions by roughly equal amounts, namely by 0.0765 applications per day in response to the *flexibility* treatment, and by 0.0997 applications per day in response to *career* treatment. Among female applicants (Figure 2.3e), only the *flexibility* treatment leads to an increase in applications (by 0.0424 applications per day). The *career* treatment leads to a slight, but insignificant, decrease of -0.0025 applications per day (Figure 2.3c). Considering Figures 2.3b, 2.3d and 2.3f, we observe almost no changes for the *career* treatment and slight, but insignificant, increases for the *flexibility* treatment (0.0592 overall, 0.0503 for males, and 0.009 for females).

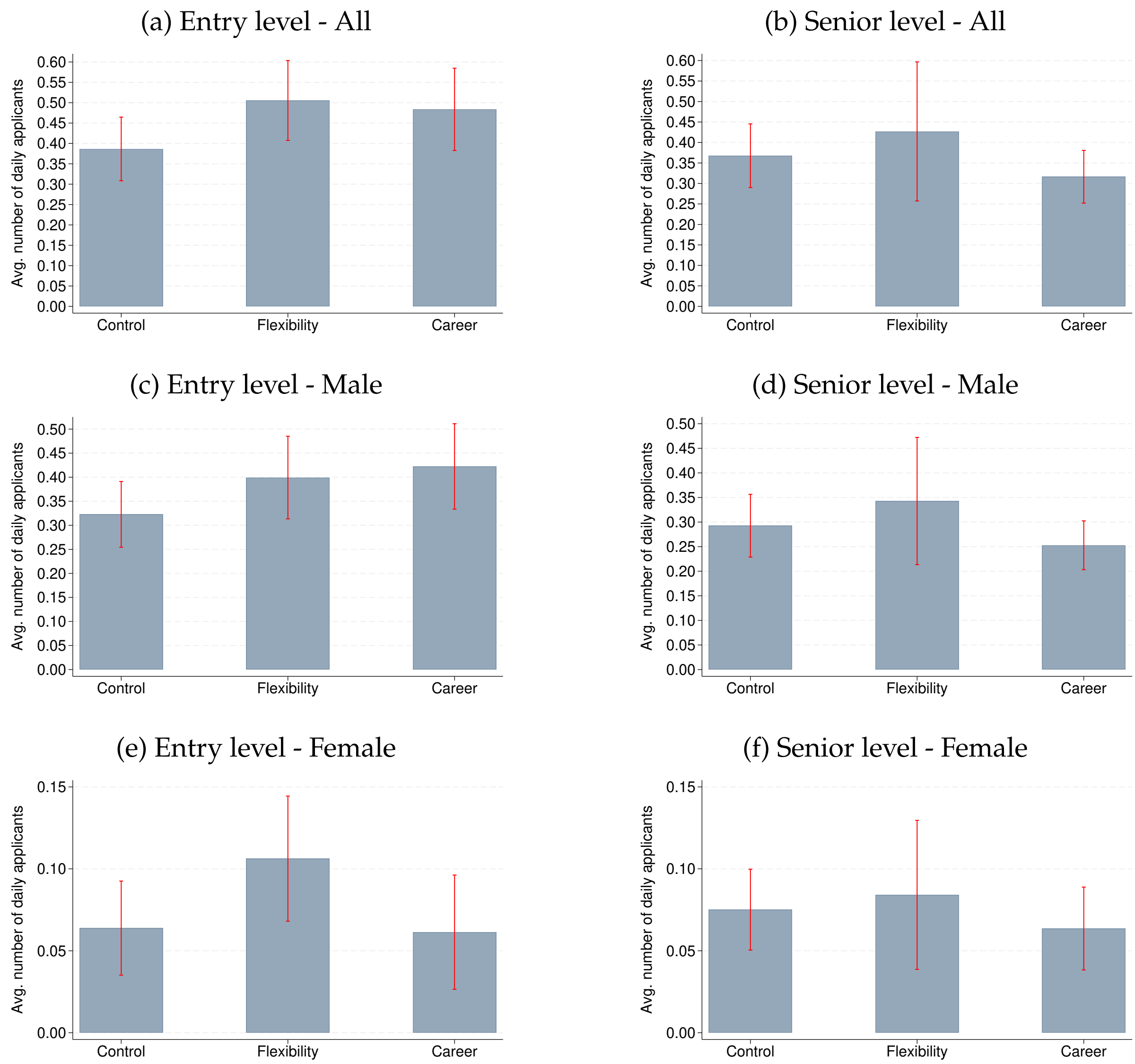

Figure 2.3: Average Number of Daily Applications

Note: This figure shows the average number of daily applications for each treatment by gender and experience level of the job ad. The bar represents the mean, while red lines show 95% confidence bands for the mean. We denote by $\bar{y}_c$ the mean estimator for the *control* treatment, and by $\bar{y}_f$ we denote the mean estimator for the *flexibility* treatment, while by $\bar{y}_{ca}$ we denote the mean estimator for the *career* treatment. Figure 2.3a shows the mean of daily applicants for entry-level positions, with $\bar{y}_c = 0.3865$, $\bar{y}_f = 0.5055$, and $\bar{y}_{ca} = 0.4838$. Figure 2.3c shows the mean of daily male applicants for entry-level positions, with $\bar{y}_c = 0.3227$, $\bar{y}_f = 0.3992$, and $\bar{y}_{ca} = 0.4224$. Figure 2.3e shows the mean of daily female applicants to entry-level positions, with $\bar{y}_c = 0.0638$, $\bar{y}_f = 0.1062$, and $\bar{y}_{ca} = 0.0613$. Figure 2.3b shows the mean of daily applicants to senior-level positions, with $\bar{y}_c = 3678$, $\bar{y}_f = 0.4270$, and $\bar{y}_{ca} = 0.3164$. Figure 2.3d shows the mean of daily male applicants to senior-level positions, with $\bar{y}_c = 0.2926$, $\bar{y}_f = 0.3429$, and $\bar{y}_{ca} = 0.2528$. Figure 2.3f shows the mean of daily female applicants to senior-level positions with $\bar{y}_c = 0.07512$, $\bar{y}_f = 0.0841$, and $\bar{y}_{ca} = 0.0636$.

2.5.2 Empirical Strategy

Our goal is to uncover the causal effect of highlighting flexibility or career advancement on the number of daily applications. Each job ad is observed for both treatments, *flexibility* and *career*, and the control period. Our data thus follow a panel structure that allows us to exploit variation within each of the 105 job ads over a period of 30 days per ad. To uncover the treatment effects of interest, we rely on the following linear specification:

$$y_{it} = \beta_f Flexibility_{it} + \beta_{ca} Career_{it} + \alpha_i + \lambda_t + \epsilon_{it} \tag{2.1}$$

Here, y_{it} denotes the number of applications received for job ad i on day t. The variables $Flexibility_{it}$ and $Career_{it}$ are dichotomous and equal to one if job ad i belongs to the *Flexibility* or *Career* group on day t. The time index $t \in \{1, 2, ...8, 9, 12, 13, ..., 18, 19, 22, 23, ..., 30\}$ denotes the number of days since the job ad first went online. In total, our estimations include 26 observations per job advertisement: on average one per day. As we cannot exactly measure the time of the treatment switch, we exclude the day t of the treatment switch and the day $t + 1$ after the treatment switch.[19] The variable λ_t accounts for time fixed effects, α_i denotes the individual job-ad fixed effect, and ϵ_{it} denotes the error term.

We rely on OLS fixed-effects regressions to derive our main results, but also provide robustness evidence based on Poisson fixed-effects regressions to account for the count-level nature of the dependent variable (see Section 2.9.2 of Appendix A).[20]

2.5.3 Main Result

We proceed by discussing the estimation results from an OLS fixed-effects regression of Equation 2.1, as presented in Table 2.2. Columns 1 to 3 show the estimated treatment effects on the total number of applications to entry-level jobs, while Columns 4 to 6 show the estimated effects for senior-level jobs. All estimations include job ad and time fixed effects and standard errors clustered on job-ad level.[21]

[19]This choice is made to mitigate concerns with respect to potential spillovers. In Section 2.9.2 of Appendix A, we present the results of our main analysis including day $t + 1$. The results are qualitatively similar. Additionally, we present a discussion including further analyses providing evidence that spillovers do not pose an identification threat.

[20]Specifically, due to overdispersion and the presence of inflated zeros, we rely on the Poisson Pseudo Maximum Likelihood estimator. The estimation is implemented in Stata using the *ppmlhdfe* command from the *ppml* package; see Correia, Guimarães and Zylkin (2020).

[21]Although the number of clusters is in an acceptable range to rely on standard clustering methods, we also present the p value of wild bootstrapped standard errors (see Cameron, Gelbach and Miller 2008) in the last two rows of additional statistics of Table 2.2.

We begin to discuss the results for the entry-level job ads. We observe that the *flexibility* and the *career* treatment increase the number of applications on average. The *flexibility* treatment is estimated to increase the total number of daily applications by approximately 0.171, which, given a mean of the *control* treatment of 0.39, corresponds to a relative increase of 44%. The *career* treatment is estimated to increase the total number of daily applications by approximately 0.137, which corresponds to a relative increase of 35%.[22]

Next, we consider the results in Column 2 for male applicants. We observe that both treatments equally increase the number of applications (i.e., the Null $\beta_f = \beta_{ca}$ cannot be rejected), the point estimate for the *flexibility* treatment amounts to 0.119, which corresponds to an increase of 37% and of the *career* treatment to 0.133, which corresponds to an increase of 42%.

Column 3 shows the results for female applicants only. We observe that the *flexibility* treatment is estimated to increase the daily number of female applicants by 0.052, corresponding to an increase of 87%, but no significant increase for the *career* treatment. The null $\beta_f = \beta_{ca}$ is rejected for standard significance levels with a corresponding p-value of 0.012.[23]

Extrapolating these point estimates to a full 30-day period, the *flexibility* treatment is estimated to increase the total number of applications approximately by 5.13. Out of these 5.13 additional applications, 3.57 are estimated to be from male and 1.56 from female applicants. The *career* treatment is estimated to generate 4.11 additional applications, of which roughly all are from male applicants.

Columns 4 to 6 show the results for senior-level positions, and hence job ads requiring previous work experience. Across all three specifications, we observe no treatment effects for the total number of applications, neither in total nor separated for female and male applicants.

[22] Performing the same estimations by means of a Poisson fixed-effects regression – which is presented in Table A.1 in Section 2.9.2 of Appendix A – yields similar results, with even smaller standard deviations of the point estimates and slightly larger relative magnitudes. It is estimated that the *flexibility* treatment increases the total number of applications by 57%, and the *career* treatment is estimated to increase the total number of applications by 40%.

[23] Again, the Poisson regression yields similar results, with estimated increases for the *flexibility* treatment by 47% for males and by 102% for females. The *career* treatment is estimated to increase the number of male applicants by 44%, and no statistical significant increase for female applicants can be ascertained.

Table 2.2: Treatment Effects on the Number of Applications

	No. of applications - OLS					
	Inexperienced			*Experienced*		
	All	Male	Female	All	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Flexibility	0.171**	0.119*	0.052***	0.060	0.054	0.006
	(0.067)	(0.061)	(0.018)	(0.119)	(0.096)	(0.026)
Career	0.137*	0.133*	0.004	-0.028	-0.021	-0.007
	(0.079)	(0.072)	(0.023)	(0.033)	(0.028)	(0.017)
Observations	829	829	829	1896	1896	1896
No. of Clusters	32	32	32	73	73	73
Control mean dep. variable	0.39	0.32	0.06	0.37	0.29	0.08
Bootstrap p β_f	0.02	0.06	0.01	0.89	0.83	0.93
Bootstrap p β_{gr}	0.11	0.07	0.92	0.39	0.43	0.77

Note: This table shows the impact of the treatments on the number of applications received per day. The estimates are obtained using standard OLS fixed-effect regressions; thus, the marginal effects need to be interpreted in terms of change in the number of applications per day. All specifications include job-ad and time fixed effects. Columns 1 to 3 present the effects for job ads requiring no previous work experience, while Columns 4 to 6 present the effects for job ads requiring previous work experience. Column 1 and 4 show the effect for the total number of applications, while Columns 2 and 5 only show the effect for the number of male applicants, and Columns 3 and 6 only show the effect for the number of female applicants. Standard errors clustered on job-ad level are reported in parentheses. The last two rows show the p-values from wild bootstrapped clustered standard errors (Cameron, Gelbach and Miller, 2008).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Next, we relate the results to the predictions derived from our conceptual framework. We find mixed evidence with respect to prediction 1 regarding the increase in applications for both treatments. We find that this increase is only present for entry-level job ads requiring no previous work experience. However, this equally provides strong support for our second prediction, which states that the increase in applications should be larger for entry-level positions than for senior-level positions. We cannot reject that the treatment coefficients for male applicants are statistically distinguishable; however, we can indeed reject this hypothesis for female applicants. Thus, we find mixed evidence for prediction 3, as both treatments seem to be equally attractive for male applicants, but the *flexibility* treatment only attracts additional female applicants.

2.5.4 Further Results

In this section, we analyze how the composition of the applicant pool is affected. We present two sets of analyses. First, we analyze changes to the distribution of the applicants' region of residence. As highlighted by Moretti (2011), an increase in amenities can cause an exogenous labor-supply shock that may increase worker mobility. Second, job-specific abilities or social preferences may correlate with preferences for certain amenities affecting the quality composition of the applicant pool (e.g., Deserranno 2019; Nekoei 2023).

Region of Residence

We categorize the applications by applicants living in the federal state of the location of the firm (State), applicants living in Germany, but not in the federal state of the firm (Germany w/o state), and applicants from abroad (Abroad).

We start with a descriptive analysis by considering Figure 2.4, which presents the mean of the respective daily number of applications by each region of residence category for each treatment. Figure 2.4a shows the mean of daily applicants living in Germany w/o state, while Figure 2.4b shows the mean of daily applicants living in the federal state, and Figure 2.4c that of the number of daily applicants living abroad. Considering the bar charts, we observe strong increases of applicants from Germany w/o state (increases by 0.089 for the *flexibility* treatment and by 0.0965 for the *career* treatment), while we observe no remarkable increases for applicants from the two other categories. Already simple t-tests for mean comparison confirm this, as the difference of means for applicants from Germany w/o state is significant for both treatments, while we find no significant differences for the other two regional categories.[24] As an alternative, in Table A.3 of Appendix A, we present the re-estimation of Equation 2.1 with the applicants from a particular region category as outcome variable. The results are similar to the mean comparisons presented above.

This provides evidence that highlighting flexible work opportunities and career-progress opportunities allows the firm to source from a larger regional talent pool. However, the informational treatments do not seem to be large enough to be pivotal for an application decision for applicants living abroad, which is in line with the discussions of Moretti (2011) that worker mobility is finite.

[24] The null that $\bar{y}_f < \bar{y}_c$ and $\bar{y}_{ca} < \bar{y}_c$ for applicants from Germany w/o state can be rejected for standard significance levels. For $H_0 : \bar{y}_f < \bar{y}_c$, we reject at the 5% level ($p = 0.0106$), and for $H_0 : \bar{y}_{ca} < \bar{y}_c$, we reject at the 1% level ($p = 0.0092$).

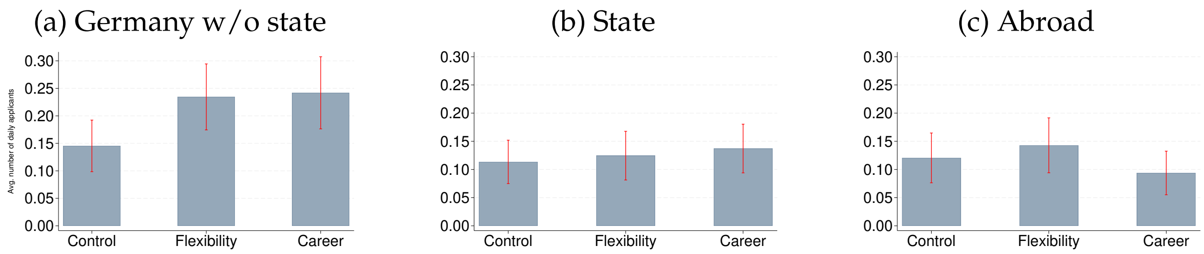

Figure 2.4: Average Number of Daily Applications by Region of Residence

Note: This figure shows the average number of daily applicants for each treatment and by region of residence of applicants. Figure 2.4a shows the numbers for applicants living in Germany, but not in the state of the firm (Germany w/o state), while Figure 2.4b shows the number of applicants living in the federal state of the firm (State), and Figure 2.4c shows the average number of applicants living abroad (Abroad). We denote the mean of the flexibility treatment by $\bar{y}_f$, of the career treatment by $\bar{y}_{ca}$, and of the control group by $\bar{y}_c$. For Figure 2.4a, $\bar{y}_c = 0.1454$, $\bar{y}_f = 0.2344$, and $\bar{y}_{ca} = 0.2419$. For Figure 2.4b, $\bar{y}_c = 0.1135$, $\bar{y}_f = 0.1245$, and $\bar{y}_{ca} = 0.1372$. For Figure 2.4c, $\bar{y}_c = 0.1206$, $\bar{y}_f = 0.1428$, and $\bar{y}_{ca} = 0.0939$. For applications from Germany w/o state, the null that $\bar{y}_f < \bar{y}_c$ and $\bar{y}_{ca} < \bar{y}_c$ for the daily applicants can both be rejected at the 5% level with p-values of 0.0106 and 0.0092, respectively. For all other applicants (from the State and Abroad), we cannot reject the null of smaller means of the treatment groups.

To check how the overall distribution of applicants is affected, we re-estimate Equation 2.1 using standard OLS fixed-effects regressions on the applicant level. Each observation now corresponds to an applicant for job ad i on a particular day t. The result is a linear probability model, which is able to detect whether probability mass is shifted to one category, as the point estimates give the marginal probability increase of belonging to a certain category upon coming from either the *flexibility* or the *career* treatment.

Table 2.3 presents the results. Column 1 shows the marginal probability change of an applicant living in Germany w/o state, while Column 2 shows the marginal probability change of an applicant living in the state, and Column 3 shows the marginal probability change of an applicant living abroad, conditional on an application coming from the *flexibility* or the *career* treatment in comparison to the *control* group. In line with the previously shown mean comparisons, we observe that, for both treatments, applicants are more likely to live in Germany w/o state (an increase of 0.133 for the *flexibility* and of 0.149 for the *career* treatment), while we observe no statistically significant changes for the *flexibility* treatment for both other regional categories. For the *career* treatment, we observe no change to the share for applicants living in the federal state of the firm, but a negative statistically significant point estimate for applicants living abroad (of -0.148). This negative point estimate does not imply an absolute reduction in the number applicants from abroad. Rather, it points towards a distributional change in favor of applicants from Germany w/o state and from the federal state.

Table 2.3: Treatment Effects on the Number of Applications by Region of Residence

	Region of residence of applicants - OLS		
	Germany w/o state	State	Abroad
	(1)	(2)	(3)
Flexibility	0.133*	-0.067	-0.050
	(0.076)	(0.070)	(0.077)
Career	0.149***	0.004	-0.148**
	(0.050)	(0.057)	(0.070)
Observations	380	380	380
No. of Clusters	32	32	32
Mean dep. variable	0.45	0.27	0.26

Note: This table shows the effect of the treatments on the distribution of region of residence of the applicants. The outcome variables are binary indicators in case applicants live in Germany, but not the federal state of the firm (Germany w/o state), live in the federal state of the firm (State), or live abroad (Abroad). All estimations include job-ad and time fixed effects and are estimated via standard OLS fixed-effects regressions. Thus, the model corresponds to a linear probability model, and the point estimates can be interpreted as marginal probability increases. The interpretation corresponds to the marginal increase in probability of an applicant belonging to one of the categories upon applying to a particular treatment. Standard errors clustered on job-ad level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Quality

Similarly to the previous section, we start with a descriptive analysis by considering Figure 2.5, which presents the average daily number of applications rated either with an A (2.5a), B (2.5b), or C (2.5c), or those applications that are screened out for each treatment (2.5d).

Comparing simply the means from the graphs, we observe that the increase in applicants in response to both treatments is quite evenly distributed across categories. We only note that the *career* treatment seems to induce an even larger increase of A-rated applicants (by 0.016 for the *flexibility* treatment and by 0.037 for the *career* treatment).[25] As an alternative, in Section 2.9.2 in Table A.4, we present the re-estimation of Equation 2.1, using the applicants with a particular rating category as outcome variable. The results are similar to the mean comparisons presented above.

[25]This is confirmed by the fact that a t-test rejects the null of equal means for the daily applicants with an A rating.

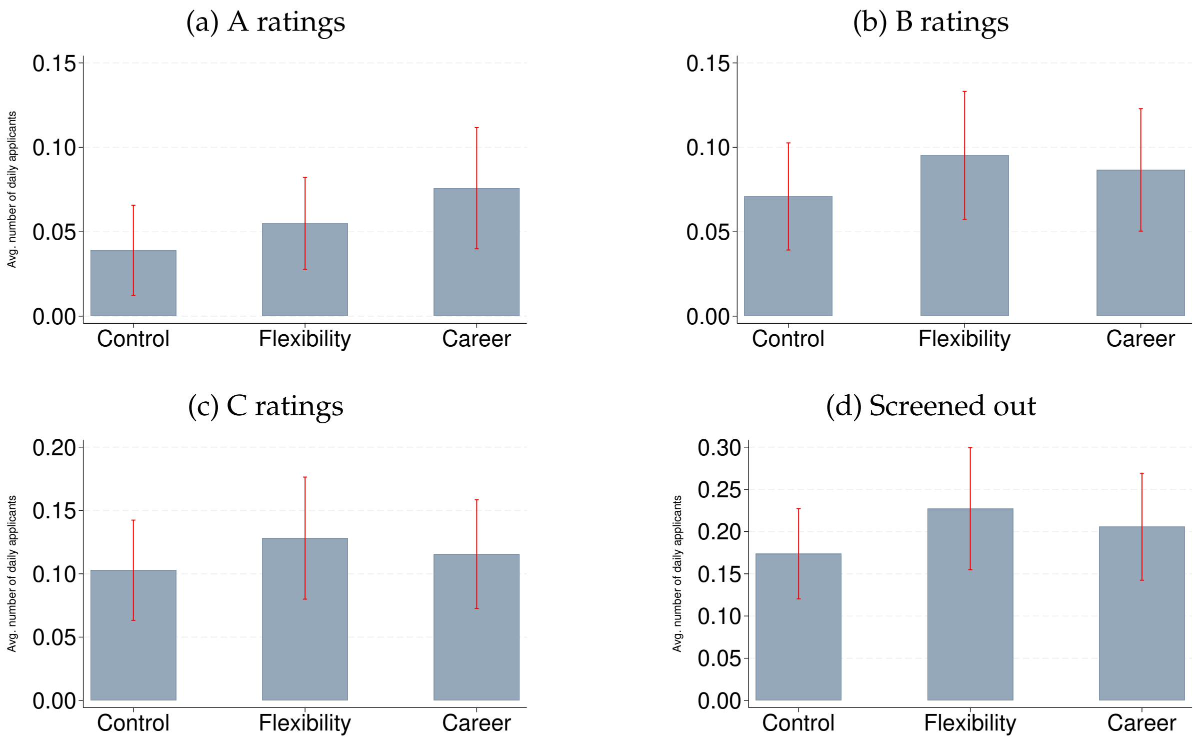

Figure 2.5: Average Number of Daily Applications by Quality

Note: This figure shows the average number of daily applicants for each treatment and by quality. Figure 2.5a shows the mean number of daily applicants with an A rating, while Figure 2.5b shows the mean number of applicants with a B rating, Figure 2.5c shows the mean number of applicants with a C rating, and Figure 2.5d shows the mean number of applicants who have been screened out. We denote the mean of the flexibility treatment by $\bar{y}_f$, of the career treatment by $\bar{y}_{ca}$, and of the control group by $\bar{y}_c$. For Figure 2.5a, $\bar{y}_c = 0.0390$, $\bar{y}_f = 0.0550$, and $\bar{y}_{ca} = 0.0758$. For Figure 2.5b, $\bar{y}_c = 0.0709$, $\bar{y}_f = 0.0952$, and $\bar{y}_{ca} = 0.0866$. For Figure 2.5c, $\bar{y}_c = 0.1028$, $\bar{y}_f = 0.1282$, and $\bar{y}_{ca} = 0.1155$. For Figure 2.5d, $\bar{y}_c = 0.1738$, $\bar{y}_f = 0.2271$, and $\bar{y}_{ca} = 0.2058$. For the mean of daily applications with an A rating the null hypothesis that $\bar{y}_f < \bar{y}_c$ cannot be rejected, while $\bar{y}_{ca} < \bar{y}_c$ can be rejected at the 10% level ($p = 0.0529$). The means of the other categories do not differ significantly from each other.

Similarly to the investigation of the region of residence of applications, we want to understand whether the treatments cause a change in the overall distribution of ratings. We approach this by re-estimating Equation 2.1 on the applicant level, i.e., conditional on having applied. This means that each observation corresponds to one applicant for job ad i on day t and that the point estimates identify marginal probability increases with respect to one rating category.

Table 2.4 presents the results. Column 1 shows the results for A ratings, Column 2 for B ratings, Column 3 for applicants with C ratings, and Column 4 for applicants who have been screened out. Overall, we observe no strong distributional changes for both treatments. Considering the point estimates for the *flexibility* treatment from Columns 1 to 4, we observe point estimates close to zero, which are insignificant. This shows that the *flexibility* treatment managed to attract additional applicants without compromising the quality distribution in terms of recruiter ratings. Considering the point estimates for the *career* treatment, we also observe no statistically significant, and point estimates are close to zero for applicants rated B and C. However, we observe a positive point estimate with a t-statistic of 1.55, for A-rated applicants and a negative point estimate of similar size for screened-out applicants with a t-statistic of -1.05, mirroring the descriptive finding of an even larger increase in A-rated applicants with a comparably lower increase of screened out applicants attracted by the *career* treatment.[26]

Table 2.4: Treatment Effects on the Number of Applications by Quality

	Rating and hiring outcomes - OLS				
	A rating	B rating	C rating	Screened out	Interview
	(1)	(2)	(3)	(4)	(5)
Flexibility	-0.002	0.001	-0.001	0.002	-0.004
	(0.035)	(0.057)	(0.049)	(0.053)	(0.056)
Career	0.060	0.003	0.001	-0.064	0.103*
	(0.039)	(0.048)	(0.055)	(0.062)	(0.057)
Observations	380	380	380	380	380
No. of Clusters	32	32	32	32	32
Mean dep. variable	0.11	0.18	0.25	0.25	0.25

Note: This table shows the effect of the treatments on the distribution of rating categories of applicants. The outcome variables are binary indicators in case applicants are rated with an A (best category), B, C, screened out (least good category), or were invited to an interview. All estimations include job-ad and time fixed effects and are estimated via standard OLS fixed-effects regressions. Thus, the model corresponds to a linear probability model, and the point estimates can be interpreted as marginal probability increases. The interpretation corresponds to the marginal increase in probability of an applicant belonging to one of the categories upon applying to a particular treatment. Standard errors clustered on job-ad level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

To complement the analysis, Column 5 of Table 2.4 presents the results of a linear probability model, for which the outcome variable is another quality indicator, namely whether an applicant is invited to an interview. We estimate the marginal probability increase of an applicant being invited to an interview upon having applied to the

[26]Furthermore, for the estimation in Column 1, the null that $\beta_f > \beta_{ca}$ can be rejected at the 10% level ($p = 0.065$).

flexibility or the *career* treatment. Corresponding to the indication of the weak positive distributional change, we observe a positive weakly significant point estimate of 0.103 for the *career* treatment, indicating a higher likelihood of being invited for an interview when applying to a *career*-treatment job ad.

Overall, we conclude that the analysis provides evidence that the additional applicants were attracted without significant changes to the quality distribution of applicants. More precisely, we find no indications of changes for the *flexibility* treatment and even weak indications of a positive shift for the *career* treatment.

2.6 Survey Experiment and Mechanisms

In Section 2.3, we predicted that the effects of highlighting flexibility or career-advancement opportunities in job ads on applicants' behavior is driven by updating beliefs among potential applicants about job characteristics and the working environment. To assess the plausibility of beliefs as an underlying mechanism of our main treatment effect among inexperienced workers, this section presents the results of a survey experiment with STEM students. Further, we analyze gender differences in workplace preferences.

2.6.1 Experimental Design

The job ads for entry-level positions are targeted at candidates who recently graduated from university in a STEM field. In line with this target group, we collected survey responses from a total of 2,136 STEM-graduates across 12 different labs in Germany and Austria.[27] As most of these participants recently graduated, or were about to graduate, they were an idea pool to elicit beliefs about the job characteristics and work-environment in entry-level STEM positions. As the presented job ads are for high-skilled and complex jobs in the technology industry, it is important to align the required educational background of the job ad with the actual educational background of the lab participant. Thus, we invited only individuals who possessed the educational background required by the job ad. This ensures more reliable answers, as those participants were better informed about the tasks outlined in the job ad and the industry overall.

[27] Detailed information about the labs and participant numbers can be found in Table A.6 of Appendix A.

The experimental procedure was as follows: Whenever an entry-level job ad was posted and part of our field experiment, we initiated a corresponding lab session. We thus conducted the survey experiment in real time, aligning it with the company's actual recruitment period for the position. This is something we communicated as part of the survey to create a more realistic atmosphere without being deceptive.[28] As the supply of students with a STEM background in economic research labs at universities was limited, we needed to contact many different labs at different universities to gather a sufficient number of responses. Due to administrative procedures and guidelines, not all the labs were available at the same time, but rather on a rolling basis over the course of our field experiment. Due to the restrictions of the size of the participant pool, only 20 out of 32 entry-level positions in our main data were part of the survey experiment.[29] Our target for each survey wave was to recruit at least 45 participants. In total, we conducted 47 different waves with a total of 20 job ads.[30] All job ads were part of more than one survey wave to ensure that we could include lab fixed effects.

The structure of the survey was as follows: The survey started with some questions about the educational background, demographics, and preferences for workplace characteristics of the participants. The second and main block of the survey showed participants a job ad from our field experiment and informed them that this was a real job currently posted by the company. The name of the firm was revealed, and we presented the job ad either with the *control*, the *flexibility*, or the *career* treatment. Thereafter, we elicited the subjects' beliefs about job characteristics as well as the working environment. We removed the information from the job ad about the workplace location to avoid confounding across lab locations. Instead, we asked participants to assume that the place of work was at a reasonable distance to their current place of living. The last block asked participants about their interest in the presented job.[31] In Section 2.9.3 of Appendix A, we present summary statistics of the variables measured as part of the survey in Table A.7.

[28]We selected job ads for real positions that were actively posted at the time, allowing students also to apply for these roles as part of the survey. Towards the end of the process, students had the opportunity to contact the firm directly in order to signal their interest in the job and to receive instructions on how to apply. It is important to note that not even a handful of students (3 out of 2136) actually availed of this opportunity. We tracked them using unique IDs that corresponded to treatment and the specific job advertisement. This method allowed us to identify these individuals in the field-experiment dataset, enabling us to filter out applications that potentially skew our treatment effects.

[29]We did not randomize the job ads. Whether a job ad was part of the survey experiment depended solely on the availability of an economics research lab, a sufficiently large participant pool, and the job ad being online during the availability of the pool.

[30]To increase the quality of respondents' answers, we removed the fastest 5% of respondents.

[31]The questionnaire of the survey experiment is available from the authors upon request.

2.6.2 Beliefs about Job Characteristics

The main focus of the survey experiment was to measure how our treatment shapes beliefs about job characteristics. To do so, we relied on a battery of questions that are based on Ronen (1994) and have also been applied in other studies investigating job characteristics (see, e.g., Gill et al. 2023). In particular, we asked questions about the expected work-life balance, avoid overtime at work, the opportunity for part-time work, for flexible scheduling, the attractiveness of the location of the job, the necessity of work-related travel, job security, provision of a high income, prospects of salary growth, salary negotiation possibilities, a family-friendly workplace, career-advancement opportunities, the firm's reputation, how challenging the tasks of the job are, the childcare support offered by the company, and the possibility to work from home (home-office). Participants were asked to rate statements about these items on a scale from 1 (does not apply at all) to 10 (fully applies) from the perspective of how accurately they expected these statements to describe the presented job.[32]

Our analysis serves two primary purposes. First, we examine the impact of our treatments on two categories: *work-life balance* and *career benefits*. The *work-life balance* category encompasses expected work-life balance, flexible scheduling, home-office opportunities, childcare support, avoidance of overtime, and family-friendly job characteristics. The *career benefits* category contains the following items: good salary, possibility of salary growth, career-advancement opportunities, the level of challenge of the individual job tasks, and the opportunity of regular salary negotiations. Our outcome variables are composite scores for each category, calculated as the standardized sum of the ratings for each item within the category. Second, we aim to identify which individual items contribute most to the composite scores and are the main driver behind our observed shift in beliefs.

To identify the treatment effects, we estimate an equation similar to 2.1 of the main analysis, with the outcome variables being our two composite scores of i) *work-life balance* and ii) *career benefits* items. In this model, we include lab fixed effects instead of time fixed effects, in addition to job-ad fixed effects.[33] Additionally, we include further control variables such as gender, high school GPA, migration background, university degree,

[32]For our analysis, we exclude the items on beliefs regarding the location, opportunity for part-time work, work-related travel, job security, and reputation of the firm. These items are not useful for our analysis, as the job security in Germany is extremely high for permanent positions, and strongly regulated; whether the job is full-time or part-time is stated in the ad; and work-related travel is also job-dependent and outlined, if applicable, in the job description. In Section 2.9.3 of Appendix A, Table A.10 presents the regression results for these excluded items in Columns 1 to 5.

[33]Our results remain the same when we use principal component analysis and apply endogenous weights to the collection survey items.

and family status. Table A.5 of Appendix A Section 2.9.3 provides detailed descriptions of the variables. As our outcome variables are standardized, the estimated marginal effects need to be interpreted in terms of standard deviations (sd) of the respective composite score.

Table 2.5: Belief-Updating about Job Characteristics

	Beliefs					
	Work-life balance			*Career benefits*		
	(1)	(2)	(3)	(4)	(5)	(6)
Flexibility	0.106**	0.131***	0.132***	-0.017	-0.008	-0.008
	(0.041)	(0.042)	(0.042)	(0.048)	(0.051)	(0.051)
Career	-0.112**	-0.096**	-0.094**	0.159**	0.163***	0.162***
	(0.044)	(0.044)	(0.044)	(0.057)	(0.056)	(0.056)
Observations	2136	2136	2136	2136	2136	2136
No. of Clusters	20	20	20	20	20	20
Job FE	Yes	Yes	Yes	Yes	Yes	Yes
Lab FE	No	No	Yes	No	No	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Bootstrap p β_f	0.02	0.01	0.01	0.67	0.94	0.85
Bootstrap p β_{ca}	0.01	0.04	0.04	0.02	0.02	0.01

Note: The table shows the impact of the treatments on the beliefs about job characteristics. Work-life balance adds up beliefs about flexibility, work-life balance, home-office, childcare support, projected overtime, and a family-friendly workplace culture. Career benefits adds up beliefs about expected salary, salary growth, career opportunities, degree of challenge of the tasks, and the possibility to negotiate salary increases on a regular basis. The outcome variables are standardized; thus, the marginal effects need to be interpreted in terms of standard deviations. The control variables include gender, high school GPA, migration background, the university degree, and family status. Standard errors are clustered on job-ad level and are reported in parentheses. The last two rows show the p-values from wild bootstrapped clustered standard errors (Cameron, Gelbach and Miller, 2008).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The results are presented in Table 2.5. Columns 1 to 3 show the effect on the composite score of *work-life balance*, while Columns 4 to 6 show the effects on the composite score of *career benefits*. Columns 1 and 4 present the most parsimonious specification and only include job-ad fixed effects, while Columns 2 and 5 further include additional control variables, and Columns 3 and 6 present the most comprehensive specification including, in addition, lab fixed effects. To interpret our results, we focus on our most comprehensive specifications in Columns 3 and 6. We observe that the *flexibility* treatment leads to an increase about 0.132 sd of the expected *work-life balance* provided by the job, while we observe small and noisy point estimates close to zero regarding the provided opportunities for *career benefits*. Considering the effect of the *career* treatment, we observe that it increases beliefs about the provided *career benefits* by 0.162 sd, while at the same time decreasing the beliefs about the provided *work-life balance* by 0.094 sd.[34]

To provide a deeper understanding of the main drivers behind the observed belief shifts in our two composite scores, we present treatment effects for each item in Table A.8 (Appendix A Section 2.9.3). The *flexibility* treatment significantly increases beliefs that the job offers more flexible scheduling, better work-life balance, and home-office opportunities. While the point estimates for childcare opportunities, family-friendly workplace, and the possibility to avoid overtime are positive, they do not appear to be statistically significant. Conversely, the *career* treatment shows negative point estimates for items contributing to the *work-life balance* score, with work-life balance itself and the ability to avoid overtime being the most affected. However, this treatment exhibits positive point estimates for all items contributing to our *career benefits* indicator. The most substantial effects are observed in perceptions of career advancement opportunities, salary growth prospects, and increased possibilities for salary negotiation.

Summarizing the results, we find evidence that the treatments indeed lead to belief-updating among potential applicants. We find strong support for the prediction, developed in our conceptual framework, that both treatments lead to a positive shift in beliefs about the provided job flexibility and career-advancement opportunities. Interestingly, we find evidence that potential applicants perceive a trade-off between the provided career benefits and work-life balance, as the *career* treatment leads to an increase of the former, but to a decrease of the latter.[35]

2.6.3 Beliefs about the Working Environment

A second purpose of the survey experiment is to analyze whether the treatments also affect beliefs about the working environment. As part of the survey, in a second battery of questions we elicited beliefs about the expected share of direct colleagues with a particular personal or character attribute. We focus on six items, the believed share of direct colleagues (i) being female, (ii) having a family, (iii) prioritizing career over family,

[34]Further, to investigate heterogeneities in belief-updating with respect to gender, we re-estimate the equation and include an interaction term for being female. No statistically significant effects were identified in this analysis.

[35]Relating this to our conceptual framework in Appendix A in Section 2.9.1, we find evidence that $\tilde{\rho} < 0$.

(iv) eager to have a career, (v) having a STEM degree, and (vi) earning a high income.[36] We allocate these items again into two categories, to calculate composite scores over two aggregated items: the first category is a *friendly* working environment (analogous to work-life balance), and *competitive* working environment (analogous to *career benefits*).

The first outcome, *friendly* working environment, is measured by the standardized sum of the scores of (i) and (ii), while the second outcome variable is *competitive* environment and is measured by the standardized sum of the scores of (iii) to (vi).[37] We identify treatment effects as in the previous section by re-estimating Equation 2.1 with lab fixed effects (instead of time fixed effects), job-ad fixed effects and additional controls. As the outcome variables are standardized, we need to interpret the marginal effects terms of standard deviations.

The results are presented in Table 2.6. Columns 1 to 3 show the effect on the composite score of a *friendly* working environment, while Columns 4 to 6 show the effects on the composite score of a *competitive* working environment. Columns 1 and 4 show the most parsimonious specification and only include job-ad fixed effects, Columns 2 and 5 further include additional control variables, and Columns 3 and 6 show the most comprehensive specification including lab fixed effects. To interpret our results, we focus on our most comprehensive specifications in Columns 3 and 6. We observe that the *flexibility* treatment leads to an increase of 0.086 sd in expected friendliness of the working environment, while we observe no effects for the *career* treatment. Considering the effect of the *career* treatment, we observe that it leads applicants to believe that the working environment is by 0.092 sd more competitive, while we observe no effects for the *flexibility* treatment. Overall, the results are smaller and statistically less precise.

[36]As a further item, we also elicited the share of colleagues over a particular age as a distraction item, which we exclude in this analysis. Table A.10 in Appendix A presents the regression result for this item in Column 6.

[37]In Table A.9 in Section 2.9.3 of Appendix A, we present estimations of the treatment effects for each single item.

Table 2.6: Belief-Updating about Working Environment

	Beliefs					
	Friendly			*Competitive*		
	(1)	(2)	(3)	(4)	(5)	(6)
Flexibility	0.081*	0.084*	0.086*	-0.050	-0.050	-0.041
	(0.045)	(0.044)	(0.044)	(0.051)	(0.054)	(0.054)
Career	0.029	0.033	0.033	0.083	0.088*	0.092*
	(0.041)	(0.040)	(0.040)	(0.050)	(0.050)	(0.051)
Observations	2136	2136	2136	2136	2136	2136
No. of Clusters	20	20	20	20	20	20
Job FE	Yes	Yes	Yes	Yes	Yes	Yes
Lab FE	No	No	Yes	No	No	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Bootstrap p β_f	0.07	0.05	0.04	0.33	0.47	0.45
Bootstrap p β_{ca}	0.46	0.41	0.47	0.12	0.11	0.09

Note: This table shows the impact of the treatments on the beliefs about the working environment. *Friendly* working environment adds up beliefs about the share of colleagues being female and having a family. *Competitive* working environment adds up survey questions about beliefs about the share of colleagues prioritizing career over family, being eager to have a career, having a STEM degree, and earning a high income. The outcome variables are standardized; thus, the marginal effects need to be interpreted in terms of standard deviations. Controls include gender, high school GPA, migration background, university degree, and family status. Standard errors clustered on job-ad level are reported in parentheses. The last two rows show the p-values from wild bootstrapped clustered standard errors (Cameron, Gelbach and Miller, 2008).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Our results in Table 2.6 indicate that our *flexibility* treatment is positively associated with a more *friendly* working environment, whereas the *career* treatment is associated with a more *competitive* working environment.

To further elucidate the main drivers behind the observed belief shifts, we present treatment effects for each item in Table A.9 (Appendix A, Section 2.9.3). Our analysis reveals that the impact of the *flexibility* treatment on our composite score of a *friendly* work environment is jointly significant. While there are no significant net effects regarding the share of females or the proportion of workers with families in the workforce, the point estimates for the *friendly* work environment indicator are positive (Columns 1 and

2). Both of these factors appear to be equally important in terms of magnitude. The *career* treatment, on the other hand, shows positive point estimates for individual items associated with a *competitive* environment. Among these items, 'ambitous' stands out as the most significant driver, both in terms of magnitude and statistical significance.

Overall, our results show that the information treatments extend beyond belief-updating about job characteristics onto belief-updating about selection into the work environment.

2.6.4 Heterogeneity of Worker Preferences by Gender

Our analysis explores gender differences in workplace preferences, particularly regarding flexibility and career advancement. Building on Wiswall and Zafar (2018), we hypothesized that flexibility-focused job advertisements would attract more female applicants, while career-focused advertisements would attract more male applicants (Prediction 3 in Section 2.3). Our main analysis in Section 2.5.3 yielded mixed evidence: while male applicants responded similarly to both treatments, female applicants were significantly more responsive to the flexibility treatment.

To identify the mechanisms behind these differential responses, we analyze our survey data on workplace preferences in Table 2.7. Our survey captured preferences for *flexibility* attributes in Columns 1 to 6 (including work-life balance, home-office, childcare support, and overtime expectations) and *career* attributes in Columns 7 to 11 (such as salary prospects, advancement opportunities, and negotiation potential). We estimate the gender differences by regressing each workplace preference item on a female indicator, controlling for high school GPA, migration background, university degree, and family status. The constant term represents male preferences, while the sum of the constant and female coefficient captures female preferences. These means by gender are provided in the additional statistics in Table 2.7. We find that while men value both sets of attributes, women demonstrate a significantly stronger preference for flexibility than men. These survey results complement our findings from Table 2.2 and suggest that these effects arise due to gender differences in workplace preferences.

Table 2.7: Gender Differences in Workplace Preferences

	Preferred workplace characteristics										
	Flexibility						*Career advancement*				
	(1) Flex-ibility	(2) Work-life balance	(3) Home-office	(4) Child-care	(5) Family	(6) Over-time	(7) Salary	(8) Career	(9) Salary growth	(10) Chall-enge	(11) Negot-iation
Female	0.422***	0.600***	0.583***	0.989***	0.678***	0.244**	-0.109	0.059	-0.128	-0.031	-0.256**
	(0.091)	(0.075)	(0.110)	(0.130)	(0.110)	(0.108)	(0.077)	(0.091)	(0.083)	(0.092)	(0.110)
Constant	6.675***	7.829***	6.133***	3.711***	6.683***	4.277***	7.674***	7.178***	7.558***	6.478***	6.256***
	(0.194)	(0.145)	(0.230)	(0.253)	(0.213)	(0.227)	(0.153)	(0.189)	(0.165)	(0.188)	(0.216)
Male	6.675	7.829	6.133	3.711	6.683	4.277	7.674	7.178	7.558	6.478	6.256
Female	7.097	8.429	6.716	4.700	7.361	4.521	7.565	7.237	7.430	6.447	5.999
Observations	2136	2136	2136	2136	2136	2136	2136	2136	2136	2136	2136
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table shows the impact of gender on workplace preferences. The outcome variables represent aspects of workplace preferences independent of the job advertisement. These include expected flexibility, work-life balance, the option to work from home, childcare support, a family-friendly workplace culture, projected overtime, expected salary, career opportunities, salary growth, and the possibility of negotiating salary increases. The control variables are high school GPA, migration background, university degree, and family status. We use robust standard errors and report them in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

2.7 Robustness

We present several robustness checks with respect to our main analyses, which are detailed in Section 2.9.2 of Appendix A.

To account for the count-level nature of the dependent variable, we present a re-estimation of Equation 2.1 to estimate the treatment effects using a Poisson fixed-effects regression. Our analyses yields similar results, with even smaller standard deviations of the point estimates and slightly larger magnitudes (see Table A.1).

Further, we investigate potential spillover effects that may arise if applicants are exposed to multiple treatment conditions over time. Such spillovers could lead to a downward bias in our main estimates. To alleviate this concern, we excluded the day of the treatment switch and the following day from our main analysis in Section 2.5.1. To further examine spillovers, we conduct two sets of analyses. First, we re-estimate our main regression model with interaction terms for each 10-day period. The results, presented in Column 1 of Table A.2, show no evidence of strong time trends in the treatment effects. This suggests a lack of spillovers, as such effects should manifest in changing treatment impacts over time. Second, we re-estimate the main model including lagged treatment variables. Columns 2 to 4 of Table A.2 show the point estimates remain stable when accounting for lagged treatments, with only a weakly significant coefficient on the lagged flexibility treatment. Overall, this provides strong evidence that spillovers do not meaningfully impact the size or significance of our main treatment effects.

In addition, we present regression-based analyses providing alternative estimators for the treatment effects on the composition of the applicant pool in terms of region of residence (Table A.3) and quality (Table A.4), which were descriptively analyzed in Section 2.5.4.

Lastly, to rule out the possibility that the treatment effects are driven by age-related changes in preferences for flexibility and career advancement, we provide evidence that the treatment effects are not attributable to changing workplace preferences due to age. As shown in Figure A.1, the average marginal effects of age on preferences for flexibility and career advancement are relatively constant and mostly insignificant. Linear age trends yield small, statistically insignificant coefficients of 0.001 for both preference dimensions ($p = 0.289$ and $p = 0.407$, respectively). This suggests preferences for flexibility and career advancement do not significantly vary with age.

2.8 Conclusion

In a rapidly growing technology industry, where high-skilled human capital is a key strategic resource, firms face significant challenges in attracting new talent (Coff 1997; Bapna et al. 2013; Del Carpio and Guadalupe 2022). By conducting a field experiment at one of the largest European technology firms, we demonstrate that highlighting flexibility and career-advancement opportunities can increase the number of applications and the regional scope of the applicant pool for entry-level positions. Importantly, this increase in applications occurred without heavy trade-offs in terms of applicant quality. At most, we find weak evidence for a relatively more positive selection compared to featuring flexible working conditions. Highlighting amenities and benefits in job advertisements is thus an effective and rather cost-efficient tool to increase the number of applications, making it an important strategy in the firms' "war for talent". Our finding that flexible work opportunities particularly attract female applicants provides valuable insights for firms and policymakers seeking to promote gender equality in organizations.

We complemented the field experiment with a survey experiment to examine how the treatment affects young professionals' beliefs and expectations about job characteristics. Highlighting flexibility in job ads shifts beliefs towards a better work-life balance, while highlighting career advancement leads potential applicants to expect higher career benefits and an inferior work-life balance. Potential applicants also update beliefs about the working environment. When flexibility is highlighted, they induce the workplace to be more family-friendly. Career advancement is associated with a more competitive workplace, which is in line with prior findings by Belot, Kircher and Muller (2022). Our results thus unveil the importance of job ads in shaping applicants' beliefs about job characteristics and the working environment, with potential implications for a firm's overall reputation.

Our findings deliver important insights on how information provision shapes the selection of workers into jobs. First, they show that very minor changes can have substantial effects on application behavior. This hints towards important information frictions on the labor market for entry-level jobs (see, e.g., Pissarides 2011; Belenzon and Tsolmon 2016) and towards astonishing consequences, given that the decision over a first job can have long-lasting implications for an individual's career (Kahn 2010). In this regard, our results speak to a literature showing that small nudges can have substantial and lasting effects on individuals and organizations (Hong, Hossain and List 2015). Second, by highlighting job amenities instead of explicitly searching or not searching for certain types of workers (Kuhn and Shen 2023; Flory, Leibbrandt and List 2015), we show that, even in regular job ads, the provided information can have important implications for the size and composition of the applicant pool. In this sense, our study provides a link between the (survey) literature on preferences for job attributes (Wiswall and Zafar 2018; Gill et al. 2023) and the literature on worker selection into firms (see, e.g., Nekoei 2023; Gill et al. 2023; DeVaro et al. 2024). Third, the fact that inexperienced and experienced workers as well as males and females reacted differently in terms of application behavior and belief-updating to the provided information provides novel evidence on the heterogeneity of worker preferences and belief-updating in a real-word setting (Del Carpio and Guadalupe, 2022; Belot, Kircher and Muller, 2022).

While our results are robust as regards the number of applications for entry-level jobs, they do not provide answers on how firms can increase their applicant pool for highly-qualified experienced jobs, i.e., in cases where the overall pool of potential applicants is small and potential employees already hold sufficiently precise beliefs about a respective company. Our results only suggest that in this case an information treatment is much less effective. Future research may also provide better and more large-scale evidence on the impact of highlighting job amenities on the quality of the applicant pool, especially regarding the long-term performance of selected employees.

Technological advances will soon enable firms to target job advertisements not only to specific groups of individuals, but even to individual candidates. Our results suggest that the targeted assignment of job ads could be highly effective in attracting suitable candidates. Combining evidence from this paper with newly developed tools in the optimal treatment assignment literature (see, e.g., Kasy and Sautmann 2021; Opitz et al. 2024) could thus open up new perspectives for hiring strategies with substantial implications for labor-market search and matching.

2.9 Appendix A

2.9.1 Conceptual Framework

In this section, we present the formal model leading to the predictions stated in Section 2.3.

Preferences and Beliefs

Assume that potential applicants are characterized by (i) belonging to a group g of experienced workers denoted by E or inexperienced workers denoted by I, such that $g \in \{E, I\}$, and by (ii) having a fixed preference for job flexibility denoted by π_w^f and career advancement denoted by π_w^{ca}, where $w \in \{F, M\}$ denotes the gender. Additionally, each potential applicant has a job-specific ability denoted by α_i. We assume that workers decide between applying for a job at our target firm or an outside offer, the utility of which we denote by $\bar{U}_g$, and depends on previous work experience g, but is otherwise constant. The utility of a job at the target firm is a function of immediate wage returns denoted by m, returns to job-specific ability denoted by δ_g, and utility from job flexibility and from career-advancement opportunities:

$$U_{g,w,i} = m + \delta_g \alpha_i + \pi_w^f \tilde{\theta}_g^f + \pi_w^{ca} \tilde{\theta}_g^{ca}. \tag{2.2}$$

The job-specific ability, α_i, might arbitrarily correlate with workplace preferences for flexibility π_w^f and/or workplace preferences for career advancement π_w^{ca}. The utility component $\pi_w^f \tilde{\theta}_g^f$ formalizes that agents derive utility from workplace flexibility which is linear in their beliefs about flexibility in a particular job. We assume that $\pi_w^f \in [0, \infty)$, meaning that – all else equal – individuals prefer working under flexible working conditions, but are heterogeneous in this preference. Similarly, the utility component $\pi_w^{ca} \tilde{\theta}_g^{ca}$ describes an agent's utility from career advancement and shows a preference for career advancement of $\pi_w^{ca} \in [0, \infty)$.

Potential applicants are ex-ante uncertain about (i) the exact workplace flexibility and (ii) the career-advancement potential at the firm. Their priors for θ_g^f and θ_g^{ca} are normally distributed with $\tilde{\theta}_g^f \sim N\left(\bar{\theta}_g^f, {\tau_g^f}^{-1}\right)$ and $\tilde{\theta}_g^{ca} \sim N\left(\bar{\theta}_g^{ca}, {\tau_g^{ca}}^{-1}\right)$. Thus, before agents of group g obtain any additional information from the job ads, they have a prior $\tilde{\theta}_g^f$ with mean $\bar{\theta}_g^f$ and precision τ_g^f about the provided workplace flexibility and a prior

$\tilde{\theta}_g^{ca}$ with mean $\bar{\theta}_g^{ca}$ and precision τ_g^{ca} about the provided career growth. Additionally, applicants have a belief about the correlation between provided flexibility and career advancement. More formally, applicants have a common belief $\tilde{\rho}$ about the correlation coefficient of their priors, $\tilde{\theta}_g^f$ and $\tilde{\theta}_g^{ca}$. Moreover, let $\tilde{\theta}_E^f \perp \tilde{\theta}_I^f$ and $\tilde{\theta}_E^{ca} \perp \tilde{\theta}_I^{ca}$ hold.

For our further analysis, we make two assumptions.

Assumption 1. We assume that, on average, more experienced workers hold weakly more positive and strictly more precise ex-ante beliefs about the provided workplace flexibility and career growth at the job.

Formally, Assumption 1 translates into $\theta^f > \bar{\theta}_E^f \geq \bar{\theta}_I^f$, and $\theta^{ca} > \bar{\theta}_E^{ca} \geq \bar{\theta}_I^{ca}$ as well as $\tau_E^f > \tau_I^f$ and $\tau_E^{ca} > \tau_I^{ca}$ hold. The assumption that inexperienced workers have less accurate beliefs is motivated by the observation that more experienced workers have better networks (see, e.g., Glitz 2017) and are likely, overall, to be more informed about the labor market in their specific sector. This corresponds to assuming that they are better informed about the working conditions provided by the firm.

Secondly, we assume the following.

Assumption 2. We assume that female applicants have a higher preference for job flexibility than males and that male applicants have a higher preference for career growth than females.

Formally, Assumption 2 translates into $\pi_F^f > \pi_M^f$ and $\pi_M^{ca} > \pi_F^{ca}$ and is motivated by the findings of Wiswall and Zafar (2018).

The Effect of Highlighting Flexibility and Career Advancement in Job Ads

Before the job ad is posted, individuals know their job-specific ability α_i, their preferences for flexibility π_w^f, and career advancement π_w^{ca}. In expectation, their prior beliefs about flexibility amount to $\bar{\theta}_g^f$, and their beliefs about career-advancement opportunities amount to $\bar{\theta}_g^{ca}$.

The employer posts job ads that either (a) contain no information about flexibility or career advancement (*control treatment*) (b) contain information about flexible working conditions (*flexibility treatment*) or (c) contain information about potential career-advancement opportunities (*career treatment*). We interpret our treatments as information treatments, which serve as a positive signal to potential applicants and results in belief-updating of their priors regarding flexibility and career advancement provided by the firm. The signal s depends on the realization with $s \in \{s_c, s_f, s_{ca}\}$ while $s_f \sim N(\theta^f, \tau^{s_f-1})$ and $s_{ca} \sim N(\theta^{ca}, \tau^{s_{ca}-1})$. As the signal is positive, it holds that $\theta^f > \bar{\theta}^f_E \geq \bar{\theta}^f_I$ and $\theta^{ca} > \bar{\theta}^{ca}_E \geq \bar{\theta}^{ca}_I$.[38] The signal s_c is assumed to be completely uninformative.[39]

After observing the signal, we assume that applicants update their beliefs. Due to the normality assumption regarding the distributions, the posterior beliefs denoted by $\hat{\theta}$ are a weighted average of their priors and signals s_f, s_{ca}. In case applicants observe the signal s_f, their posteriors are given by:

$$\hat{\theta}^f_g(\tilde{\theta}^f_g, s_f) = \frac{\tilde{\theta}^f_g \tau^f_g + \tau^{s_f} s_f}{\tau^f_g + \tau^{s_f}} \tag{2.3}$$

$$\hat{\theta}^{ca}_g(\tilde{\theta}^f_g, \tilde{\theta}^{ca}_g, s_f) = \tilde{\theta}^{ca}_g + \tilde{\rho} \cdot \sqrt{\frac{\tau^f}{\tau^{ca}}} \cdot \frac{\tau^{s_f}(s_f - \tilde{\theta}^f_g)}{\tau^{s_f} + \tau^f} \tag{2.4}$$

In case applicants observe the signal s_{ca}, their posteriors are given by:

$$\hat{\theta}^{ca}_g(\tilde{\theta}^{ca}_g, s_{ca}) = \frac{\tilde{\theta}^{ca}_g \tau^{ca}_g + \tau^{s_{ca}} s_{ca}}{\tau^{ca}_g + \tau^{s_{ca}}} \tag{2.5}$$

$$\hat{\theta}^f_g(\tilde{\theta}^{ca}_g, \tilde{\theta}^f_g, s_{ca}) = \tilde{\theta}^f_g + \tilde{\rho} \cdot \sqrt{\frac{\tau^{ca}}{\tau^f}} \cdot \frac{\tau^{s_{ca}}(s_{ca} - \tilde{\theta}^{ca}_g)}{\tau^{s_{ca}} + \tau^{ca}} \tag{2.6}$$

Note that whether applicants use information provided via s_f to update their prior $\tilde{\theta}^{ca}_g$ and equally the information provided via s_{ca} to update their prior $\tilde{\theta}^f_g$ depends on their beliefs about potential trade-offs. In case $\tilde{\rho} = 0$, the right-hand side of 2.4 and 2.6 collapses to the respective prior beliefs. Since the *control treatment* does not contain information about flexibility or career growth, such job ads do not shift agents' priors.

Applicant i applies to the job if $U_{g,w,i} > \bar{U}_g$; thus, it is reasonable to assume that each increase of $U_{g,w,i}$ translates into a higher likelihood to apply. The average treatment effect of the *flexibility treatment* depending on group membership g and the belief about the trade-off $\tilde{\rho}$ can thus be described as $\Delta U|s_f(w, g, \tilde{\rho}) = E[U_{g,w} \mid s_f] - E[U_{g,w} \mid s_c] = E[U_{g,w} \mid s_f] - E[U_{g,w}]$, and the treatment effect of the *career treatment* can be described as $\Delta U|s_{ca}(w, g, \tilde{\rho}) = E[U_{g,w} \mid s_{ca}] - E[U_{g,w} \mid s_c] = E[U_{g,w} \mid s_{ca}] - E[U_{g,w}]$. We can explicitly

[38] We may interpret θ^f and θ^{ca} as the true level of flexibility and career-advancement opportunities provided by the firm.

[39] This only holds due to the exogenous nature of the signals.

formulate both expressions as

$$\Delta U|s_f(w,g,\tilde{\rho}) = \frac{\tau^{s_f}}{\tau_g^f + \tau^{s_f}}(\theta^f - \bar{\theta}_g^f) \cdot \left(\pi_w^f + \pi_w^{ca}\sqrt{\frac{\tau^f}{\tau^{ca}}}\tilde{\rho}\right) \tag{2.7}$$

$$\Delta U|s_{ca}(w,g,\tilde{\rho}) = \frac{\tau^{s_{ca}}}{\tau_g^{ca} + \tau^{s_{ca}}}(\theta^{ca} - \bar{\theta}_g^{ca}) \cdot \left(\pi_w^{ca} + \pi_w^f\sqrt{\frac{\tau^{ca}}{\tau^f}}\tilde{\rho}\right) \tag{2.8}$$

Given our previous discussion, we can now analyze the expected utility change in more detail. Considering 2.7 and 2.8, we observe that both expressions are positive if $\tilde{\rho}$ is not too small or more precisely, if $\tilde{\rho} > -\frac{\pi_w^f}{\pi_w^{ca}} \cdot \sqrt{\frac{\tau^{ca}}{\tau^f}}$ holds. Additionally, given our assumptions, $(\theta^f - \bar{\theta}_g^f)$ and $(\theta^{ca} - \bar{\theta}_g^{ca})$ are larger for $g = I$ than for $g = E$. Due to the assumed difference in prior precision, the same is true for $\frac{\tau^{s_f}}{\tau_g^f + \tau^{s_f}}$ and for $\frac{\tau^{s_{ca}}}{\tau_g^{ca} + \tau^{s_{ca}}}$. This leads to Proposition 1, which is the basis for prediction 1 and 2 in our conceptual framework in Section 2.3.

Proposition 1. *If $\tilde{\rho}$ is not too small, both treatments increase on average the total number of applications, and the increase is on average larger for applicants from group $g = I$.*

Considering 2.7 and 2.8 further, we observe that π_w^{ca} and π_w^f enter the expressions positively. Thus, the larger both are, the larger the total expected utility change is. Due to the assumed differences in gender preferences, it holds that $\pi_F^f > \pi_M^f$ and $\pi_M^{ca} > \pi_F^{ca}$, and thus the increases following the flexibility signal are expected to be larger for female applicants, while the expected increases following the career-advancement signal are expected to be larger for male applicants. This finding leads to Proposition 2 and serves as a basis for prediction 3 in our conceptual framework in Section 2.3.

Proposition 2. *It holds that $\Delta U|s_f(g,\tilde{\rho}) > \Delta U|s_{ca}(g,\tilde{\rho})$ for $w = F$, i.e., female applicants, and $\Delta U|s_f(g,\tilde{\rho}) < \Delta U|s_{ca}(g,\tilde{\rho})$ for $w = M$, i.e., male applicants.*

2.9.2 Robustness

In this section of the Appendix, we present several robustness checks with respect to our main analyses. First, we present a re-estimation of Equation 2.1 to estimate the treatment effects using a Poisson fixed-effects regression. Next, we provide several analyses providing evidence that spillover effects do not pose an identification threat to our empirical investigation in Section 2.5.1. Further, we present regression-based analyses providing alternative estimators for the treatment effects on the composition of the applicant pool in terms of region of residence and quality, which were descriptively analyzed in Section 2.5.4. Last, to rule out the possibility that the treatment effects are driven by age-related changes in preferences for flexibility and career advancement, we provide evidence that the treatment effects are not attributable to age differences in these workplace preferences.

Alternative Estimator

Table A.1 presents the results of a re-estimation of Equation 2.1 using a Poisson fixed-effects estimator. More precisely, we use a Pseudo-Poisson-ML estimator relying on the *ppmlhdfe* package in Stata (Correia, Guimarães and Zylkin, 2020). Columns 1 to 3 show the estimated treatment effects on the total number of applications to entry-level jobs, while Columns 4 to 6 show the estimated effects for jobs that require previous work experience. All estimations include job-ad and time fixed effects, and standard errors clustered on job-ad level.

The results are quite similar compared to the OLS fixed-effects regressions presented in Table 2.4 in Section 2.5.1. We begin to analyze the effect on the total number of applications for entry-level job ads. The point estimate of the *flexibility* treatment is 0.449 and highly significant. To interpret these coefficients, we consider the incidence ratio, which is the exponential of the coefficient, and gives the marginal estimated factor change of the mean of the dependent variable. For the *flexibility* treatment, this ratio is 1.57, which means that the estimated increase of applications amounts to 57%. Compared to the estimate in Section 2.5.1 of an increase of 44%, this estimate is quite similar in magnitude. For the *career* treatment, the point estimate is 0.333, corresponding to an incidence ratio of 1.40 and thus an estimated increase of 40%. Again, this estimate is quite similar in magnitude to the OLS estimate, amounting to an increase of 35%.

Equally, the results in Columns 2 and 3 are comparable to those presented in Table 2.2. In Column 2, which estimates treatment effects for male applicants only, we estimate an incidence ratio of 1.47 for the *flexibility* treatment, corresponding to an increase of 47% (compared to 37% from the OLS estimation). For the *career* treatment, we estimate an incidence ratio of 1.44, corresponding to an increase of 44% (compared to an estimate of 42% from the OLS regression). In Column 3, which presents the treatment effects for female applicants, we observe an incidence ratio of 2.02 for the *flexibility* treatment, amounting to an estimated increase of 102% (compared to an increase of 87% from the OLS regression). Equally to the results in Section 2.5.1, we can reject $\beta_f = \beta_{ca}$ for female applicants, but not for male applicants.

Table A.1: Treatment Effects on the Number of Applications – Poisson

	No. of applications - Poisson					
	Inexperienced			*Experienced*		
	All	Male	Female	All	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Flexibility	0.449***	0.388**	0.704**	-0.005	0.037	-0.205
	(0.147)	(0.162)	(0.315)	(0.161)	(0.182)	(0.203)
Career	0.333**	0.364**	0.095	0.031	0.034	-0.093
	(0.163)	(0.163)	(0.382)	(0.119)	(0.116)	(0.254)
Observations	827	827	569	1662	1610	908
Mean dep. variable	0.46	0.38	0.08	0.37	0.37	0.37
No. of Clusters	32	32	24	64	62	35
IRR Flexibility	1.57	1.47	2.02	0.99	1.04	0.81
IRR Career	1.40	1.44	1.10	1.03	1.04	0.91

Note: This table shows the impact of the treatments on the number of received applications per day. The estimates are obtained using a Poisson fixed-effect regressions; thus, to interpret these coefficients, we consider the incidence ratio, which is the exponential of the coefficient, and gives the marginal estimated factor change of the mean of the dependent variable. All specifications include job-ad and time fixed effects. Columns 1 to 3 present the effects for job ads requiring no previous work experience, while Columns 4 to 6 present the effects for job ads requiring previous work experience. Columns 1 and 4 show the effect for the total number of applications, Columns 2 and 5 only for the number of male applicants, and Columns 3 and 6 only for the number of female applicants. Standard errors clustered on job-ad level are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Spillover Analysis

In this section, we investigate potential spillover effects that may arise if interested applicants look at a job ad on several days, with a change in treatment in between. Such spillovers may arise as applicants are exposed to more than one treatment, and may lead to a downward bias in our main estimates. Moreover, it is conceivable that particularly effective or ineffective treatments have lasting effects beyond the ten-day period being shown on the job ad. To alleviate the risk of spillovers in the first place, we have excluded the day of the treatment switch and the day after in our main analysis in Section 2.5.1.

To investigate spillovers, we provide two types of analyses: i) We investigate time heterogeneities in treatment effects; and ii) We investigate the relevance of lagged treatments.

First, we present time heterogeneities in treatment effects with respect to the ten-day periods. More precisely, we re-estimate Equation 2.1 and include interaction terms for each ten-day period. The result is presented in Column 1 of Table A.2. We observe that the baseline point estimates of the treatment effects (measuring the effect for the first ten days) is slightly larger than the ones presented in the main part, while we observe noisy point estimates of the time-interaction effects. Overall, we conclude that there is no evidence for strong time trends in the estimated treatment effect. This evidence speaks against the existence of spillover effects, as spillovers should be more likely to occur over time leading to changes in treatment effects over time.

In Columns 2 to 4, we re-estimate Equation 2.1 including the first lag of the treatment. In Column 2, we include only the lag for the *flexibility* treatment, and in Column 3 only the lag for the *career* treatment; in Column 4, we include both. Including lags allows us to investigate whether a particular treatment is predictive of the number of applications beyond the ten-day period. It also allows us to see if including lags changes the magnitude of the estimates of our main treatment effects. From Column 2 to 4, we observe that the point estimates are relatively stable and in size all very close to the estimation in Table 2.4. We only note a weakly significant point estimate of the *flexibility* lag in Column 2. Overall, this provides strong evidence that spillovers are not relevant for the estimated size or significance of our main treatment effects.

Table A.2: Robustness – Time Heterogeneity and Lags

	No. of applications - OLS			
	(1)	(2)	(3)	(4)
Flexibility	0.227	0.223***	0.173**	0.217**
	(0.161)	(0.075)	(0.067)	(0.080)
Career	0.221	0.137*	0.110	0.123
	(0.169)	(0.077)	(0.089)	(0.088)
Flexibility×Day 11-20	-0.166			
	(0.194)			
Flexibility×Day 21-30	0.003			
	(0.273)			
Career×Day 11-20	-0.069			
	(0.242)			
Career×Day 21-30	-0.169			
	(0.203)			
Lag1 Flexibility		0.141*		0.122
		(0.079)		(0.081)
Lag1 Career			-0.098	-0.049
			(0.080)	(0.081)
Observations	829	829	829	829
No. of Clusters	32	32	32	32
Mean dep. variable	0.46	0.46	0.46	0.46

Note: This table shows the impact of the treatments on the number of received applications per day. The estimates are obtained using standard OLS fixed-effect regressions; thus, the marginal effects need to be interpreted in terms of change in the number of applications per day. All specifications include job-ad and time fixed-effects. Column 1 includes interactions of the treatment dummies with time-period dummies. More precisely, we interact each treatment dummy with a dummy being equal to one for treatment days 11 to 20, and one being equal to one for treatment days 21 to 30. Column 2 includes the first lag for the *flexibility* treatment, Column 3 includes it for the *career* treatment, and Column 4 includes both. These dummies are equal to one in case in the period before the current treatment period, either the *flexibility* or the *career* treatment was online. Standard errors clustered on job-ad level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Further Results

In this section, we provide a regression-based analysis of the mean differences analyzed descriptively in Section 2.5.4. In particular, we re-estimate Equation 2.1 and use either the total number of applications of each place-of-residence category as an outcome variable (these results are presented in Table A.3), or the total number of each recruiter-rating category as an outcome variable (these results are presented in Table A.4).

Table A.3: Treatment Effects by Category – Region of Residence

	Region of residence of applicants - OLS		
	Germany w/o state	State	Abroad
	(1)	(2)	(3)
Flexibility	0.121**	0.020	0.033
	(0.047)	(0.031)	(0.042)
Career	0.125**	0.034	-0.025
	(0.050)	(0.039)	(0.026)
Observations	829	829	829
No. of Clusters	32	32	32
Mean dep. variable	0.21	0.13	0.12

Note: This table shows the impact of the treatments on the number of received applications per day by region of residence of the applicants. The outcome variable of Column 1 is the number of daily applicants who live in Germany, but not in the federal state of the firm (Germany w/o state). The outcome variable of Column 2 is the daily applicants living in the federal state of the firm (State). The outcome variable of Column 3 is the daily number of applicants living abroad (Abroad). The estimates are obtained using standard OLS fixed-effect regressions; thus, the marginal effects need to be interpreted in terms of change in the number of applications per day. All specifications include job-ad and time fixed effects. Standard errors clustered on job-ad level are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Treatment Effects by Category – Quality

	Quality of applicants - OLS				
	A	B	C	Screened out	Interview
	(1)	(2)	(3)	(4)	(5)
Flexibility	0.018	0.035	0.045	0.073	0.039
	(0.014)	(0.027)	(0.038)	(0.053)	(0.027)
Career	0.035*	0.027	0.030	0.045	0.071*
	(0.018)	(0.018)	(0.026)	(0.060)	(0.038)
Observations	829	829	829	829	829
No. of Clusters	32	32	32	32	32
Mean dep. variable	0.06	0.08	0.12	0.20	0.12

Note: This table shows the impact of the treatments on the number of received applications per day depending on the quality of the applicant measured by means of recruiter ratings. The outcome variable of Column 1 is the number of daily applicants with an *A* rating. The outcome variable of Column 2 is the number of daily applicants with a *B* rating. The outcome variable of Column 3 is the daily number of applicants with a *C* rating. The outcome variable of Column 4 is the number of daily applicants who were screened out and the outcome variable of Column 5 of applicants being invited to an interview. The estimates are obtained using standard OLS fixed-effect regressions; thus, the marginal effects need to be interpreted in terms of change in the number of applications per day. All specifications include job-ad and time fixed effects. Standard errors clustered on job-ad level are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Age

This section examines how preferences for workplace flexibility and career changes vary across age. Figure A.1 graphically illustrates the average marginal effects of age on preferences for *flexibility* and *career*. Participants rated their preferences for attributes on a scale from 1 (does not apply at all) to 10 (fully applies). We created two composite scores to capture preferences independent of specific job contexts. *Flexibility* aggregates ratings for flexible work schedule, work-life balance, home-office, childcare, family-friendly environment, and overtime avoidance; while *career* combines ratings for high income, career opportunities, salary growth, and the possibility of negotiating salary increases. These composite scores mirror the items used to categorize beliefs about job characteristics in Section 2.6.2. Both scores are calculated as the standardized sum of the ratings for each item within the category.

In Figure A.1 we observe a relatively constant, and if anything, slightly increasing effect of age on having a preference for flexibility (A). Linear age trends yield coefficients of 0.001, for both, *flexibility* and *career*, with p-value = 0.289 and 0.407, respectively. These results indicate that preferences for both flexibility and career opportunities may, if anything, slightly increase with age.

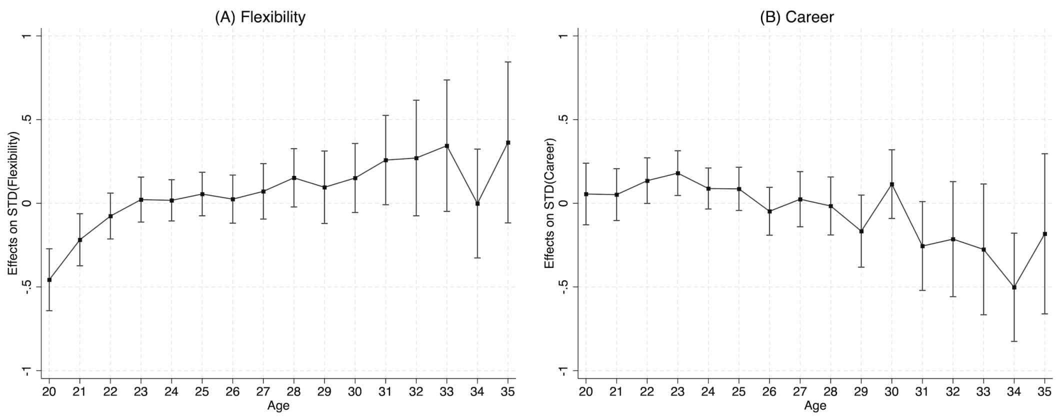

Figure A.1: Predictive Margins of Age on Preferences

Note: The figure displays average marginal effects on the preference for *flexibility* (A) and *career* (B) for individuals aged 20 to 35. To generate these results, we conducted a regression analysis of preferences for *flexibility* and *career* on age dummies. We then estimated predicted values for the respective outcomes across the observed age range from 20 to 35, in one-year increments. Both outcomes are standardized.

2.9.3 Survey Experiment

This section presents additional material of our survey experiment. Table A.5 provides descriptions of the control variables that are used in our survey experiment. Table A.6 presents the different labs and the corresponding participant numbers, while Table A.7 presents summary statistics of the data collected. Table A.8 presents the results for each single item contributing to the composite score for the beliefs about job characteristics (Section 2.6.2), while Table A.9 presents the results for each single item of the composite scores of the beliefs about the working environment (Section 2.6.3). Last, Table A.10 presents the treatment effects for items not used in our analysis in Section 2.6.

Table A.5: Variable Definitions

Variable	**Description**
Female	Dummy that equals 1 if the individual is female, 0 else
Migration background	Dummy that equals 1 if at least one parent is born outside of Germany, 0 else
University degree	Dummy that equals 1 if the individual is enrolled in a bachelor's program or has at least a bachelor's degree
Family status	Dummy that equals 1 if the individual has at least one child, 0 else

Table A.6: Survey – Laboratory and Treatment

Laboratory	Control	Flexibility	Career	Total Participants
RWTH Aachen	112	112	107	331
FU Berlin	161	166	160	487
University of Bonn	50	51	53	154
HHU Düsseldorf	8	9	8	25
University of Göttingen	2	3	2	7
University of Hannover	39	38	37	114
University of Heidelberg	14	14	13	41
University of Innsbruck	15	14	15	44
University of Cologne	98	97	95	290
KIT Karlsruhe	49	60	52	161
LMU Munich	79	79	82	240
TUM Munich	79	80	83	242
Total	706	723	707	2,136

Note: This table shows the number of participants in our survey by laboratory and treatment.

Table A.7: Summary Statistics by Treatment

Variable	Control Mean	Control SD	Flexibility Mean	Flexibility SD	Career Mean	Career SD
A. Background variables						
Female	0.421	0.494	0.375	0.484	0.380	0.486
University degree	0.609	0.488	0.527	0.500	0.556	0.497
Family status	0.038	0.192	0.043	0.203	0.031	0.174
Migration background	0.458	0.499	0.402	0.491	0.451	0.498
B. Beliefs about job characteristics						
Work-life balance	36.297	8.354	37.170	8.732	35.317	8.727
Career benefits	25.565	6.141	25.419	6.305	26.594	5.760
C. Beliefs about working environment						
Friendly	80.414	23.883	82.274	24.445	80.987	22.968
Competitive	130.561	45.175	128.640	43.424	134.330	44.403
Observations	706		723		707	

Note: This table presents summary statistics categorized by treatment status. Panel A provides an overview of background variables. Panel B focuses on our two indicators characterizing beliefs about job characteristics: work-life balance and career benefits. Work-life balance adds up ratings about expected flexibility, work-life balance, home-office, childcare support, projected overtime, and family-friendly workplace culture. Career benefits adds up ratings about expected salary, provided salary growth, career opportunities, how challenging the tasks of the jobs are, and the possibility of negotiating salary increases. Panel C provides a summary of beliefs about the working environment, including whether it is either friendly or competitive. Friendly working environment adds up beliefs about the share of colleagues who are female and have a family. Competitive working environment adds up survey questions about beliefs about the share of colleagues who prioritize their career over having a family, who are eager to have a career, who have a STEM degree, and earn a high income.

Table A.8: Belief-Updating about Job Characteristics

	Beliefs										
	Work-life balance						*Career benefits*				
	Flexibility	Work-life balance	Home-office	Childcare	Family	Overtime	Salary	Career	Salary growth	Chall-enge	Negot-iation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Flexibility	0.305**	0.206***	0.225*	0.133	0.108	0.164	-0.026	-0.038	0.016	0.023	-0.033
	(0.110)	(0.069)	(0.125)	(0.111)	(0.117)	(0.107)	(0.080)	(0.092)	(0.105)	(0.109)	(0.101)
Career	-0.038	-0.229**	-0.021	-0.155	-0.131	-0.232*	0.050	0.233*	0.601***	0.088	0.173**
	(0.107)	(0.099)	(0.131)	(0.111)	(0.105)	(0.130)	(0.083)	(0.123)	(0.088)	(0.137)	(0.080)
Observations	2136	2136	2136	2136	2136	2136	2136	2136	2136	2136	2136
No. Clusters	20	20	20	20	20	20	20	20	20	20	20
Job FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lab FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bootstrap p β_f	0.02	0.00	0.08	0.29	0.37	0.14	0.71	0.70	0.90	0.83	0.76
Bootstrap p β_{ca}	0.71	0.01	0.87	0.16	0.22	0.09	0.53	0.08	0.00	0.57	0.04

Note: The table shows the impact of the treatments on the beliefs about the individual items describing job characteristics: expected flexibility, work-life balance, home-office, childcare support, family-friendly workplace culture, projected overtime, expected salary, career opportunities, salary growth, and the possibility to negotiate salary increases. Controls include gender, high school GPA, migration background, university degree, and family status. Standard errors are clustered on job-ad level and are reported in parentheses. The last two rows show the p-values from wild bootstrapped clustered standard errors (Cameron, Gelbach and Miller, 2008).

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: Belief-Updating about Work Environment

	Beliefs					
	Friendly		*Competitive*			
	Female (1)	Family (2)	Income (3)	Ambitious (4)	Career (5)	STEM (6)
Flexibility	0.874	1.255	-0.704	-0.148	-1.019	-0.860
	(0.575)	(0.910)	(1.104)	(1.251)	(1.156)	(0.889)
Career	0.881	-0.037	1.330	2.161**	0.555	0.610
	(0.831)	(1.018)	(1.315)	(0.932)	(1.318)	(1.166)
Observations	2136	2136	2136	2136	2136	2136
No. Clusters	20	20	20	20	20	20
Job FE	Yes	Yes	Yes	Yes	Yes	Yes
Lab FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bootstrap p β_f	0.16	0.16	0.56	0.89	0.39	0.34
Bootstrap p β_{ca}	0.34	0.93	0.32	0.04	0.72	0.62

Note: This table shows the impact of the treatments on the beliefs about the individual items describing the working environment: the share of colleagues being female, having a family, the share of colleagues prioritizing career over family, who are eager to have a career (ambitious), have a STEM degree, and earn a high income. Controls include gender, high school GPA, migration background, university degree, and family status. Standard errors clustered on job-ad level are reported in parentheses. The last two rows show the p-values from wild bootstrapped clustered standard errors (Cameron, Gelbach and Miller, 2008).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.10: Distractor Items

	Part-time (1)	Travel (2)	Location (3)	Security (4)	Reputation (5)	Old (6)
Flexibility	-0.076	-0.261	0.016	-0.104	-0.053	0.091
	(0.141)	(0.162)	(0.103)	(0.123)	(0.124)	(1.064)
Career	-0.212	-0.198	-0.104	-0.136	0.087	-1.376
	(0.191)	(0.244)	(0.146)	(0.099)	(0.152)	(1.096)
Observations	2136	2136	2136	2136	2136	2136
No. Clusters	20	20	20	20	20	20
Job FE	Yes	Yes	Yes	Yes	Yes	Yes
Lab FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bootstrap p β_f	0.55	0.13	0.92	0.42	0.65	0.93
Bootstrap p β_{ca}	0.32	0.47	0.50	0.17	0.59	0.25

Note: This table illustrates the impact of the treatments on the individual items excluded from our indicators: opportunity to work part-time, travel requirements for the job, attractive work location, secure workplace, reputation of the employer, and share of old employees. Controls include gender, high school GPA, migration background, university degree, and family status. Standard errors clustered on job-ad level are reported in parentheses. The last two rows show the p-values from wild bootstrapped clustered standard errors (Cameron, Gelbach and Miller, 2008).

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3 Relative Grades and Gender Differences in STEM Enrollment

joint with Pia Pinger and Philipp Seegers*

Abstract

Based on novel administrative and survey data from Germany, this study investigates the importance of relative STEM performance in high school for the gender gap in STEM enrollment. We first document that males display a higher relative STEM performance than females, which however mainly emerges from females' stronger achievement in non-STEM subjects. Our findings further reveal that a one-standard-deviation increase in grade-based STEM advantage raises the likelihood of pursuing a STEM degree by approximately 19 percentage points for males, but only by half as much for females. A decomposition analysis shows that 26% of the STEM gender gap could be attributed to differences in grade-based STEM performance if major preferences resembled those of males. However, relative grades are largely unimportant in an environment where preferences mirror those of females. This suggests that STEM performance differences have limited influence on females' decisions to pursue STEM degrees. While STEM advantage significantly impacts observed gender gaps in STEM enrollment, this effect is primarily driven by males.

Keywords: Gender Gap, STEM Enrollment, Relative Grades, Ranks
JEL classification: I21, I24, J16, J24

*We would like to express our gratitude to the "Qualitäts- und UnterstützungsAgentur – Landesinstitut für Schule (QUA-LiS NRW)" for their support throughout this process and for providing the necessary data. Financial support through the German Research Foundation Collaborative Research Centre TR 224 (Project A02) is gratefully acknowledged. Pia Pinger thankfully acknowledges support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1-390838866.

3.1 Introduction

Women are underrepresented in math-intensive fields. In Germany, only 22% of graduates in science, technology, engineering, and mathematics (STEM) are female, compared to 32% in the OECD overall. Moreover, the gender gap in STEM attendance has not decreased in most developed countries over the past decades (OECD 2024). The underrepresentation of women in STEM is concerning to the extent that STEM graduates tend to earn high wages and have above-average career prospects (e.g., Anger and Plünnecke 2022; Blau and Kahn 2017). Moreover, females shying away from STEM-related fields may limit the talent pool in occupations that are often viewed as key contributors to a country's growth and national competitiveness (Carnevale, Smith and Melton 2011; Bianchi and Giorcelli 2020; Del Carpio and Guadalupe 2022).

Why is it that women are so much less likely to choose a math-intensive field of study? The underlying reasons are manifold, ranging from a difference in preferences and expectations (Zafar 2011, 2013; Wiswall and Zafar 2021; Niederle and Vesterlund 2010), via norms (Guiso et al. 2008; Nosek et al. 2009; Nollenberger, Rodríguez-Planas and Sevilla 2016; Del Carpio and Guadalupe 2022; Carlana 2019; Terrier 2020; Nicoletti, Sevilla and Tonei 2022), to a lack of female role models (Breda et al. 2023; Bettinger and Long 2005; Winters et al. 2013; Dee 2005, 2007; Canaan and Mouganie 2023), or peer effects (Murphy and Weinhardt 2020; Elsner, Isphording and Zölitz 2021; Elsner and Isphording 2017; Duflo, Dupas and Kremer 2011). PISA results reveal persistent gender differences in academic performance across OECD countries (OECD 2019*a*,*b*). First, boys tend to do slightly better in STEM subjects on average (1.4% higher scores), although this difference does not suffice to explain prevailing gender disparities in math-intensive professions (Ceci et al. 2014). Second, girls outperform boys in terms of verbal abilities (6% higher scores), and they tend to outperform boys in school, overall.

As a consequence, a girl that performs well in STEM is likely to perform even better in non-STEM subjects. She thus receives more positive signals about her non-STEM abilities than about her STEM abilities compared to a boy, and may conclude that a non-STEM occupation suits her abilities best. To the extent that such perceptions lead to disparities in enrollment decisions, this could explain persistent gender differences in STEM choices and human-capital investments. Recent evidence (Breda and Napp 2019; Goulas, Griselda and Megalokonomou 2022) suggests that relative performance in STEM versus non-STEM can indeed have important implications for female intentions to continue an education in STEM and that relatively stricter grading policies in STEM courses might reinforce this tendency (Ahn et al. 2024). However, there remains a gap in our understanding of how these relative performance differences and their interpretation affect actual decision-making, beyond mere intentions. Our study addresses this gap by investigating three key questions: First, how does relative performance in STEM versus non-STEM subjects, measured by grades or ranks at the end of high school, influence the decision to pursue STEM-related subjects in higher education? Second, to what extent do males and females differ in their assessment of these relative performance indicators? At an aggregate level, how much of the gender gap in STEM enrollment at university can be attributed to differences in relative performance?

To answer these questions, we rely on two sources of data from Germany. First, administrative data documenting grade distributions from upper secondary education, including both overall performance and subject-specific achievement in the final exit exams. Second, survey data that contains information on background, university enrollment and performance, high school grade point averages (GPAs) of exit exams, and subjects chosen in high school exit exams. Based on these data, we construct two measures of relative STEM performance following Goulas, Griselda and Megalokonomou (2022). First, *grade-based STEM advantage* is calculated as the ratio of STEM over non-STEM GPA achieved in final exit exams, minus one. A grade-based STEM advantage greater than zero indicates that an individual has a higher GPA in STEM subjects than in non-STEM subjects, reflecting a relative proficiency in STEM based on grades. Second, *rank-based STEM advantage* is computed as the ratio of the school-cohort rank of STEM GPA to the school-cohort rank of non-STEM GPA, minus one. This describes an individual's relative grade position as compared to her classmates, that is, considering the individual's position within the school and the year based on their grades. Our analysis is based on a sample of 573 observations that allow us to link these measures of relative STEM advantage of upper secondary education school leavers to enrollment choices in tertiary education. We also conduct a decomposition analysis to quantify the extent to which gender differences in STEM enrollment can be attributed to variations in grade- and rank-based performance indicators.

Germany offers a perfect setting to study the effect of ability signals on human-capital investment and selection. First, Germany offers an educational landscape where over 90% of institutions, including schools and universities, are publicly funded and tuition-free. Public schools maintain exceptionally high quality, with private institutions holding no marked advantage. Consequently, financial constraints exert little influence over educational choices. Second, uniform compensation schemes for teachers and standardized curricula yield consistent educational quality across schools of a particular type. Variations in schooling levels and educational intensity primarily emanate from school tracking, which is transparent to students, parents, educators, and researchers alike. Lastly, gender-based disparities in tertiary-education outcomes are particularly persistent in Germany (OECD 2024), which may reflect substantial non-monetary variations in educational decision-making across groups.

We present two sets of results. First, we provide descriptive evidence of a significant gender gap in high school grades between STEM and non-STEM subjects. Females exhibit smaller STEM to non-STEM grade differences compared to males, which we refer to as "grade-based STEM advantage". This advantage stems from females achieving comparable grades in STEM subjects while outperforming in non-STEM subjects, aligning with existing literature (OECD 2019*b*; Breda and Napp 2019; Goulas, Griselda and Megalokonomou 2022). In our sample, we also identify a 24% gender gap in STEM enrollment in higher education programs. Our analytical findings reveal that grade-based STEM advantage increases the likelihood of choosing a STEM subject for both genders, but with a notably smaller effect for females. A one-standard-deviation increase in grade-based STEM advantage raises the probability of pursuing a STEM degree by 19 percentage points for males, but only half as much for females. A one-standard-deviation increase in rank-based STEM advantage raises the probability of pursuing a STEM degree by 4.2 percentage points for males, but there is no effect for females. Second, a decomposition of the STEM enrollment gap into relative-STEM-performance-related differences and differences in preferences reveals that if female major preferences resembled those of males, 26% of the gender gap in STEM enrollment could be attributed to disparities in grade-based performance indicators. However, rank-based performance indicators do not significantly affect the gender gap in STEM choices. In a scenario with female choice preferences, neither grade-based nor rank-based STEM performance differentials significantly influence STEM enrollment differences. We find that males are more likely than females to specialize in STEM fields if they have a relative advantage in STEM-related subjects, whether based on grades or ranks. This suggests that non-performance-related factors, such as preferences or anticipated discrimination, may discourage females from choosing STEM occupations despite positive ability signals and we provide suggestive evidence that females enrolled in STEM subjects indeed anticipate more gender-based discrimination in their future careers.

Our study contributes to the existing literature in at least four ways. First, we extend research on STEM advantage in educational decisions by examining actual choices rather than intentions (Breda and Napp 2019; Goulas, Griselda and Megalokonomou 2022). Second, we provide the first analysis of grade- and rank-based performance indicators' relative importance across genders in Germany, a setting, where grades are crucial for university enrollment.[1] Third, there is evidence that shows that students have imperfect knowledge of their own ability (Zafar 2011; Stinebrickner and Stinebrickner 2012, 2014; Bobba and Frisancho 2016) and are uncertain about their returns to education (Jensen 2010; Attanasio and Kaufmann 2014; Wiswall and Zafar 2015). We are able to show that in their education decisions, female students seemingly place too little weight on their relative advantage. In our decomposition exercise, we are able to delineate effects that stem from (gender) preferences to those from performance differences. This approach enables us to quantify the contribution of observed performance differences in STEM and non-STEM fields on the overall gender STEM-enrollment gap. Extending the literature (e.g., Delaney and Devereux 2019; Card and Payne 2021; Riegle-Crumb et al. 2012), we are able to add an explanation on the paradox of women selecting lower-wage non-STEM fields despite demonstrating equal or superior academic performance across disciplines. Lastly, we extend research on ability cues in decision-making (Stinebrickner and Stinebrickner 2012; Murphy and Weinhardt 2020; Elsner, Isphording and Zölitz 2021; Bond et al. 2018; Li and Xia 2024; Tan 2023) and gender differences in grade responsiveness. Prior work shows females' persistence in subjects correlates with strong performance (Owen 2010), yet they exit male-dominated and STEM fields more readily after poor performance than males do (Kugler, Tinsley and Ukhaneva 2021; Rask and Tiefenthaler 2008). While existing studies examine absolute grades in individual subjects, we demonstrate that females respond less than males to relative performance differences across subjects. Overall, women may thus require stronger signals than males to decide for a career in STEM.

The remainder of the paper is organized as follows. In the next section, we provide information on the institutional setting, the data, measures, and descriptive statistics. In Section 3.3, we present the main results. Section 3.4 concludes.

[1] Access to tertiary education is determined by the acquisition of the high school degree only. Admission restrictions in competitive fields such as business and administration, psychology, or medicine, are generally determined by the final high school GPA.

3.2 Institutional Setting, Data, and Descriptive Statistics

3.2.1 The German School System

The German education system distinguishes itself by assigning the responsibility for education to each federal state. In this study, we investigate GPAs in central examinations from high schools situated within the federal state of North Rhine-Westphalia (NRW), using data of high school leavers between 2010 to 2019. These central exams make up an important fraction of the upper secondary degree GPA. It opens doors for future education and career paths by determining eligibility for tertiary education.

The exams are centrally provided by the federal state of NRW, aiming to enhance comparability and to ensure equitable treatment for all students. The grade information we observe in our sample stems from standardized exams across all upper secondary schools. In the final examination, students select four subjects, consisting of three written exams and one oral exam. For grading consistency, we focus exclusively on the GPAs obtained from written exams. The GPAs range from 0 to 15 points, with 15 denoting the highest grade and 0 the lowest. The minimum passing grade is 4 points. The final high school GPA, computed from these points, then ranges from 1.0 to 4.0, where 1.0 is the best grade and 4.0 the lowest one.[2] In 2005, there was a shift in the educational system from the G9 to the G8 system. The G8 system reduces time spent at at school from 13 years to 12 years. Since the first cohort participating in the G8 system finished upper secondary school in 2012, we need to account for graduation-year effects. The institutional background is presented in more detail in Section 3.5.1 of Appendix B.

3.2.2 Data

Our dataset combines survey data from the German student study "Fachkraft 2013" with administrative records of GPA distributions from NRW. The survey, conducted in March 2021 and March 2022, collected comprehensive information about students' background, university enrollment, performance, and for a subsample, detailed high school information including course selection, grades, and IQ scores. Students were recruited through a major nationwide job board platform, with participation incentivized through Amazon vouchers.[3] The sample closely compares to the overall population of German students in terms of region, university type, study fields, and likelihood to hold

[2]See APO-GOSt in the version of 12 March 2009 [Article 1, Paragraph 20].

[3]The job board jobmensa.de is operated by Studitemps GmbH (jobvalley) and is the largest platform for student jobs. Participation was incentivized using Amazon vouchers amounting to 1,950 EUR (29 x 50, 1 x 500 vouchers).

a student job (Hemkes et al. 2016). The administrative data, obtained from the Qualitäts- und UnterstützungsAgentur – Landesinstitut für Schule (QUA-LiS NRW), covers GPA distributions from central final exams and school-leaving grades for all upper secondary schools in NRW from 2010-2019. These records include school characteristics (legal status and type) and GPA frequency distributions across 22 subjects[4], ranging from STEM fields like mathematics and physics to humanities and arts.

Merging these administrative records with our survey data yields a final sample of 573 individuals with information on high school performance and major choices in higher education.

3.2.3 Measures

Tertiary Education Sorting Students in Germany directly enroll for a field of study when they first enter university. We elicited the current study field as a choice out of a list of 14 majors.[5] We adopt a STEM definition that emphasizes strong quantitative rigor. For the purposes of classifying tertiary education choices, we consider the following disciplines as STEM-related: computer sciences, engineering, mathematics, chemistry, and physics.

Secondary Education and Grades We identify five subjects as STEM subjects in high school – computer sciences, mathematics, physics, and chemistry – in alignment with the STEM classification used for categorizing tertiary education choices. Since it is compulsory to take at least one subject from a STEM field, we are able to observe STEM and non-STEM GPAs for all of our sample. For simplicity, we reversed the order in our analysis such that a higher always GPA indicates better grades. To assess individual competence in STEM relative to non-STEM subjects, we follow Goulas, Griselda and Megalokonomou (2022). Our first measure of relative performance, *grade-based STEM advantage*, is based on grades and constructed for each student i in the following way:

$$\textit{Grade-based STEM advantage}_i = \frac{\textit{STEM GPA}_i}{\textit{Non-STEM GPA}_i} - 1 \tag{3.1}$$

[4]Specifically, the subjects include math, chemistry, physics, computer science, technology, German, English, French, Dutch, biology, history, geology, social sciences, Chinese, educational science, art, Latin, music, Spanish, sport, psychology, and business administration.

[5]Majors comprise educational sciences, computer sciences, engineering, art, music, mathematics, media sciences, medicine, health sciences, natural sciences, psychology, legal sciences, social sciences, humanities, sports science, linguistics, cultural studies, and economics.

A grade-based STEM advantage exceeding 0 indicates that individual *i* has achieved a higher GPA in STEM subjects than in non-STEM subjects, signifying a relative proficiency in STEM based on grades. A negative value would be interpreted inversely.

To construct our second measure of relative performance, *rank-based STEM advantage*, we need to construct two separate rank measures – one based on STEM GPA and another based on non-STEM GPA – to capture students' relative standing within each domain.. Since school cohorts and classes vary in size, we do not use the raw rank of students in each subject s in their school cohort c but transform the rank position (n_{ijsc}) into a local percentile rank (R_{ijsc}) to make it comparable across schools j, following Murphy and Weinhardt (2020).

$$R_{ijsc} = \frac{n_{ijsc} - 1}{N_{jsc} - 1} \times 100 \tag{3.2}$$

where N_{jsc} is the cohort size of school j in cohort c of subject s. We multiply this measure by 100, resulting in a rank scale from 0 to 100, where the lowest-ranked student in each school cohort has R = 0 and the highest-ranked student has R = 100. In the case of ties, both students are assigned the lower rank.

Rank-based STEM advantage is defined for each student i in the following way:

$$\textit{Rank-based STEM advantage}_i = \frac{R_{ijsc} \textit{ of STEM GPA}}{R_{ijsc} \textit{ of non-STEM GPA}} - 1 \tag{3.3}$$

where R_{ijsc} are the local percentile ranks we compute in Equation 3.2 where $s \in \{STEM, \textit{non-STEM}\}$.

3.2.4 Descriptive Statistics

Our final sample is drawn from North Rhine-Westphalia's student population. NRW, being the largest federal state in Germany and comparable in size to the Netherlands, offers a rich context for examining key characteristics of German pupils, including their university preferences, fields of study, and regional distribution. For a more detailed explanation of the variables, please refer to Section 3.5.2 of Appendix B.

Panel A of Table 3.1 shows that female students perform better than male students as regards their overall high school GPA. Moreover, while females display a slightly lower performance in STEM subjects (0.06 sd), they achieve 0.13 sd better grades in non-STEM subjects compared to males. Gender differences in high school GPAs are statistically

Table 3.1: Summary Statistics

	Females		*Males*		*Females - Males*		
Variable	(1) Mean	(2) SD	(3) Mean	(4) SD	(5) Norm.Δ	(6) Abs.Δ	(7) p-val
A. Performance in high school							
High school GPA	2.713	0.622	2.634	0.588	0.09	0.08	0.12
STEM GPA	9.204	3.239	9.470	3.259	-0.06	-0.27	0.33
Non-STEM GPA	11.015	2.218	10.594	2.262	0.13	0.42	0.03
B. Constructed variables in high school							
Grade-based STEM advantage	-0.140	0.309	-0.071	0.346	-0.15	-0.07	0.01
Rank-based STEM advantage	0.067	1.346	0.169	1.900	-0.04	-0.10	0.45
Rank STEM GPA	58.304	27.525	59.991	27.655	-0.04	-1.69	0.47
Rank non-STEM GPA	68.298	24.024	66.056	24.477	0.07	2.24	0.28
C. Background variables							
High school GPA (cohort)	2.561	0.180	2.533	0.172	0.11	0.03	0.07
STEM GPA (cohort)	7.954	1.505	7.967	1.629	-0.01	-0.01	0.92
Non-STEM GPA (cohort)	8.805	1.060	8.472	1.062	0.22	0.33	0.00
Low SES	0.292	0.455	0.474	0.500	-0.27	-0.18	0.00
Migration status	0.088	0.284	0.094	0.292	-0.01	-0.01	0.82
IQ	2.310	1.691	1.782	1.896	0.21	0.53	0.00
D. Tertiary education							
STEM degree	0.198	0.399	0.435	0.497	-0.37	-0.24	0.00
Law degree	0.041	0.199	0.039	0.194	0.01	0.00	0.88
Economics and Business degree	0.109	0.312	0.220	0.415	-0.21	-0.11	0.00
Humanities and Social Sciences degree	0.469	0.500	0.228	0.421	0.37	0.24	0.00
Health degree	0.183	0.387	0.078	0.268	0.22	0.11	0.00
University GPA	2.859	0.592	2.700	0.528	0.20	0.16	0.00
University STEM GPA	2.674	0.565	2.532	0.510	0.19	0.14	0.11
University non-STEM GPA	2.899	0.592	2.805	0.514	0.12	0.09	0.11

Note: This table reports statistics of variables by gender for a set of 573 observations. Columns 1 and 3 show the mean for each group, while Columns 2 and 4 present the standard deviation (sd). Column 5 reports normalized differences between females and males (Imbens and Wooldridge 2009). Normalized differences are calculated as averages by group status scaled by the square root of the sum of the variances. Column 6 presents the absolute differences, while Column 7 provides the p-values from a two-sided t-test for comparing means.

significant for non-STEM subjects ($p<.05$) but not for STEM subjects. The probability density distributions in Figure 3.1 confirm this pattern, showing significant gender disparities in distributions of non-STEM GPAs ($p<.05$) but smaller, non-significant differences in STEM fields.

Panel B of Table 3.1 presents gender-based disparities in scholarly achievement of our constructed metrics. The outcomes reveal a significant male advantage in STEM subjects relative to non-STEM subjects, measured by grade-based STEM advantage. Males have a 0.15 sd higher STEM advantage ($p<.01$) based on grades. This finding is substantiated by the empirical evidence presented in Figure 3.2. Importantly, the higher grade-based STEM advantage for males is mainly driven by *worse performance in non-STEM GPAs of males compared to females*. Furthermore, we find no significant difference in rank-based STEM advantage. As shown in Figure 3.1, the variation in grade- versus rank-based

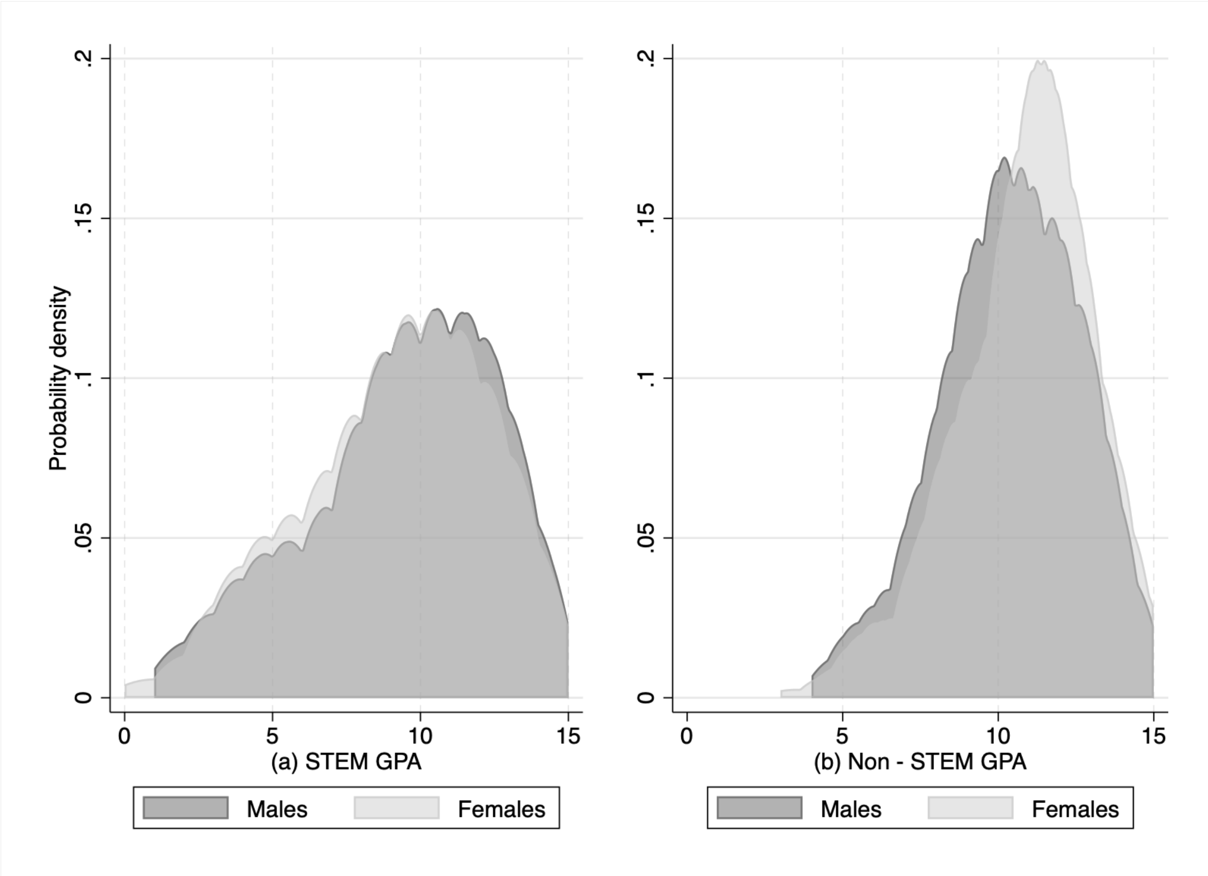

Figure 3.1: Kernel Densities of STEM and Non-STEM Performance

Note: The figure displays kernel density plots of the number of points achieved in STEM and non-STEM subjects in the final exams of upper secondary school for males and females, respectively. Kernel = Epanechnikov; a two-sample Kolmogorov-Smirnov test indicates a significant difference between distributions of non-STEM GPA by gender ($p<.05$), rejecting the null hypothesis of equal distributions; the test fails to reject the null hypothesis for the distributions of STEM GPA by gender, indicating no significant difference between distributions; (a) optimal bandwidth = 0.896 for males, 0.906 for females; (b) optimal bandwidth = 0.672 for males, 0.614 for females.

STEM advantage comes from the fact that there is much more mass at the upper end of the non-STEM GPA distribution, indicating that it is easier to get a top grade in these subjects when compared to STEM subjects. We find no significant gender differences in STEM or non-STEM GPA ranks. Since ranks do not vary across subjects, the grade-based STEM advantage seems to be driven by generally higher grades in non-STEM subjects compared to STEM subjects.

Panel C of Table 3.1 displays gender differences in background variables. While we observe considerable variation in our key variables of interest, the selected nature of our sample – individuals from upper secondary education who have enrolled in tertiary education – may limit representativeness. In particular, the selection process for university enrollment could differ between males and females. To address potential

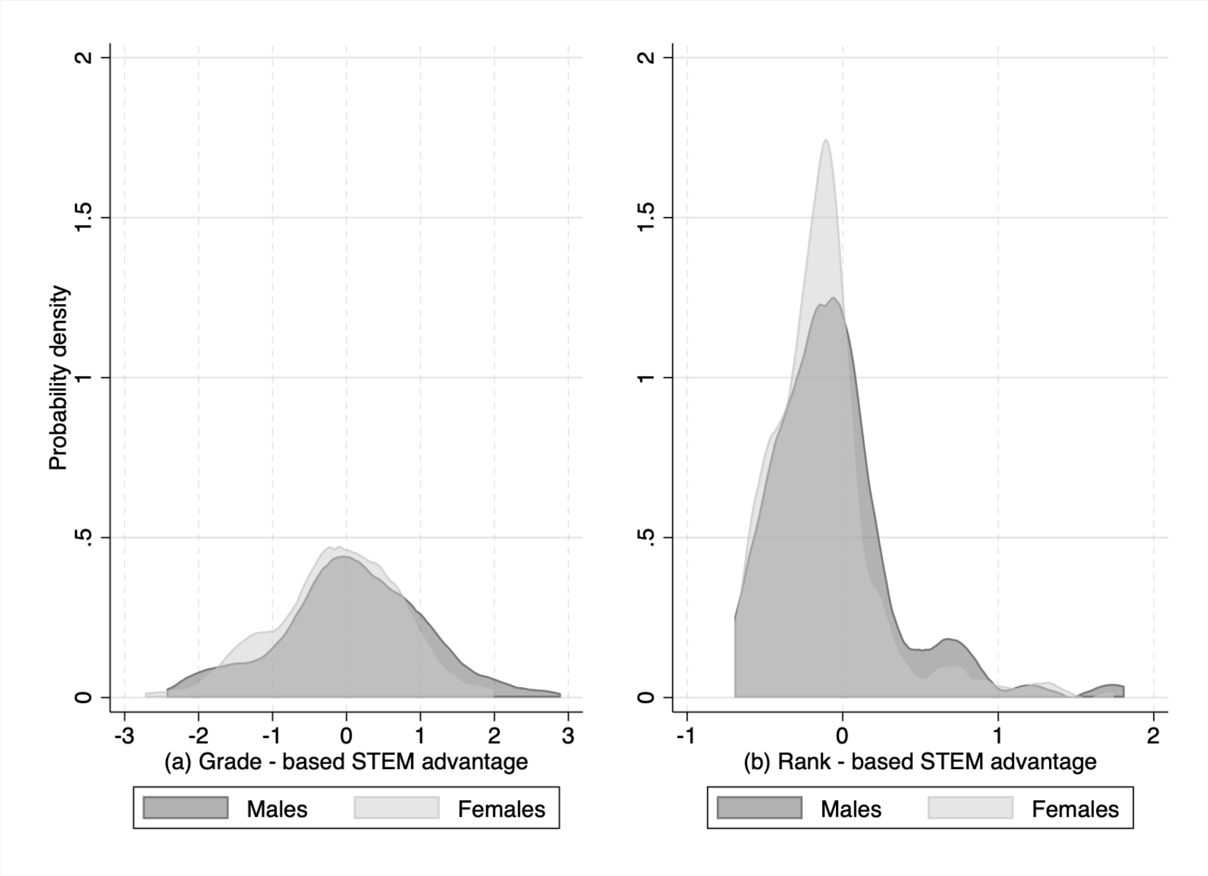

Figure 3.2: Kernel Densities of Grade- and Rank-Based STEM Advantage

Note: The figure displays kernel density plots of our measures of relative STEM advantage for males and females, respectively. Kernel = Epanechnikov; a two-sample Kolmogorov-Smirnov test indicates a significant difference between distributions of a grade-based STEM advantage by gender($p<.10$) and rank-based STEM advantage ($p<.05$) , rejecting the null hypothesis of equal distributions; (a) optimal bandwidth = 0.267 for males, 0.222 for females; (b) optimal bandwidth = 0.097 for males, 0.068 for females. We remove 14 outliers by dropping observations where grade-based STEM advantage exceeds 3, or rank-based STEM advantage exceeds 2. Both measures are standardized.

sample selection issues, we control for cohort-level performance using average GPAs in overall high school performance, STEM subjects, and non-STEM subjects. Furthermore, we find low SES and IQ to be statistically significant factors, so we also control for these background characteristics in our empirical analysis.

Panel D of Table 3.1 provides summary statistics on major choice and academic performance. It shows the distribution of male and female students across STEM, Law, Economics and Business, Humanities and Social Sciences, and Health and Natural Sciences degree programs. STEM majors, characterized by their math-intensive nature, attract a significantly larger proportion of male students, resulting in a gender gap of 24%. In contrast, there is a reversed gender gap of 24% for the choice of majors in Humanities and Social Sciences, where there is a substantially higher representation of female students. Further, we see that female students outperform their male counterparts in terms of academic grades (0.20 sd).

Figure 3.3 graphically illustrates average marginal effects of STEM GPA and grade-based STEM advantage on the probability to pursue a STEM degree by gender. We observe a consistently higher likelihood of males choosing STEM majors over females based on STEM GPA and grade-based STEM advantage. Males' likelihood of choosing a STEM degree increases with their STEM advantage, while females show a more modest response.

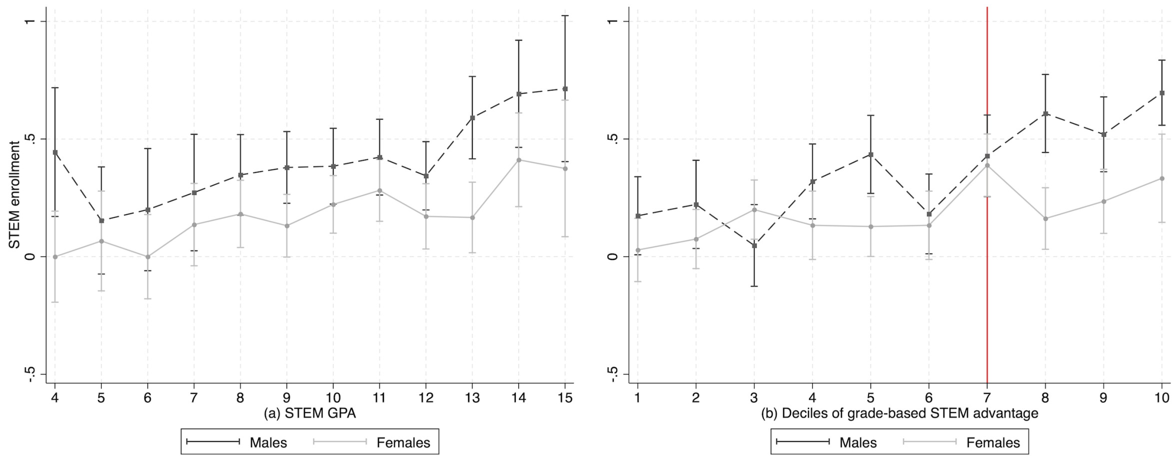

Figure 3.3: STEM Enrollment by STEM GPA and Grade-Based STEM Advantage

Note: For (a) STEM GPA, we predict average marginal effects on the probability of studying STEM across a range of values from 4 (minimum passing grade) to 15 points, separately for males and females. For (b) deciles of grade-based STEM advantage, we regress STEM enrollment on the deciles of grade-based STEM advantage interacted with a female indicator. We then calculate predicted probabilities of STEM enrollment across the deciles of grade-based STEM advantage by gender. The vertical line at the 7^{th} decile represents the threshold, where the right-hand side indicates having a positive grade-based STEM advantage.

3.3 Empirical Results

3.3.1 The Relationship between STEM Enrollment and Relative Performance Indicators

We estimate a linear probability model to investigate the relationship between our relative STEM performance indicators and the choice of a STEM major.

$$\begin{aligned} Y_i = \alpha + \beta_0 \times Female_i + \\ \beta_1 \times Performance\ indicator_{pi} + \\ \beta_2 Female_i \times Performance\ indicator_{pi} + \\ \gamma' X_i + \delta_t + \varepsilon_i \end{aligned} \tag{3.4}$$

Y_i is a binary variable indicating whether individual i enrolled into a STEM major. $Female_i$ is a dummy variable equal to one when i is female, zero if male. $Performance\ indicator_{pi}$ is a placeholder for either grade- or rank-based STEM advantage. We introduce an interaction term of performance indicators and the female dummy. Both performance measures capture the proficiency of individual i in STEM subjects compared to non-STEM subjects, with grade-based STEM advantage focusing on performance differences based on grades and rank-based STEM advantage being based on information from local percentile ranks. Our control variables, denoted as X_i, include both GPAs and ranks in STEM and non-STEM subjects to account for students' absolute performance levels. We incorporate school-cohort performance by including average STEM and non-STEM GPAs, along with high school GPAs. To control for ability, we include measures of IQ[6] and the individual's high school GPA. Additionally, we include personal background information on socioeconomic status and migration status. To accommodate potential graduation-year characteristics, we introduce graduation-year dummies denoted as δ_t. The error term is represented by ε_i. We estimate the model using robust standard errors. Both performance indicators are standardized in order to compare effect sizes across variables.

The coefficient of primary interest is β_2, denoting potential heterogeneity in the importance of our performance indicators for the likelihood of pursuing a STEM degree across genders. The results from estimating this model are displayed in Table 3.2. The full table is displayed in Table B.2 of Appendix B, Section 3.5.3.

First, we inspect the results without interaction terms in Columns 1 and 2 for grade-based STEM advantage and Columns 4 and 5 for rank-based STEM advantage. Following this, we explore gender differences in the effects of grade-based STEM advantage in Column 3 and rank-based STEM advantage in Column 6. The results indicate a substantial and statistically significant gender gap of approximately 17 percentage points in STEM enrollment across all specifications ($p<.01$). A one-standard-deviation increase in grade-based STEM advantage boosts the probability of pursuing a STEM degree by 15-19 percentage points ($p<.01$), ceteris paribus.

The results from the interaction model displayed in Column 3, indicate that being female reduces the positive effect of a one-standard-deviation increase in grade-based STEM advantage by 9.7 percentage points relative to males ($p<.01$). Thus, while grade-based STEM advantage increases the likelihood of choosing a STEM subject for both genders, the effect is around 50% smaller for females. Hence, although females show similar performance in STEM subjects, the influence of grade-based STEM advantage

[6]We measured IQ based on ten items from a Raven-type matrices IQ test (Raven and Court 1998).

Table 3.2: STEM Enrollment and Relative Performance Indicators

	Grade-based STEM advantage			*Rank-based STEM advantage*		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.184***	-0.165***	-0.160***	-0.200***	-0.169***	-0.168***
	(0.036)	(0.039)	(0.039)	(0.037)	(0.039)	(0.039)
Grade-based STEM advantage	0.174***	0.152***	0.188***			
	(0.055)	(0.054)	(0.052)			
Female×Grade-based STEM advantage			-0.097***			
			(0.034)			
Rank-based STEM advantage				0.023	0.020	0.042***
				(0.016)	(0.015)	(0.011)
Female×Rank-based STEM advantage						-0.058**
						(0.028)
Grades	Yes	Yes	Yes	No	No	No
Ranks	No	No	No	Yes	Yes	Yes
Other controls	No	Yes	Yes	No	Yes	Yes
Observations	573	573	573	573	573	573
Adjusted R^2	0.134	0.168	0.178	0.109	0.151	0.153

Note: Columns 1-2 and Columns 4-5 present estimated effects of grade- or rank-based STEM advantage on STEM enrollment in tertiary education, respectively. Columns 3 interacts grade-based STEM advantage with gender to identify heterogeneous effects. Column 6 repeats this analysis for rank-based STEM advantage. Regressions are estimated with a constant, control for STEM and non-STEM GPAs, and ranks of STEM and non-STEM GPA. Other controls include school-cohort performance as measured by STEM, non-STEM, and high school GPA, individual ability approximated by IQ and high school GPA, personal background such as socioeconomic status, migration status, and graduation-year dummies, which are omitted in this table for brevity. We use robust standard errors, reported in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

on choosing a STEM major is disproportionately smaller for them compared to males. That is, females require significantly stronger signals of relative ability based on grades in STEM compared to non-STEM subjects to pursue a STEM degree. In fact, to attain an equivalent probability of pursuing a STEM degree as males, females require a grade-based STEM advantage that is almost four standard deviations higher.[7] In Column 6, the interaction term (Female × Rank-based STEM advantage) largely offsets or even reverses the main effect. This is a discouraging finding: Females' relative STEM performance ranking seems to have no influence on their likelihood of choosing a STEM major.

Our results point towards substantial selection costs of choosing a STEM occupation among females. Further, one interpretation of our findings is that non-performance-related factors, such as field preferences, perceived future working conditions, and perceived discrimination strongly discourage females from choosing a STEM occupation, even if they obtain very positive signals about their STEM abilities. As a consequence, even strong signals about STEM performance hardly affect female choices compared to males.

3.3.2 Quantifying Decision-Making Differences: Male vs. Female Choice Worlds

In the previous sections, we identified gender disparities in STEM performance and observed varying effects of STEM advantage across genders, indicating that strong signals of having a grade-based STEM advantage hardly induce females to choose a STEM major. In this section, we conduct a twofold Kitagawa-Oaxaca-Blinder-type decomposition (Jann 2008) to assess the contributions of performance differences to the gender gap. Our goal is to run the following thought experiment: How would observed performance differences among males and females affect the sorting into STEM-related fields, in a world where female (male) major preferences in STEM-related fields resembled that of males (females)? To assess the relative importance of performance measures in a non-discriminatory "male-choice" or a discriminatory "female-choice" world, we respectively categorize potential drivers of the STEM gender gap into two groups. The first group represents the grade-based performance indicators, incorporating grade-based STEM advantage, STEM GPA, and non-STEM GPA. The second group represents the rank-based performance indicators, comprising rank-based STEM advantage, rank of STEM GPA, and rank of non-STEM GPA.

[7] We want to set the probability of males to pursue a STEM degree equal to the probability for females conditional on grade-based STEM advantage. Given Table 3.2, we solve for x in 0.188=-0.160+x(0.188-0.097), which leads to an x=3.8.

Consider two groups, *male* and *female*, with our outcome variable Y representing STEM major, and our set of predictors X driving the STEM gender gap categorized above in grade- and rank-based performance indicators. We define the mean outcome difference as follows,

$$R = E(Y_{male}) - E(Y_{female}). \tag{3.5}$$

Given $E(Y)$ as the expected value of the outcome variable, we seek to understand the extent to which group differences in predictors contribute to the mean outcome difference. The twofold decomposition dissects outcome differences into an explained part (“quantity effect") and an unexplained component (which we term “preferences effect").

Consider the following linear model:

$$Y_l = X_l' \beta_l + \epsilon_l. \tag{3.6}$$

Assuming $E(\epsilon_l) = 0$ for each group $l \in$ (male, female), within the linear model framework, where X represents a vector containing predictors and a constant, β encompasses slope parameters and intercept, and ϵ denotes the error term, the mean outcome difference in the twofold decomposition can be expressed as:

$$\begin{aligned} R =& \{E(X_{male}) - E(X_{female})\}'\beta^* \\ &+ \{E(X_{male})'(\beta_{male} - \beta^*) + E(X_{female})'(\beta^* - \beta_{female})\}. \end{aligned} \tag{3.7}$$

The first component,

$$Q = \{E(X_{male}) - E(X_{female})\}'\beta^*, \tag{3.8}$$

is the part of the outcome differential that is explained by group differences in the predictors, the *quantity effect*. The second component,

$$U = \{E(X_{male})'(\beta_{male} - \beta^*) + E(X_{female})'(\beta^* - \beta_{female})\}, \tag{3.9}$$

is the *preferences effect*, i.e., the part that reflects differences in decision-making that can be due to (anticipated) discrimination, considerations about fit, unobserved factors or preference heterogeneity between males and females.

Table 3.3 presents two set of results. In Columns 1 and 2, following the literature on STEM disparities, we posit that the STEM enrollment gap is biased against women rather than men.[8] Thus, in Equation 3.7, we use the coefficients of males, denoted as β_{male}, for β^*, evaluating the relative significance of performance measures in a non-discriminatory "male-choice" world. In Columns 2 and 3, we set $\beta_{female} = \beta^*$ to analyse the influence of performance measures in a "female-choice" scenario.

Table 3.3: Kitagawa-Oaxaca-Blinder Decomposition of the STEM Gender Gap

	Male coefficients		*Female coefficients*	
	(1) Absolute	(2) Share	(3) Absolute	(4) Share
Difference	0.211***	100.000	0.211***	100.000
	(0.039)		(0.039)	
Explained difference	0.045*	21.327	0.059***	27.962
	(0.025)		(0.017)	
Composition effects attributable to				
(A) Grade-based performance indicator	0.054**	25.592	0.013	6.161
	(0.023)		(0.010)	
(B) Rank-based performance indicator	-0.010	4.739	0.001	0.474
	(0.009)		(0.007)	
Control variables	-0.001	0.474	0.045***	21.327
	(0.015)		(0.015)	
Observations	573		573	

Note: This table decomposes differences in STEM subject choice in tertiary education attributable to differences in grade- and rank-based performance indicators using a twofold Kitagawa-Oaxaca-Blinder decomposition. We control for IQ, socioeconomic status, and migration status. Columns 1 and 2 use male coefficients for the unknown non-discriminatory coefficients vector β^*, while Columns 3 and 4 use female coefficients. For each decomposition, we also present the share of the difference that is attributable to the respective component. Robust standard errors in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[8]Determining the components of the twofold decomposition, as shown in Equation 3.7, requires an estimate for the unknown non-discriminatory coefficients vector β^*. Oaxaca (1973) suggests that β^* can equal either β_{male} or β_{female} based on the direction of discrimination towards a particular group.

Conceptually, the β-coefficients can be interpreted as preference parameters, reflecting the decision-making tendencies of males and females when confronted with distinct abilities and constraints. By using male coefficients in Columns 1 and 2, we gain insights into the role of gender differences in decision-making processes. Assuming that females may decide differently due to preference-related factors such as discouragement, lack of role models, or concerns about penalties related to family responsibilities in STEM fields, we can assess the extent to which relative STEM performance disparities would persist if these barriers were eliminated.

Abstracting from these barriers, 21% of the STEM gender gap stems from group differences in predictors (quantity effect). Our decomposition reveals that gender differences in grade-based performance indicators account for 26% ($p<.05$) of the STEM enrollment gap, while rank-based metrics show no significant impact. When examining a counterfactual scenario where preferences are "female" (Columns 3-4), neither grade- nor rank-based performance disparities significantly influence STEM-field selection. Instead, the gender gap in STEM enrollment is largely driven by differences in our control variables: IQ, migration status, and socioeconomic status. This observation aligns with our previous finding presented in Table 3.2 , where rank-based STEM advantage does not demonstrate significant economic relevance in relation to the STEM enrollment gap. Since the coefficients of rank-based STEM advantage and its interaction with gender offset each other, we do not find an overall effect of rank-based STEM advantage on STEM choice in a "female-preference" world.

To address potential concerns about the choice of reference coefficients in decomposition analyses (Neumark 1988; Oaxaca and Ransom 1994), we additionally estimate a pooled model where coefficients are derived from a regression. Results presented in Table B.3 of Appendix B largely align with our main specifications: the explained portion of the gender gap remains substantial at 21%, with grade-based performance indicators continuing to be the primary driver, accounting for 14% of the gap ($p<.05$). The insignificant role of rank-based measures persists across all specifications. This is unsurprising given that there are no male-female differences in rank-based STEM advantage.

Given that performance differentials minimally influence female STEM enrollment, we examine alternative drivers. Prior literature suggests anticipated gender discrimination may deter STEM pursuit (e.g., Porter and Serra 2020). Using a linear probability model (Table 3.4), we examine whether female students in STEM programs report higher levels of anticipated discrimination. Our preferred specification (Column 4) shows females experience a 32 percentage point higher probability of anticipated gender discrimination (p<.01), with an additional 19.8 percentage point increase among STEM-enrolled females (p<.05). This implies that females pursuing STEM degrees expect additional obstacles. To the extent that only those women select into STEM fields who expect less discrimination in a STEM-related occupation, our estimates provide a lower bound estimate of perceived barriers or discrimination in the STEM occupations. For detailed table contents, we refer to Table B.4 of Appendix B in Section 3.5.3.

Table 3.4: Anticipated Discrimination and STEM Enrollment

	Anticipated gender-based discrimination			
	(1)	(2)	(3)	(4)
Female	0.382***	0.375***	0.336***	0.319***
	(0.036)	(0.038)	(0.041)	(0.043)
STEM major	0.043	0.037	-0.036	-0.057
	(0.040)	(0.044)	(0.041)	(0.043)
Female×STEM major			0.169**	0.198**
			(0.082)	(0.087)
Relative STEM advantages	No	Yes	No	Yes
Grades	No	Yes	No	Yes
Ranks	No	Yes	No	Yes
Other controls	No	Yes	No	Yes
Observations	573	573	573	573
Adjusted R^2	0.150	0.138	0.154	0.144

Note: Columns 1-2 present estimated effects of gender and STEM enrollment on the expectation of gender-based discrimination. Columns 3-4 interact STEM enrollment with gender to identify heterogeneous effects. Regressions are estimated with a constant. 'Relative STEM advantages' includes our grade- and rank-based measures of STEM advantage. 'Grades' controls for STEM and non-STEM GPA and 'Ranks' controls for rank of STEM and non-STEM GPA. Other controls include school-cohort performance as measured by STEM, non-STEM, and high school GPA, individual ability approximated by IQ and high school GPA, personal background such as socioeconomic status, migration status, and graduation-year dummies, which are omitted in this table for brevity. We use robust standard errors, reported in parenthesis.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.4 Conclusion

Gender disparities in relative performance across STEM and non-STEM fields have long-lasting effects, potentially affecting not only educational decisions, but also leading to wage disparities (e.g., Blau and Kahn 2017; Kleven, Landais and Søgaard 2019), especially in sectors such as science and technology (Goldin 2014) and countries like Germany where occupational mobility is low (SOEP 2022). The scarcity of STEM graduates, particularly among women, poses a significant challenge to the tech industry and innovation in general (Carnevale, Smith and Melton 2011; Bianchi and Giorcelli 2020; Del Carpio and Guadalupe 2022; Coff 1997).

Our study examines how performance indicators affect human-capital investment and selection into STEM-related fields in Germany. Germany's publicly funded, tuition-free education system, coupled with standardized curricula and compensation schemes for teachers, minimizes financial constraints on educational choices. Persistent gender-based disparities in tertiary-education choices thus seemingly reflect significant non-monetary factors in educational decision-making.

Our study identifies important gender dynamics in educational decision-making. Women graduate from high school with better GPAs. Although there are no significant or large gender disparities in STEM subject performance, males perform significantly worse than females in non-STEM subjects. Additionally, we observed a 24% gender gap in male-female STEM major choices, aligning with trends observed across OECD countries, including Germany (OECD 2024). Our analysis reveals that an grade-based STEM advantage is positively associated with STEM enrollment among both genders, though the impact is much smaller for females compared to males. This findings indicate that, despite similar proficiency in STEM subjects, females require substantially stronger grade-based performance signals in STEM relative to non-STEM subjects to pursue a STEM degree. However, women who do enter STEM programs ultimately outperform their male peers in terms of GPA.

Our decomposition analysis examines how differences in relative performance indicators and control variables contribute to the STEM gender gap under different preference scenarios. We show that in a male-choice world, where female preferences mirror those of males, 21% of the STEM gender gap can be attributed to group differences in predictors. Specifically, differences in grade-based STEM advantage and performance gaps across STEM and non-STEM subjects account for 26% of the gender gap, while rank-based differences explain only 5%. Conversely, in a female-choice world grade-and rank-based performance differences account for merely 6% and 0.5%, respectively. These results underscore that performance variations play a minimal role in females' STEM choices, with rank-based STEM advantage showing consistently low economic significance across specifications.

We interpret these findings of being indicative of substantial perceived selection costs to entering a STEM occupation among females and we show that women who selected into STEM majors indeed expect more gender-related on-the-job discrimination.

Our analysis is informative as regards policies that address the underrepresentation of women in STEM and the ongoing discourse on gender inequality in education and labor markets (Goldin 2014; Marianne 2011; Blau and Kahn 2017; Kleven and Landais 2017; Goulas, Griselda and Megalokonomou 2022; Breda and Napp 2019; Breda et al. 2019; Breda, Jouini and Napp 2018, 2023; Francesconi and Parey 2018; Zafar 2013; Wiswall and Zafar 2018). Our results suggest that a change in STEM grades or respective grading policies in secondary school will have little impact on reducing the gender gap in STEM. On the positive side, this also implies that "easier grades" as observed mostly in non-STEM subjects might not systematically drive females out of STEM fields, although this could differ when comes to grades in higher education (Ahn et al. 2024).

Our results further indicate that, despite similar proficiency in STEM subjects, females seem to be held back by differential preferences or barriers when it comes to pursuing a STEM degree. Non-performance-related factors such as field preference, perceived future working conditions, and perceived discrimination seemingly discourage females from choosing a STEM occupation even if they obtain very positive signals about their STEM abilities. Specifically, females require a STEM advantage that is four standard deviations higher than males to have the same probability of studying a STEM subject, highlighting the need for more encouragement and support for women in science. In line with Breda, Jouini and Napp (2023), our findings suggest significant overselection costs, emphasizing the importance of addressing gender-specific barriers in STEM fields to ensure equitable opportunities for all aspiring students.

A systematic analysis of measures and programs aiming to effectively counter perceived on-the-job discrimination in STEM occupations would be an interesting endeavor for future research. In light of our findings, such policies could reduce the STEM-enrollment gap, improve the talent pool in STEM occupations, and may ultimately improve a country's growth and competitiveness.

3.5 Appendix B

3.5.1 Institutional Background

Upon completing primary school, students are channeled into different secondary-school types. The four main types of secondary schools are: *Gymnasium* (academic secondary school/high school, ISCED Level 3), *Realschule* (intermediate secondary school, ISCED Level 2), *Hauptschule*, and *Gesamtschule* (comprehensive school, ISCED Level 2).[9] In our analysis we focus on the *Gymnasium*, which can either last eight (G8) or nine (G9) years depending on the state and cohort. It culminates in the *Abitur*, the highest secondary-school certificate and a prerequisite for admission to tertiary education. While transitions between these four types of tracks are theoretically possible at any time, the frequency and structure of such transitions vary by state, with most upward movements occurring after the completion of lower secondary programs. This tracking system plays a crucial role in shaping students' future academic and professional paths within the German education landscape. For those interested in further details, we direct their attention to the regulatory framework governing both the upper secondary level of *Gymnasium* and the *Abitur* examination in North Rhine-Westphalia. This framework is established by the "Verordnung über den Bildungsgang und die Abiturprüfung in der gymnasialen Oberstufe" (APO-GOSt).[10] Enacted on 5 October 1998, this foundational regulation provides the legal and structural basis for the educational processes and assessment methods analysed in our study. The APO-GOSt serves as a cornerstone document, outlining the curriculum structure, examination procedures, and qualification requirements for students in the upper secondary level of *Gymnasiums* in North Rhine-Westphalia.

[9] ISCED-97 definitions provided by the OECD (2017).

[10] This translates to "Ordinance on the Educational Path and *Abitur* Examination in Upper Secondary Education" and can be accessed here: `https://recht.nrw.de/lmi/owa/br_text_anzeigen?v_id=10000000000000000186`.

3.5.2 Variable Descriptions

Table B.1: Variable Definitions

Variable	Description
High school GPA	Between 1.0 and 4.0 (higher better)
STEM GPA	0-15 points (higher better)
Non-STEM GPA	0-15 points (higher better)
Rank of STEM GPA	Between 0 and 100 (higher better)
Rank of non-STEM GPA	Between 0 and 100 (higher better)
High school GPA (cohort)	Between 1.0 and 4.0 (higher better)
STEM GPA (cohort)	0-15 points (higher better)
Non-STEM GPA (cohort)	0-15 points (higher better)
Low SES	Dummy that equals 1 if at least one parent has a high school diploma, 0 else
Migration status	Dummy that equals 1 if individual grew up in another country, 0 else
University GPA	Between 1.0 and 4.0 (higher better)
University STEM GPA	Between 1.0 and 4.0 (higher better)
University non-STEM GPA	Between 1.0 and 4.0 (higher better)

3.5.3 Tables

Table B.2: STEM Enrollment and Relative Performance Indicators

	Grade-based STEM advantage			*Rank-based STEM advantage*		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.184***	-0.165***	-0.160***	-0.200***	-0.169***	-0.168***
	(0.036)	(0.039)	(0.039)	(0.037)	(0.039)	(0.039)
Grade-based STEM advantage	0.174***	0.152***	0.188***			
	(0.055)	(0.054)	(0.052)			
Female×Grade-based STEM advantage			-0.097***			
			(0.034)			
Rank-based STEM advantage				0.023	0.020	0.042***
				(0.016)	(0.015)	(0.011)
Female×Rank-based STEM advantage						-0.058**
						(0.028)
STEM GPA	-0.015	-0.003	0.003			
	(0.018)	(0.019)	(0.020)			
Non-STEM GPA	0.015	0.023	0.020			
	(0.016)	(0.017)	(0.017)			
Rank STEM GPA				0.004***	0.004***	0.004***
				(0.001)	(0.001)	(0.001)
Rank non-STEM GPA				-0.001	-0.001	-0.001
				(0.001)	(0.001)	(0.001)
STEM GPA (cohort)		0.005	0.006		0.044***	0.043***
		(0.015)	(0.015)		(0.016)	(0.016)

Continues next page

Table B.2: STEM Enrollment and Relative Performance Indicators (Continued)

	Grade-based STEM advantage			*Rank-based STEM advantage*		
	(1)	(2)	(3)	(4)	(5)	(6)
Non-STEM GPA (cohort)		-0.006	-0.007		-0.021	-0.020
		(0.020)	(0.020)		(0.020)	(0.020)
High school GPA (cohort)		-0.312**	-0.291*		-0.330**	-0.332**
		(0.155)	(0.153)		(0.157)	(0.155)
IQ		0.010	0.009		0.009	0.008
		(0.010)	(0.011)		(0.011)	(0.011)
High school GPA		-0.049	-0.053		-0.056	-0.055
		(0.048)	(0.048)		(0.044)	(0.044)
Low SES		0.094**	0.104***		0.093**	0.096**
		(0.040)	(0.040)		(0.041)	(0.041)
Migration status		0.099	0.120*		0.096	0.105
		(0.067)	(0.068)		(0.068)	(0.068)
Grad.-year	No	Yes	Yes	No	Yes	Yes
Observations	573	573	573	573	573	573
Adjusted R^2	0.134	0.168	0.178	0.109	0.151	0.153

Note: Columns 1-2 and Columns 4-5 present estimated effects of grade- or rank-based STEM advantage on STEM enrollment in tertiary education, respectively. Columns 3 interacts grade-based STEM advantage with gender to identify heterogeneous effects. Column 6 repeats this analysis for rank-based STEM advantage. Regressions are estimated with a constant, control for STEM and non-STEM GPAs, and ranks of STEM and non-STEM GPA. Other controls include school-cohort performance as measured by STEM, non-STEM, and high school GPA, individual ability approximated by IQ and high school GPA, personal background such as socioeconomic status, migration status, and graduation-year dummies. Graduation-year dummies are omitted for brevity. We use robust standard errors, reported in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.3: Kitagawa-Oaxaca-Blinder Decomposition of the STEM Gender Gap

	Pooled regression model	
	(1) Absolute	(2) Share
Difference	0.211***	100.000
	(0.038)	
Explained difference	0.045***	21.327
	(0.016)	
Composition effects attributable to		
(A) Grade-based performance indicator	0.030**	14.218
	(0.013)	
(B) Rank-based performance indicator	-0.002	0.948
	(0.007)	
Control variables	0.017*	8.057
	(0.010)	
Observations	573	

Note: This table decomposes differences in STEM subject choice in tertiary education attributable to differences in absolute and relative performance indicators using twofold Kitagawa-Oaxaca-Blinder decomposition from a pooled regression model. We control for IQ, socioeconomic status, and migration status. For each decomposition, we also present the share of the difference that is attributable to the respective component. Robust standard errors in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.4: Anticipated Discrimination and STEM Enrollment

	Anticipated gender-based discrimination			
	(1)	(2)	(3)	(4)
Female	0.382***	0.375***	0.336***	0.319***
	(0.036)	(0.038)	(0.041)	(0.043)
STEM major	0.043	0.037	-0.036	-0.057
	(0.040)	(0.044)	(0.041)	(0.043)
Female×STEM enrollment			0.169**	0.198**
			(0.082)	(0.087)
Grade-based STEM advantage		0.004		0.008
		(0.066)		(0.065)
Rank-based STEM advantage		0.029		0.031
		(0.025)		(0.026)
STEM GPA		0.031		0.033
		(0.032)		(0.031)
Non-STEM GPA		0.000		0.001
		(0.028)		(0.027)
Rank STEM GPA		-0.003		-0.004
		(0.003)		(0.003)
Rank non-STEM GPA		0.000		0.000
		(0.002)		(0.002)
STEM GPA (cohort)		-0.013		-0.017
		(0.027)		(0.027)
Non-STEM GPA (cohort)		0.009		0.004
		(0.025)		(0.025)
High school GPA (cohort)		-0.114		-0.097
		(0.175)		(0.174)
IQ		0.004		0.008
		(0.011)		(0.011)
High school GPA		0.034		0.035
		(0.055)		(0.055)
Low SES		0.007		-0.001
		(0.042)		(0.042)
Migration status		-0.068		-0.074
		(0.057)		(0.055)
Grad.-year	No	Yes	No	Yes
Observations	573	573	573	573
Adjusted R^2	0.150	0.138	0.154	0.144

Note: Columns 1-2 present estimated effects of gender and STEM enrollment on the expectation of gender-based discrimination. Columns 3-4 interact STEM enrollment with gender to identify heterogeneous effects. Regressions are estimated with a constant, control for grade- and rank-based STEM advantage, STEM and non-STEM GPAs, and ranks of STEM and non-STEM GPA. Other controls include cohort performance as measured by STEM, non-STEM, and high school GPA, individual ability approximated by IQ and high school GPA, and personal background such as socioeconomic status and migration status. Graduation-year dummies are omitted in this table for brevity. We use robust standard errors, reported in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4 Climate Stress Tests, Bank Lending, and the Transition to the Carbon-Neutral Economy

joint with Huyền Nguyễn, Trang Nguyễn and Klaus Schaeck*

Abstract

We ask if bank supervisors' efforts to combat climate change affect bank lending and their borrowers' transition to the carbon-neutral economy. Combining information from the French supervisory agency's climate pilot exercise with borrowers' emission data, we first show that banks that participate in the exercise increase lending to high carbon emitters, but simultaneously charge higher interest rates. Second, participating banks collect new information about climate risks and boost lending for green purposes. Third, receiving credit from a participating bank facilitates the borrowers' efforts to improve environmental performance. Our findings establish a hitherto undocumented link between banking supervision and the transition to net zero.

Keywords: Climate Stress Test, Carbon Risk, Banking Supervision, Syndicated Loans, Green Finance

JEL classification: G21, G28, K11

*We are grateful for helpful comments and suggestions from Andreas Beyer, Giorgia Barboni, Ricardo Correa, Hans Degryse, Klaus Duellmann, Andrea Enria, Bill English, Ivan Ivanov, Tristan Jourde, Moqi Groen-Xu, Reint Gropp, Nadja Günster, Andreas Fuster, Xavier Freixas, Thomas Kick, Michael Koetter, Philipp Klein, Kai Li, Jose Lopez, Michala Marcussen, Christoph Meinerding, Ralf Meisenzahl, Louis Nguyen, Steven Ongena, Pia Pinger, Andreas Pfingsten, Martin Oehmke, Larissa Schaefer, Merih Sevilir, Christoph Schneider, Alexander Schulz, Zacharias Sautner, Ulrich Wagner, Shuo Xia, and conference and seminar participants at the Bank of England, the Deutsche Bundesbank, the inaugural ECB Research Conference on Banking Supervision, the ECB-IMF Macroprudential Policy and Research Conference, the EFI Network Meeting at the National Bank of Belgium, the Federal Reserve Bank of Boston Conference on Stress Tests, the University of St Andrews, the SURF Interagency Seminar, the University of Münster, the University of Nottingham, and the University of Tübingen.

4.1 Introduction

Central banks and regulatory and supervisory agencies are at the forefront of the fight against climate change.[1] Droughts and floods pose physical risk, and the changing policies and preferences in the behavior of economic agents affect the valuation of assets and liabilities, thus posing a transition risk as borrowers from banks are ill-prepared for the decarbonization of their business models. Therefore, supervisory agencies start to conduct climate stress tests to assess the resilience of banking systems to climate change. Despite the key role of supervisory agencies to combat climate change, little is known about how such efforts affect bank lending, nor whether they play a role in the transition to a carbon-neutral economy.

In this paper, we exploit plausibly exogenous variation in climate stress tests, approximated by the participation of banks in the French banking regulators' climate pilot exercise, as a proxy for supervisory efforts to tackle climate change. Our aim is to estimate the effect of banking supervision on the borrowers' environmental performance via the banks' lending decisions. While climate stress tests are primarily driven by concerns about financial stability, we characterize climate stress tests as an information production exercise that uncovers new information about the exposure of banks to climate change. We then combine data from the French climate pilot exercise with borrower-specific information on carbon emissions to understand whether the participation of banks in this climate pilot exercise affects the borrowers' environmental performance.

Our approach is econometrically appealing because it enables us to examine how supervisory efforts to understand risks arising from climate change affect lending decisions conditional on the borrowers' exposure to transition risk. Our setup allows us to disentangle the information value contained in carbon emissions of borrowers from the incremental reduction in information asymmetries related to transition risk arising from participation in the climate pilot exercise. This information advantage enables banks to improve their understanding, assessment, and management of the long-term consequences of transition risk. To do so, we compare bank lending to high carbon emitters (brown borrowers) with bank lending to low carbon emitters conditional on whether banks participate in the French bank supervisors' climate pilot exercise.

[1] Regulation focuses on the development and promulgation of rules under which financial intermediaries operate (Eisenbach, Lucca and Townsend 2016), whereas supervision is concerned with the monitoring of financial firms to ascertain compliance with laws and regulations to ensure safe and sound operations. The organization of regulation and supervision varies across jurisdictions, with regulation and supervision being either orchestrated within the central bank or by separate authorities. While climate change affects all dimensions of the regulatory and supervisory environment, climate stress tests are typically performed by supervisory agencies, and we therefore refer to 'supervisory efforts' or 'supervisory actions' to combat climate change in this research.

We find that the climate pilot exercise informs the participating banks' lending decisions above and beyond the publicly available information on borrower-specific exposure to transition risk, approximated by carbon emissions. Most importantly, high carbon emitters whose banks take part in the climate pilot exercise obtain more credit, albeit at higher loan rates. Such borrowers also take actions to make their business models more resilient toward transition risk. In contrast, borrowers whose banks do not participate receive less credit, and show little progress to decarbonize their business models. A further novel insight from our research is that we are able to document empirically the production of new information following the climate pilot exercise. We present evidence that participating banks display a significantly higher tendency to discuss climate stress test scenarios in their earning calls, that they communicate more frequently with borrowers about transition risk issues, and that they discuss carbon emissions more often in their earning calls.

While a growing literature examines how banks incorporate climate change into lending decisions (Murfin and Spiegel 2020; Nguyen et al. 2022; Ouazad and Kahn 2022), little is known about how borrowers' business models are affected by bank supervisors' actions to address climate change. Borio, Claessens and Tarashev (2023) argue that it is unrealistic to expect financial institutions to finance the green transition without clear expectation on regulatory changes. Oehmke and Opp (2022) find that carbon-related capital requirements allow banks to manage transition risk, but that these requirements are inferior to carbon taxes in reducing carbon emissions. By leveraging data from the French supervisory agency and combining them with the exposure of borrowers to transition risk, our analysis of lending decisions allows us to establish a hitherto undocumented mechanism through which supervisory actions related to climate change affect bank borrowers in their efforts to transform their businesses on the way to the carbon-neutral economy.

Our starting points are theories by Goldstein and Sapra (2014) that predict that stress tests reduce information asymmetries, as well as uncover and release new information; and by Dang, Gorton and Holmstrom (2009) and Gorton and Ordonez (2014), who posit that sudden information shocks trigger information production.

We hypothesize that new information collected during the climate pilot exercise influences how banks lend to brown firms. Changes in bank lending can either facilitate or impede the borrowers' transition to a carbon-neutral economy. While the borrowers' carbon emissions allow banks to assess the borrowers' transition risk, we argue that participation in the climate pilot exercise reduces information asymmetries beyond the information obtained via public information about the borrowers' carbon footprints.

The information collected during the climate pilot exercise, together with supervisory feedback, deepens and refines the participating banks' understanding of climate change and the long-term consequences of transition risk. This may motivate banks to support borrowers in the transformation of their business models by continuing to provide credit. In contrast, non-participating banks are more likely to evaluate transition risk with a short-term perspective and reduce their exposures to such borrowers.

Our findings underscore that supervisory efforts concerning climate change affect borrowers' actions related to climate change. Concerning short-term adjustments, we show that high carbon emitters that received loans from banks participating in the climate pilot exercise are more likely to have eco-friendly products, develop emission policies, are more likely to commit to carbon emission reduction targets, and have higher environmental, social, and governance (ESG) scores, compared to borrowers of non-participating banks. Regarding adjustments that require a longer time to achieve their goals, we document that borrowers of participating banks use higher shares of renewable energy. However, such borrowers do not show improvements in total carbon emissions or direct carbon emission growth. They neither terminate supply chains with environmentally harmful suppliers, nor do they source more environmentally friendly materials.

Funding by banks that participate in the climate pilot exercise is the key driver behind these changes. Despite these borrowers' greater transition risks, banks increase lending by 38% but simultaneously incorporate a transition-risk premium of 12 basis points (bps).[2] Our tests underscore an undocumented role of climate stress tests beyond the identification of the banks' vulnerabilities to climate change. Reading through conference calls of banks in our sample, we find that climate stress tests encouraged banks to communicate more with their borrowers about carbon risk, collect more information on the borrowers' carbon emissions, and discuss more about climate risk scenarios. As a result, the participating banks' deeper understanding of climate change, and transition risk in particular, enables them to support their borrowers on the way to reducing carbon emissions. Evidence on the origination of green loans supports our hypothesis. Climate stress-tested banks are more likely to grant loans to brown borrowers with green purposes or with sustainability-linked provisions compared to non-stress-tested banks. These loans also have longer maturity, reflecting that they are more likely to be used for strategically important projects.

[2] After the climate pilot exercise, participants charge high carbon emitters 8% higher interest rates compared to low carbon emitters. As the mean loan spread in our sample is 150 bps, this effect is equivalent to 8*150/100 = 12 bps.

Climate stress tests, approximated in our setting with the French supervisors' climate pilot exercise, are ideal for examining supervisory efforts to address climate change. While similar to financial stability stress tests in terms of resource intensity and objective of identifying vulnerabilities, climate stress tests take a longer-term horizon to evaluate potential losses when borrower activities do not align with the transition to a carbon-neutral economy. They do not trigger capital charges either, and consequently do not mechanically affect the cost of lending.[3] However, they shift attention to climate change and, importantly, require the participating banks to collect extensive information about exposures to physical and transition risk using scenarios based on carbon prices. This focus on carbon prices reinforces our choice to capture transition risk with the borrowers' total carbon emissions. Climate stress tests therefore can also promote the transition towards the carbon-neutral economy because the information acquired during the climate pilot exercise raises the banks' awareness of climate-transition risks, improving their ability to assess such risks, with corresponding effects on the banks' business strategies and risk management reflected in their lending practices.

To isolate the causal effect of the climate pilot exercise, we built a novel data set. We exploit the first climate stress test, the data of which are publicly available from the French Prudential Supervision and Resolution Authority (Autorité de contrôle prudentiel et de résolution, ACPR), and combine it with syndicated loan data for banks and borrowers, merging this information with the borrowers' carbon emissions, data on their borrowers' environmental performance from Refinitiv, and transcripts of the banks' conference calls from S&P Capital IQ.

The participating nine banking groups operate a universal banking model and represent 85 percent of total assets in the French banking system. Our sample is also representative of other banking systems. Similarly to other European countries, France has a highly developed bank-based financial system with hundreds of smaller banks that, together with foreign banks and a limited number of large institutions supervised by the Single Supervisory Mechanism, provide credit to the economy. These large French banks account for the vast proportion of total assets in the banking system, are represented in our sample, and participated in the climate pilot exercise. Importantly, the availability of data from the climate pilot exercise helps us to identify the role of banking supervision for the transition to the carbon-neutral economy, which is distinct from prior work that examines the banks' commitments to reducing carbon emissions

[3]Oehmke and Opp (2022) show that regulating bank capital to address climate risks may not reduce carbon emissions. Higher capital requirements for carbon-intensive borrowers may crowd out lending to green borrowers and increase bank fragility.

(Kacperczyk and Peydró 2022), carbon emission intensity (Ehlers, Mojon and Packer 2020), the banks' responses to information about physical risk (Correa et al. 2022; Nguyen et al. 2022; Meisenzahl 2023), or news about borrowers harming the environment (Chava 2014; Anginer et al. 2023).

A critical step in our identification strategy is to assess whether our estimates truly reflect lenders' updating beliefs about carbon transition risk as the result of climate stress tests instead of capturing the effects of differences in bank characteristics or other shocks affecting high and low carbon emitters differently. Placebo tests indicate that the lending and pricing behaviour to brown firms of banks that do not participate in climate stress tests does not change. We also rule out that our results are driven by the ECB climate stress tests, or other events happening in our research periods, such as the COVID-19 pandemic and the Russian war against Ukraine. Heckman's selection model shows that our results are not biased because of banks selecting themselves into participating in the climate stress tests. Alternative measurements of transition risk such as the use of carbon emission intensity, Sautner et al. (2023)'s transition risk measures, and Reprisk's environmental risk index (ERI) do not change our findings. We do not find any evidence either that our results can be explained by differences in the borrowers' financial constraints or bank characteristics.

Further analyses reveal heterogeneity in the data. Banks that are members of the United Nations Environment Programme (UNEP) and banks with higher shares of institutional investors are the ones that respond more strongly to climate stress tests.

Our research is important because banks in the EU generate more than 65 percent of their interest income from carbon-intensive industries (European Central Bank 2022). Equally, it is important to evaluate whether actions by bank supervisors who predominately focus on financial-stability concerns arising from climate change also play a role for the transition to the carbon-neutral economy. Moreover, although many banks already started incorporating sustainability concerns into lending activities, they currently lack detailed business strategies, risk management processes, and governance systems to address the challenges related to climate change. Many banks also reveal deficiencies about how to quantify transition risk correctly (European Central Bank 2022). Our work illustrates how supervisory agencies, via climate stress tests, contribute to reducing uncertainties related to climate change and influence banks to promote an orderly transition to the carbon-neutral economy. Finally, in contrast to previous studies that document negative effects, for borrowers, arising from transition risk, our work highlights that banks that participate in a climate stress test reaffirm their commitment to borrowers despite such borrowers' exposure to transition risk. This finding underscores that banking supervision can actively support the transition to a carbon-neutral economy.

We contribute to several different strands in the literature. First, numerous studies examine how supervisory resources and coverage (Eisenbach, Lucca and Townsend 2016; Hirtle, Kovner and Plosser 2020; Goldsmith-Pinkham, Hirtle and Lucca 2016; Ivanov, Kruttli and Watugala 2024), standards (Kiser, Prager and Scott 2012; Bassett, Lee and Spiller 2015), intensity (Agarwal et al. 2014; Rezende and Wu 2014), and enforcement actions affect the performance of banks and their borrowers (Delis and Staikouras 2011; Danisewicz et al. 2018). We contribute to this literature by estimating how supervisory efforts to address climate change produce new information that enables the participating banks better to assess information about borrowers' transition risk and revise lending decisions accordingly.

Second, we also contribute to the literature on stress tests. Morgan, Peristiani and Savino (2014) and Flannery, Hirtle and Kovner (2017) find that stress tests generate valuable information about participating banks. Acharya, Berger and Roman (2018) and Cortés et al. (2020) show stress-tested banks reduce credit, reallocate lending towards safer borrowers, and raise interest rates for small and medium-sized firms, respectively. Gropp et al. (2019) document that stress-tested banks reduce risk-weighted assets to meet capital requirements, and Kok et al. (2023) find that banks participating in stress tests reduce credit risk. Unlike these studies, our research establishes a direct link from supervisors' climate stress tests to borrowers' actions to make their business models resilient to climate change via the banks' lending decisions, without triggering capital surcharges. Recently, Acharya et al. (2023) review climate stress scenarios employed by regulators and call for more research to be done in this topic to understand the real implications of climate stress tests.

Third, we advance the literature on how the lending behavior of banks reacts to climate change. A paucity of studies shows that banks respond to information that conveys signals about borrowers' climate-change risk by reducing credit supply, charging higher interest rates, or securitizing loans (Chava 2014; Delis et al. 2024; Anginer et al. 2023; Mueller and Sfrappini 2022; Mueller, Nguyen and Nguyen 2022; Kacperczyk and Peydró 2022; Bruno and Lombini 2023; Nguyen et al. 2022; Correa et al. 2022; Meisenzahl 2023). While our empirical work confirms prior findings that information shocks that signal greater transition risk trigger reductions in credit supply, banks that participate in the climate pilot exercise increase lending. This result is consistent with the view advocated in the policy community that climate stress tests are a learning exercise for banks to understand and assess the climate-transition risk better. The tests inform the banks' business strategies with implications for lending behavior. Our results therefore underscore the beneficial effect of conducting climate stress tests, going beyond their immediate objective of preserving financial stability.

Finally, our work also addresses the scant literature on the role of financial constraints with regard to the propensity of firms to decarbonize their business models. Accetturo et al. (2022) highlight that credit availability is a key impediment to the willingness of borrowers to invest in green technologies. In contrast to their work, we show that credit availability increases as a result of the banks' participation in the climate pilot exercise, underscoring real effects of supervisory efforts to tackle climate change.

The remainder is structured as follows. We describe the institutional setting in Section 4.2 and illustrate empirical implications in Section 4.3. Section 4.4 describes the data and presents summary statistics. Section 4.5 describes our empirical strategy, Section 4.6 discusses results, and Section 4.7 presents robustness checks. We draw conclusions in Section 4.8.

4.2 Institutional Background

4.2.1 The French Climate Pilot Exercise

The climate pilot exercise in France, conducted between July 2020 and April 2021, is the first one of its kind. Its findings inform activities by various other central banks and international bodies concerning climate change. The main objectives of the pilot climate exercise are to boost the banks' and insurance companies' understanding of climate-change risks and to strengthen the ability to anticipate and manage such risks in the long run. Another benefit is to identify gaps in terms of data availability related to climate change. Contrary to financial stability stress tests, the pilot exercise does not establish the solvency of the participating institutions. Therefore, the exercise cannot be failed. It also does not trigger regulatory capital requirements, and no bank-specific results are published. These characteristics of the climate pilot exercise avoid the regulators' reputation-building behavior in traditional stress tests documented by Shapiro and Zeng (2024), which result in soft or tough stress test regimes that trigger corresponding changes in the banks' lending behavior.

Appendix C, Section 4.9.2, provides information on the nine bank groups in France that participated in the climate pilot exercise. The intention of this pilot exercise is to raise awareness of physical and transition risks among financial institutions. However, the exercise uncovered a lack of data concerning physical risk, which requires modeling the impact of rising temperatures between 1.4 and 2.6°C by 2050. One problem arises from the lack of location information of funded or collateralized retail and corporate properties. A further problem arises from lack of data on the location of businesses' production sites and value chains. Both these problems resulted in a focus on the banks' exposure to transition risk in the pilot exercise.[4] The French setting is therefore particularly well-suited for our analysis that centers on the borrowers' environmental risk profiles that convey information about transition risk.

To establish the effects of transition risk, the climate pilot exercise required banks to simulate three different scenarios based on recommendations by the Network for Greening the Financial System (NGFS) and described in detail in Section 4.9.1 of Appendix C. The scenarios concentrate primarily on the evolution of carbon prices over a 30-year period from 2020-2050. Although carbon prices are the main drivers of the transition (Bolton and Kacperczyk 2023), and climate stress tests focus on them, prices of other non-renewable energy sources such as oil, gas, and coal, and any industry using these sources, are affected by them (European Central Bank 2022). Therefore, carbon prices have vast-ranging implications for banks and their borrowers. In particular, they affect the long-term viability of borrowers' business models, their creditworthiness, and the values of assets and collateral (Baudino and Svoronos 2021).

The French climate pilot exercise is forward-looking, follows a bottom-up approach, and combines qualitative and quantitative approaches. The qualitative aspect of the climate pilot exercise highlights the learning dimension for banks and supervisors. Throughout the duration of the exercise, the participating institutions took part in Q&A sessions, culminating in bilateral interviews and feedback sessions that helped clarify, refine, and correct risk assessments and issues related to methods, data, reporting consistency, and exposures. Moreover, this process improved the banks' understanding of the limits of existing risk-management models, bolstered their comprehension of the role of climate change for business models, and mobilized resources to tackle climate change.

[4] ACPR (2020) states that the banks' assessments of physical risk significantly lacked an analysis of transition risk, reflecting difficulties related to obtaining precise information about the geographical location of their exposures.

The quantitative dimension requires banks to estimate losses they may incur for credit and market risk based on the three transition scenarios; to assess their impact; and to carry out balance-sheet projections. Unlike traditional stress tests that use time frames of three to five years, the French climate pilot exercise takes a long-term perspective from 2020 to 2050 to accommodate better the effects of climate change. It therefore combines a static balance-sheet assumption until 2025 with a dynamic balance-sheet assumption from 2025 to 2050. The former requires projections for banks' credit risk based on changes in carbon prices applied to loan and investment portfolios. The latter involves predicting losses using not only changes in carbon prices, but also changes in the composition of the balance sheets. This allows us to analyze the strategies the banks took in order to mitigate climate risks by becoming able to consider new risks and corrective actions. Another distinct feature of the exercise is its granular focus. While financial-stability stress tests use aggregate asset classes to model expected losses, the climate pilot exercise examines 55 activity sectors to consider heterogeneities across different businesses in the transition to the carbon-neutral economy.

4.3 Empirical Implications

Our goal is twofold. First, we aim to establish how the climate pilot exercise initiated by bank supervisors, with its feedback effects to participating banks, shapes the banks' view of transition risk and affects lending decisions. Second, we wish to estimate the causal effect of bank participation in the climate pilot exercise on their borrowers' environmental performance.

4.3.1 Implications: Bank Lending

Of course, it is plausible to expect that the emphasis of the climate pilot exercise on raising the awareness of banks for climate risks with feedback sessions and bilateral interviews fosters a profound understanding of climate change in participating banks. Therefore, the climate pilot exercise has the potential to motivate banks to reconsider policies and revenue generation in their lending business with borrowers that display high transition risks, which results in either favorable or unfavorable adjustments in loan-contract terms.

The effort of collating data concerning risk exposures generates new and private information that facilitates loan-monitoring, and the availability of such information may also trigger loan reviews. Our argument is nested in theories by Goldstein and Sapra (2014); Dang, Gorton and Holmstrom (2009); and Gorton and Ordonez (2014), according to which stress tests and sudden shocks produce new and unique information. It is also consistent with the theory by Diamond (1984) and corresponding empirical evidence by James (1987) and Lummer and McConnell (1989), which highlight the role of banks for reducing information asymmetries by monitoring borrowers and, importantly, for using such information to renegotiate loan-contract terms.

Moreover, the climate pilot exercise also facilitates information flows with feedback effects for banks, supervisors, and borrowers, and enables a revelation and quantification of hitherto undocumented risks. The exercise also reduces opacity related to transition risks. The interactions between supervisors and banks also spread best practices about assessing and managing climate-change risks. The participation of banks in the climate pilot exercise may also affect employees' attitudes, beliefs, and values concerning climate change. Further, insights about limits of current risk-management models, granular sectoral exposures, insufficient data, and incomplete reporting systems that do not allow assessing climate-change risk may result in additional technology investments and greater sensitivity towards climate-change risk. Prior work reinforces this view. Hirtle, Kovner and Plosser (2020) state that supervisory concerns related to risk management motivate banks to make technology investments. Tarullo (2019) underscores that supervisory expectations related to stress tests encourage banks to upgrade information and risk-management systems, boosting the efficiency of lending decisions and allowing more precise assessments of borrowers' transition risks with a long-term perspective.

The specific nature of transition risk further adds to the complexity of assessing the exposure of borrowers to such risks. Banks need to consider two key aspects. One, they need to form an opinion about the borrowers' ability, willingness, and likelihood to decarbonize their business models, and simultaneously gauge the evolution of carbon-neutral technologies over the maturity of a loan (Bolton and Kacperczyk 2023; Mueller and Sfrappini 2022; Mueller, Nguyen and Nguyen 2022). Two, the fact that banks generate more than 65 percent of their interest income suggests that banks also need to consider the high dependency from, and correlated exposures to, carbon-intensive sectors which entails considerable potential for loan losses during the transition process (European Central Bank 2022). The lending decisions of banks should therefore not only consider current levels of carbon emissions, but should also reflect whether borrowers are able to reduce carbon emissions in the transition process over the long run, consistent with the 30-year horizon of the climate pilot exercise. Related to this concern, in robustness checks, we also use other measurements of the borrowers' exposure to transition risk, such as the index developed by Sautner et al. (2023), which captures opportunities and risks firms face related to climate change, and the Reprisk environmental risk index, which signals whether borrowers are struggling with the transition to the carbon-neutral economy (Duan, Li and Wen 2023).

Against this background, it remains an empirical question whether the reduction in information asymmetries related to the borrowers' transition risk arising from the climate stress test triggers changes in bank-lending behavior.

If the climate pilot exercise shifts the banks' awareness for transition risk towards greater risk-sensitivity; if it increases uncertainty about borrowers' future cash flows from the projects funded by loans, collateral values, and aggravates concerns about stranded assets; then the participating banks may initiate reviews of their lending relationships with high-transition risk borrowers. The new information signals acquired during the climate pilot exercise may highlight a systematic underestimation of transition risk, resulting in reductions of exposures to borrowers with high transition risk and higher risk premiums. Such negative effects from tougher supervision for bank lending have been documented in prior work by Peek and Rosengren (1995) and Ivanov, Kruttli and Watugala (2024).

On the other hand, the greater awareness for climate-change risks, with its corresponding investments in better risk-management systems and an evolving culture towards helping borrowers in the transition to the carbon-neutral economy, may dominate the greater risk sensitivity for these risks. To the extent that the reduction in information asymmetries triggered by the climate pilot exercise results in a favorable updating of the banks' beliefs about the borrowers' ability to adjust to the carbon-neutral economy, the banks may expand lending to such borrowers, potentially at lower loan rates. Supervision could therefore, in line with Chaly et al. (2017) contribute to a stable provision of financial services.

These two countervailing effects will only be reflected in the data as long as other factors, such as resource constraints, executives' personal views on climate change and short-term incentives that shape the banks' lending policies, concerns about inflating green bubbles, long-term relationships with high transition risk borrowers, and legacy assets do not interfere with and mute the information signals gleaned during the climate pilot exercise. Another factor that may dampen the effect of the climate pilot exercise is that higher exposures to climate risks do not attract regulatory capital surcharges. Our empirical estimates will pick up the net effect of these competing forces.

4.3.2 Implications: The Borrowers' Environmental Performance

We next turn to the effect of banks participation in the climate pilot exercise on their borrowers' environmental performance. Answering this question illuminates a key issue in the debate on climate change – whether the banking sector, and bank supervision more specifically, can help the transition to the carbon-neutral economy.

A widely accepted view among economists is that supervision imposes costs and constraints on banks (Bernanke 2006). Even in the absence of capital requirements, as in our setting, these costs and constraints transmit via the banks' lending decisions to the real economy (Ivanov, Kruttli and Watugala 2024). Costs arise from investments in data collection related to climate-change risk, enhancements of information and risk-management systems, and, importantly, a review of exposures motivated by revisions of the estimates on credit and market risk during the transition process. Constraints come in the form of the banks' greater awareness for climate-change risks reflected in higher expectations and pressure on borrowers to decarbonize their business models, and the banks' anticipation of future capital requirements against climate-related losses that result in reductions in credit supply. In response, it is plausible to expect that borrowers from banks participating in the climate pilot exercise try to and are encouraged to boost environmental performance.

Whether borrowers from banks participating in climate stress tests indeed boost environmental performance, however, remains an open question. It is equally plausible that borrowers face formidable obstacles and impediments in the transition to the carbon-neutral economy and therefore make little or no effort to make their business models resilient to climate change. Potential challenges range from executives' short-term incentives, who delay restructuring business models and shy away from investments that deplete earnings in the short run, to lack of control of supply chains, immaturity with regard to carbon-neutral technologies and infrastructure, and to industry-specific reasons where the transition to net zero is difficult to achieve, e.g., in coal mining.

4.4 Data and Descriptive Statistics

We combine several different datasets for this research. We start by manually collecting the list of banks that participate in the French climate pilot exercise conducted by the ACPR from the Banque de France. The climate stress test takes place at parent or headquarter level. We carefully check each bank's name and location details to identify these banks.

Participation in the climate exercise is described in official documents by the ACPR as "voluntary". Despite this official stance, our frequent discussions within the policy community suggest that participation for these nine banks was effectively mandatory. This discrepancy underscores pressure on banks to comply, regardless of the formal presentation of voluntariness. Appendix C, Section 4.9.2, provides an overview of the nine participants in the climate pilot exercise.

To understand whether bank participation in the climate pilot exercise affects borrowers' actions to decarbonize their business models, we establish a link between banks and their borrowers via lending activities. We retrieve data on loan contracts from Thomson Reuters LPC's Dealscan. We include all Euro-denominated syndicated loans provided by all French and non-French banks extended to French borrowers between 2015 and 2023. Syndicated loans are well-suited for our analysis because Gustafson, Ivanov and Meisenzahl (2021) show that such loans are actively monitored with lead banks demanding information from borrowers on a regular basis. We exclude SIC codes from 6000 to 6999 to remove financial firms, and focus on lead arranger(s) following the approach used by Ivashina (2009). Participants are excluded from our sample because lead arrangers play the key role in setting and negotiating loan terms with borrowers before turning to participant lenders that can be characterized as passive investors (Correa et al. 2022).

Our unit of observation to test bank-lending behavior is the loan level. We allocate a loan into the treatment group if the name(s) of the participating bank(s) in the climate pilot exercise matches the name of the lead arranger(s) in the Dealscan data. The control group consists of loans provided by French banks that did not participate in the climate pilot exercise, and banks headquartered outside France that cannot participate in the climate pilot exercise, but supply credit to French borrowers. The benefit of this setup is that we can compare borrowers operating in the same macroeconomic environment that differ in terms of their lenders' awareness and ability to comprehend and assess risks arising from climate change.

We further augment the loan-level data with bank characteristics using the Dealscan-Compustat link from Schwert (2018) for the period from 2015 to 2020 and manually check lenders that appear in the sample in the later period. Borrower characteristics are extracted from Compustat Global by manually checking all borrowers' names to identify their GVKEYs and ISINs.

For carbon emissions and environmental performance, we merge our loan-level data with Refinitiv. In a robustness check, we also use the environmental risk index from RepRisk and climate-risk exposures, as in Sautner et al. (2023), as alternative measurements of transition risk. Our final sample for the loan-level analyses consists of 1,673 unique loans that have information on loan amount, spread, borrower carbon emissions, and borrower characteristics.

Table 4.1 reports summary statistics for our main variables, and Section 4.9.3 of Appendix C shows variable descriptions for 1,673 French loans in our sample. Our sample consists of 45% of loans originated by banks that participated in the climate stress tests. The average loan amount granted to French borrowers over the sample period is 618 million USD, with an average maturity of 5 years and an average loan spread of 150 basis points.

Figure 4.1 shows average carbon emissions across eight industries. Mining, oil, and gas, followed by transportation and utilities have the highest carbon emissions (9.97 million tons, and 6.47 million tons of carbon dioxide per industry, respectively). On the contrary, on average, wholesale trading and services have the lowest carbon emissions.

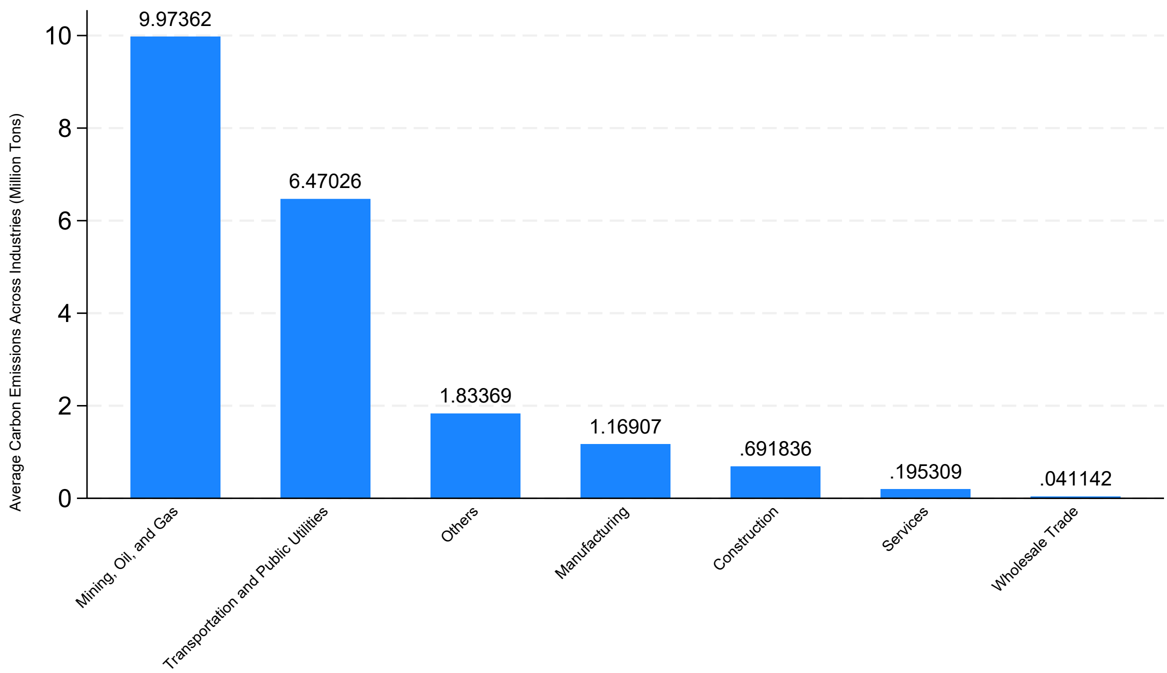

Figure 4.1: Average Carbon Emissions Across Industries

A further goal of this research is to compare the borrowers' environmental performance conditional on their banks' participation in the climate pilot exercise. For this purpose, we retrieve detailed data from Refinitiv for 2015 to 2023 on 'short-term' and 'long-term' dimensions of borrowers' environmental profiles. As Refinitiv only provides annual information on the firms' environmental performance, we aggregate information from syndicated loans to the firm-year level to observe whether a firm gets at least one loan from a participating bank at year *t-1*, and merge this information into the borrower-year level information from Refinitiv. We focus on whether borrowers have eco-friendly products, experience changes in ESG scores, environmental and emission scores, have emission policies, set emission reduction targets in their production process and view them as short-term performance dimensions because such dimensions are likely to reflect the borrowers' efforts to address climate change in the short run. In contrast, we classify the use of renewable energy over total energy sources, total emissions growth, direct emissions growth, the probability of having supply-chain environmental policies, terminations of contracts with suppliers who are considered to be environmentally unfriendly, as well as having environmental criteria for material sourcing as longer-term dimensions, as it may take longer before one can observe these changes.

Table 4.1 illustrates substantial heterogeneity across the borrowers' environmental performance. While 85% of firms have carbon-emission policies, only 63% have a target concerning the figure by which carbon emissions should be reduced by 2050. On average, firms in our sample have an ESG score of 62, an Emission score of 78, total carbon-emissions growth of 1.6%, a direct carbon-emissions growth of 1.8%, and 32% of our firms terminate contracts with suppliers that are considered environmentally unfriendly. Our final data set for the analyses of the borrowers' environmental performance results in 843 observations from seven industries between 2015 and 2023.

Table 4.1: Summary Statistics

Variable	Mean	SD	Min.	Max.	N
Loan-level data					
Loan Amount (Ln)	6.43	1.30	1.89	8.74	1,673
Spread (Ln)	5.01	0.87	3.22	6.55	1,673
All In Spread Drawn (bps)	150.14	136.42	25.00	600.00	1,673
Treat	0.45	0.50	0.00	1.00	1,673
Post	0.41	0.49	0.00	1.00	1,673
Maturity	4.62	1.88	0.17	14.00	1,673
High Emitter	0.22	0.42	0.00	1.00	1,673
Carbon Emission (Ln)	7.20	6.19	0.00	15.02	1,673
High Climate Change Exposure	0.36	0.48	0.00	1.00	1,673
High Reprisk ERI	0.79	0.41	0.00	1.00	1,673
Borrower Size	14.15	6.71	4.82	25.07	1,673
Borrower Leverage (%)	14.75	19.90	0.01	58.83	1,673
Borrower ROA (%)	1.91	2.93	-1.00	15.59	1,673
Bank Size	20.34	0.96	15.77	28.64	1,673
Bank Equity (%)	8.73	6.23	2.75	78.51	1,673
Bank ROA (%)	0.51	0.37	-0.59	5.13	1,673
Green	0.12	0.33	0.00	1.00	1,673
Green Share	0.13	0.30	0.00	1.00	749
SA Index	-3.84	0.38	-4.39	-2.57	1,425
Bank-level data					
Mentioning Climate Stress Test	0.007	0.084	0	1	1,125
Communication with Borrowers	0.005	0.073	0	1	1,125
Discussion about Emission Data	0.173	1.092	0	18	1,125
Firm-level data					
Treat	0.56	0.50	0.00	1.00	943
Post	0.40	0.49	0.00	1.00	943
Eco-Friendly Product	0.29	0.45	0.00	1.00	943
ESG Score	0.62	0.17	0.02	0.91	943
Environmental Score	0.69	0.22	0.00	0.99	943
Emission Score	0.78	0.23	0.00	1.00	943
Emission Policies	0.85	0.36	0.00	1.00	943
Target Emission	0.63	0.48	0.00	1.00	943
Renewable Energy	0.12	0.20	0.00	0.90	943
Total Emission Growth (%)	1.60	26.94	-50.33	142.30	943
Direct Emission Growth (%)	1.88	23.90	-50.46	114.91	943
Supply Chain Policy	0.77	0.42	0.00	1.00	943
Termination of Env. Unf. Suppliers	0.32	0.47	0.00	1.00	943
Materials Sourcing Criteria	0.65	0.48	0.00	1.00	943

Note: This table reports the summary statistics for the variables used in Equation 1. The initial sample consists of 1,673 loan observations between 2015 and 2023 from the DealScan database matched with borrower financial information from Compustat Global and borrower environmental performance from Refinitiv, Reprisk, and Sautner et al. (2023). The latter part of the table shows the variables on soft and hard dimensions of the firms' environmental profiles. Appendix C, Section 4.9.3, provides the variable definitions in detail.

4.5 Identification Strategy

4.5.1 Borrowers' Transition Risk and Bank Lending

We start with a simple model that explores the relationship between the banks' lending behavior and the borrowers' carbon emissions in the absence of participation in the climate pilot exercise for the period between 2015 and 2023. Results from this initial analysis inform us about how banks decide on credit supply and loan pricing depending on public information about the firms' exposure to transition risk without the influence of the climate pilot exercise.

$$Y_{lbft} = \beta \times CarbonEmissions_{f,t-1} + \gamma F_{ft} + \theta L_{lbft} + \delta_b + \delta_{it} + \delta_l + \varepsilon_{lbft}, \tag{4.1}$$

where Y_{lbft} is the loan volume or loan spread for a given loan by bank b to a borrower f at time t. $CarbonEmissions_{f,t-1}$ is the natural logarithm of total carbon emissions (measured in tons) of firm f the year before; γF_{ft} is a vector of quarterly borrower characteristics including firm size, and firm leverage; θL_{lbft} is loan maturity. For regressions with loan volume as a dependent variable, we also control for loan spread, and conversely control for loan volume when loan spread is the dependent variable.

We include bank-fixed effects, δ_b, to capture bank-specific time-invariant effects; δ_l are loan-type fixed effects to ensure that our results do not reflect differences in loan-contract features, such as whether a loan is revolving or a term loan. In addition, industry-year fixed effects δ_{it} capture differences in loan demand across different industries and industry characteristics over time during our sample period. A further benefit of including the interaction of industry- and year fixed effects is that our sample period coincides, at least partially, with the COVID-19 pandemic (which affects services and manufacturing in particular), rising inflation rates, and the war in Ukraine (that affects energy, oil, and gas). These factors have varying effects on different industries that are absorbed by these fixed effects. ε_{lbft} is the idiosyncratic error term. We double-cluster standard errors at the bank and borrower level to reflect that participation in the climate pilot exercise is at the bank level, but carbon emissions are measured on the borrower level. The main coefficient of interest is β, which identifies whether the banks change loan volume or spread if the borrowers' carbon emissions change.

4.5.2 Difference-in-Difference-in-Differences Specification

The ideal setup to establish the causal effects of the climate pilot exercise on bank lending and its corresponding effects on borrowers' environmental performance assigns the climate pilot exercise to banks in a random fashion. The institutional setup with participation in the climate exercise, officially described as voluntary, but characterized in the policy community as de facto mandatory, constitutes an empirical challenge. Banks could be nudged to participate in the climate exercise for reasons that may correlate with their lending policies and the composition of the loan portfolio.[5]

Our most feasible empirical approximation to generate plausibly exogenous variation in the assignment of the climate pilot exercise is therefore to compare the participating (treatment group) banks with French and non-French (control group) banks that cannot participate in the exercise, but also provide credit to borrowers in France. The composition of this control group mitigates concerns that foreign banks retrench to their home countries during a shock like the Ukraine crisis and that such behavior (Giannetti and Laeven 2012) interferes with our key coefficients of interest.

Having restricted our sample to participating French banks and non-participating French and foreign banks, we apply a triple difference strategy. Ultimately, we are interested in the causal relationship between the French climate pilot exercise and the banks' lending behavior towards borrowers with different levels of transition risk reflected in their carbon emissions. We identify this relationship with the following equation:

$$\begin{aligned} Y_{lbft} = {} & \beta_1 \times HighEmitter_f \times Post_t \times Treated_b + \\ & + \beta_2 \times HighEmitter_f \times Post_t + \beta_3 \times HighEmitter_f \times Treated_b \\ & + \beta_4 \times HighEmitter_f + \beta_5 \times Post_t \\ & + \gamma F_{ft} + \theta L_{lbft} + \delta_b + \delta_l + \delta_{it} + \varepsilon_{lbft} \end{aligned} \tag{4.2}$$

[5] To mitigate concerns about a possible selection problem embedded in the participation of the climate exercise, in Table C.3 of Appendix C Section 4.9.4, we present a Heckman selection model (Heckman 1979), where our first stage models participation in the pilot exercise using a dummy variable *Green lender* that takes on the value of one if a bank signed the UN Principles for Responsible Banking (0 otherwise) prior to the climate pilot exercise before 2020. This predictor variable can be plausibly excluded from the second stage because being a signatory has no bearing on a bank's overall lending activity. Table C.3 shows that the green lender dummy is significant and intuitively related to participation in the climate exercise. However, the inverse Mills ratio remains insignificant, supporting our argument that no selection problem exists.

where $Post_t$ is a dummy variable equal to 1 for the period after the French climate climate pilot exercise (2020Q3 onwards), 0 otherwise; $Treated_b$ is a dummy taking on the value 1 for a bank participating in the climate pilot exercise, 0 otherwise; all other variables are identical as in Equation 4.1, except for $HighEmitter_f$, which is a dummy variable equal to 1 if the average carbon emissions of borrower f before 2020 are above the median, and 0 otherwise. Using pre-shock measurement of the borrowers' carbon emissions allows us to capture the direct effect of the climate pilot exercise rather than the change in the firms' risk exposure. Thus, our main coefficient of interest is now β_1, which indicates whether banks that participate in the climate pilot exercise change loan volume or spread for higher carbon emitters compared to lower carbon emitters, holding everything else constant.

Last, using annual borrower-level information on their environmental performance from Refinitiv, we explore the relationship between the climate pilot exercise and changes in the borrowers' environmental performance. We use the following specification:

$$\begin{aligned} Y_{ft} = {} & \beta_1 \times HighEmitter_f \times Post_t \times Treated_{f,t-1} \\ & + \beta_3 \times HighEmitter_f \times Post_t + \beta_4 \times HighEmitter_f \times Treated_{f,t-1} \\ & + \gamma F_{ft} + \alpha_f + \tau_t + \varepsilon_{ft} \end{aligned} \tag{4.3}$$

where Y_{ft} captures either short-term adjustments for environmental performance such as having eco-friendly products, having emission policies, having emission targets, ESG scores, environmental and emission scores, or longer-term adjustments such as the share of renewable energy, total emissions growth, direct emissions growth, having supply-chain environmental policies, termination of environmentally unfriendly suppliers, materials sourcing environmental criteria of borrower f at time t; $Treated_{f,t-1}$ is a dummy taking on the value 1 if borrower f received any loan from a participating bank the year before, 0 otherwise; $HighEmitter_f$ is a dummy variable equal to 1 if the average carbon emissions of borrower f before 2020 is above the median, and 0 otherwise.; γF_{ft} is a vector of borrower characteristics including firm size and leverage; α_f and τ_t are firm- and time fixed effects, respectively.

4.5.3 Parallel Trends

A causal interpretation of the parameters in Equation 4.2 relies on the parallel trends assumption. This assumption states that, in the absence of the climate pilot exercise, participating banks and non-participating banks provide loans to borrowers of similar environmental risk profiles and their characteristics evolve in similar fashions. We examine the evolution of loan, firm, and bank characteristics using *t*-tests.

First, we ask whether changes in bank lending and interest rates from treated and control banks differ before the climate pilot exercise. If they do, one may be worried about differences in business models between these groups of banks. Table 4.2 shows loan volumes and interest rates of treated and control banks are similar before the climate pilot exercise. Likewise, changes in the share of high carbon-emitting borrowers linked to the two groups of banks are similar prior to the climate pilot exercise. We also compare other bank and firm characteristics (size, equity ratio, leverage ratio, and profitability) and find that these characteristics evolved in similar patterns before the climate pilot exercises.

Table 4.2: Comparisons Between Treated and Control Banks

Variable	Mean Treated (1)	Mean Control (2)	Diff. (3)	*t*-stat (4)
Loans Characteristics				
Δ Loan Amount (Ln)	-0.043	0.084	-0.126	-0.49
Δ Spreads (Ln)	0.099	0.186	-0.087	-0.48
Banks' characteristics				
Δ Share of High-Emitting Borrowers	0.024	0.008	0.016	0.15
Δ Bank Size	0.657	-0.222	0.881	0.98
Δ Equity/Total Assets	-0.093	-0.403	0.310	0.29
Δ Loans Growth (%)	-0.338	-0.391	0.052	0.06
Δ ROA	-0.018	-0.022	0.004	0.05
Firms' characteristics				
Δ Firm Size	0.657	-0.222	0.880	0.98
Δ Leverage	-0.001	0.230	-0.232	0.05
Δ ROA	0.103	0.044	0.059	0.12

Note: This table reports statistics of relevant variables over the period 2015 to 2020 dividing the sample between treated and control banks. Column 1 reports the mean of changes in characteristics among treated banks. Column 2 reports the mean of changes in characteristics among control banks. All loan, firm, and bank characteristics are reported as first differences. For loan and firm characteristics, we aggregate the mean of these variables to the bank-year level. In Column 3, we report the mean differences of changes in characteristics of treated and control banks. Column 4 reports *t*-statistics.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Next, we ask whether changes in bank lending and interest rates to high emitters compared to low emitters from treated and control banks differ before the climate pilot exercise. The concern here would be that our results may be driven purely by green preferences of these two groups of banks. Following the convention in the literature, we test this assumption by inspecting the dynamic effects of the climate pilot exercise on lending to high vs. low emitters for all years before the exercise. In Table 4.3, we interact $HighEmitter$ and $Treat$ with a set of yearly dummies between 2015 and 2019 and find that treated banks and control banks lend and price loans in similar patterns for *High Emitters* compared to *Low Emitters*.

Table 4.3: Comparisons Between Treated and Control Banks by Emitter Status

	Loan Amount (Ln) (1)	Spread (Ln) (2)
Treat × High Emitter × Y2015	0.041	0.046
	(0.113)	(0.074)
Treat × High Emitter × Y2016	-0.127	0.022
	(0.168)	(0.069)
Treat × High Emitter × Y2017	-0.424	-0.042
	(0.335)	(0.124)
Treat × High Emitter × Y2018	-1.387	-0.128
	(0.914)	(0.191)
Treat × High Emitter × Y2019	0.000	0.053
	(0.123)	(0.064)
Observations	992	992
Loan Controls	Yes	Yes
Firm Controls	Yes	Yes
Bank FE	Yes	Yes
Loan Type FE	Yes	Yes
Industry × Year FE	Yes	Yes
Adjusted R^2	0.872	0.888

Note: This table tests for parallel trends in loan amounts and spreads given to high emitters compared to low emitters between 2015 and 2019. $Y2015$, $Y2016$, $Y2017$, $Y2018$, and $Y2019$ are dummy variables equal to 1 if the loan is originated in the year 2015, 2016, 2017, 2018, and 2019, respectively, and 0 otherwise. Standard errors are double-clustered at the bank and borrower level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.6 Results

Our first test focuses on how banks react to their borrowers' carbon emissions. Next, we evaluate the effect of the climate pilot exercise on bank lending to brown borrowers. As part of this analysis, we also explore heterogeneous adjustments by inspecting the role of green banks and ownership structures. A further analysis explores whether banks indeed produce new information during the climate pilot exercise. The final set of tests explores whether the climate pilot exercise has real effects and whether lending by participating banks triggers behavioral changes among their borrowers.

4.6.1 Bank Lending and Firms' Exposure to Transition Risk

Table 4.4 reports the results from estimating Equation 4.1 using data between 2016Q1 and 2020Q2 when no climate pilot exercise had taken place yet.

Columns 1 and 2 show the effect of the firms' carbon emissions on loan volumes (in natural logs), whereas Columns 3 and 4 focus on loan spreads (in natural logs). In Column 1 and 3, we perform the estimation with loan characteristics, bank fixed effects, loan-type fixed effects, and industry-year fixed effects without controlling for firm characteristics. We include a vector of borrower control characteristics including size, leverage, and return on assets (ROA) in Columns 2 and 4. The result from this exercise illustrates that in the absence of the climate pilot exercise, there is no evidence that banks limit their exposure to high transition firms by reducing credit supply or increasing loan rates.

Table 4.4: How Do Banks Respond to the Firms' Carbon Emissions?

	Loan amount (Ln)		Spread (Ln)	
	(1)	(2)	(3)	(4)
Carbon Emission (Ln)	0.066	0.068	-0.002	-0.005
	(0.042)	(0.042)	(0.023)	(0.024)
Maturity	0.016	0.018	0.063	0.066
	(0.103)	(0.106)	(0.052)	(0.049)
Borrower Size		0.006		0.017
		(0.014)		(0.014)
Borrower Leverage		-0.001		-0.001
		(0.004)		(0.003)
Borrower ROA		-0.028		-0.026
		(0.018)		(0.016)
Loan Amount (Ln)			-0.236**	-0.232***
			(0.088)	(0.083)
Loan Spread (Ln)	-0.722***	-0.724***		
	(0.134)	(0.139)		
Observations	992	992	992	992
Bank FE	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes
Loan Type FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.874	0.875	0.888	0.891

Note: This table shows the relationship between the banks' lending behavior and the firms' total carbon emissions. *Loan Amount (Ln)* and *Spread (Ln)* are dependent variables. *Carbon Emissions (Ln)* is the firm total carbon emissions from Refinitiv database. Standard errors are double-clustered at the bank and borrower level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.6.2 Bank Lending after the Climate Pilot Exercise

We now turn to our analysis that focuses on how participation in the climate pilot exercise affects lending to high transition risk firms compared to low-transition risk firms.

We estimate Equation 4.2 and report the results in Table 4.5. Columns 1 and 2 report on loan volumes (in natural logs), and Columns 3 and 4 examine loan spreads (in natural logs). We control for loan characteristics, bank fixed effects, loan-type fixed effects, and industry-year fixed effects in all specifications. Additionally, we control for borrower characteristics in Columns 2 and 4, our preferred specifications.

The estimates for our coefficient of interest, β_1, are significant and positive for both dependent variables. Following the climate pilot exercise, participating banks increase loan volumes significantly by 38% for high carbon emitters. They also significantly increase loan spreads by 8% (equivalent to 12 bps), ceteris paribus. This result indicates that banks adjust their risk-pricing to reflect the greater transition risk in sticking with brown borrowers.

We do not view our results to contradict prior findings by Kacperczyk and Peydró (2022) that banks reduce credit for high transition risk firms. In contrast, we propose that the climate pilot exercise with its long-term horizon changes the banks' risk perspective. Instead of immediately reducing exposure to transition risk, participating banks may want to aid borrowers in the transition towards greener activities. Given their exposure to potential financial losses in future if borrowers fail to adopt their business models for the carbon-neutral economy, they stick with these firms and provide larger loan volumes. To compensate for the greater risk, they in turn demand higher spreads.

Table 4.5: Climate Pilot Exercise and Bank Lending to Brown Firms

	Loan amount (Ln)		Spread (Ln)	
	(1)	(2)	(3)	(4)
Treat × High Emitter × Post	0.390**	0.380**	0.082**	0.080**
	(0.190)	(0.186)	(0.031)	(0.033)
Treat × High Emitter	-0.326*	-0.321*	-0.050*	-0.049*
	(0.179)	(0.176)	(0.027)	(0.028)
High Emitter × Post	-0.039	-0.039	-0.034***	-0.033**
	(0.059)	(0.059)	(0.009)	(0.013)
High Emitter	-0.256	-0.293	-0.394**	-0.419**
	(0.520)	(0.512)	(0.163)	(0.167)
Spread (Ln)	-0.521***	-0.524***		
	(0.148)	(0.151)		
Maturity	-0.141	-0.140	0.103*	0.106**
	(0.115)	(0.117)	(0.052)	(0.050)
Borrower Size		0.004		0.011
		(0.009)		(0.010)
Borrower Leverage		-0.001		-0.002
		(0.003)		(0.003)
Borrower ROA		-0.027*		-0.015
		(0.014)		(0.010)
Loan Amount (Ln)			-0.152**	-0.152**
			(0.074)	(0.073)
Observations	1,673	1,673	1,673	1,673
Bank FE	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes
Loan Type FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.855	0.856	0.906	0.907

Note: This table shows the effect of participation in the climate pilot exercise on the banks' lending behavior. *Loan Amount (Ln)* and *Spread (Ln)* are dependent variables. $Treated$ is a dummy taking on the value 1 if a bank participates in the climate pilot exercise and 0 otherwise. *High Emitter* is a dummy variable equal to 1 if the average carbon emissions of borrower f before 2020 is above the median, and 0 otherwise. $Post$ is a dummy variable equal to 1 for the period after the French climate pilot exercise (2020Q3 onwards), 0 otherwise. Standard errors are double-clustered at the bank and borrower level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.6.3 Do Participating Banks Aid the Transition to Net Zero?

Next, we ask whether banks involved in the climate pilot exercise are more willing to help the transition to net zero. To this end, we first investigate whether the participating banks provide loans with longer maturity to their borrowers after the climate pilot exercise. The rationale behind this test is that green investments may take time, and if loans given to brown borrowers have longer maturity, they have a higher chance to be loans with positive impacts. Given the average maturity of approximately five years of the syndicated loans in our sample, we split the data into loans with a maturity of three years or less. Columns 1 to 4 of Table 4.6 illustrate that our findings are driven by loans with maturities above 3 years, suggesting that the lending activities that are of greater relevance for the green transition have longer maturities.

Second, we examine whether the climate pilot exercise leads to a higher likelihood of banks providing green loans for their brown borrowers. We define a loan that is green if it is a sustainability-linked loan where the loan-contract terms indicate costs borne by borrowers for funding change depending on future environmental performance, or, alternatively, if the loan is originated to fund energy efficiency projects. Typical examples for sustainability-linked loans are ones with higher interest rates if borrowers do not commit to net-zero scientific targets. Typical examples for loans with green purposes are ones to fund investment in windmills, production of solar panels, and energy-efficient products. We obtain this information by performing textual analysis on tranches and deal remarks from the Dealscan data.

Columns 5 and 6 of Table 4.6 examine the likelihood of a green loan being originated. The two columns differ in their inclusion of control variables for borrower characteristics. After the climate pilot exercise, treated banks are 7.6 to 9.1 percentage points more likely to provide a green loan for high carbon emitters. Given that 12% of our loans are classified as green loans, the magnitude of the effect is big and equivalent to a 63% increase in the probability that a bank provides a green loan to aid its brown borrowers in the transition to net zero. In Columns 7 and 8 that again differ in terms of the inclusion of borrower control variables, we aggregate the data to the bank-quarter level to test whether the climate pilot exercise provokes participating banks to supply more green credit to brown borrowers. We find that the share of green loans over total loans that participating banks provide increases by 4.3 to 4.6 percentage points after the pilot exercise, suggesting some rebalancing of the loan portfolio towards green lending.

Table 4.6: Does the Climate Pilot Exercise Aid the Green Transition?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Loan Amount (Ln)		Spreads (Ln)		P (Green)		Green Share	
Sample	> 3Y	≤ 3Y	> 3Y	≤ 3Y	All	All	All	All
Treat × High Emitter × Post	0.379**	-0.228	0.107***	0.021	0.076*	0.091*	0.046**	0.043**
	(0.180)	(0.145)	(0.039)	(0.026)	(0.041)	(0.051)	(0.022)	(0.018)
Treat × High Emitter	-0.259*	0.193	-0.067**	-0.018	-0.025	-0.041	-0.026**	-0.023**
	(0.154)	(0.137)	(0.032)	(0.024)	(0.025)	(0.034)	(0.012)	(0.011)
High Emitter × Post	0.257	0.000	0.167	0.000	0.015	0.068	-0.112	-0.180**
	(0.880)	(0.000)	(0.208)	(0.000)	(0.154)	(0.145)	(0.075)	(0.069)
Treat × Post	-0.013	0.278	-0.004	-0.026	-0.015	-0.008	-0.008	-0.009
	(0.037)	(0.187)	(0.011)	(0.032)	(0.014)	(0.014)	(0.006)	(0.010)
High Emitter	-0.797	0.000	-0.324*	0.000	0.180	0.197*	-0.035	-0.024
	(0.533)	(0.000)	(0.178)	(0.000)	(0.113)	(0.103)	(0.033)	(0.034)
Post	0.000	0.000	0.000	0.000	-0.276***	-0.287***	-0.712***	-0.689***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.097)	(0.084)	(0.054)	(0.091)
Observations	1,288	408	1,288	408	1,673	1,673	749	749
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan Type FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Industry FE	No	No	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes	No	No	No	No
Loan Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes	No	Yes	No	Yes

Note: We investigate whether the climate pilot exercise aids the transition to net zero. Columns 1 and 3 show the effect of the exercise on loan volumes and spreads for brown borrowers compared to green borrowers for loans with maturity of more than three years. Columns 2 and 4 show the effect of the climate pilot exercise on loan volumes and spreads for loans with maturity of three years and less. Columns 5 and 6 show the effect of the exercise on the probability that a bank originates a green loan to brown borrowers. Column 5 excludes control variables for firm characteristics. Column 6 includes these borrower control variables. Columns 7 and 8 show the effect of the exercise on the share of green loans to total loans a bank originates. Column 7 excludes control variables for borrower characteristics. Column 8 includes these controls. In Columns 1 to 6 we run regressions at the loan level. In Columns 7 and 8, we aggregate data at the bank-year level to obtain the share of green loans to total lending. $Treated$ is a dummy taking on the value 1 if a bank participates in climate stress tests, and 0 otherwise. *High Emitter* is a dummy equal to 1 if the average carbon emissions of borrower f before 2020 is above the median, and 0 otherwise. $Post$ is a dummy variable equal to 1 for the period after the French climate pilot exercise (2020Q3 onwards), and 0 otherwise. Standard errors are double-clustered at the bank and borrower level for Columns 1-6 and at the bank level for Columns 7 and 8 and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.6.4 Further Heterogeneities

Next, we explore additional heterogeneities in the data. It is plausible to assume that banks that are signatories of the UN Principles for Responsible Banking prior to their participation in the climate exercise are more willing to support borrowers in the transition to net zero. The intuition is that such banks already have a positive predisposition towards being green relative to banks that did not sign up to these principles.

Table 4.7 supports this assertion. The tests in Columns 1 and 3 highlight that our results are driven by signatories of the UN Principles for Responsible Banking. This subsample of banks displays statistically greater responsiveness in terms of lending volumes and loan spreads. In contrast, our key coefficients of interest in Columns 2 and 4 are rendered insignificant.

Likewise, ownership structure may play a role for the observed effects. The remaining tests in Table 4.7 split the sample at the level of institutional investors. Columns 5 and 7 highlight that the effects on loan volumes and spreads are concentrated in banks whose ownership structure is dominated by institutional investors, whereas the effects cannot be observed in Columns 6 and 8. This is consistent with the findings of Krueger, Sautner and Starks (2020), and Ceccarelli, Ramelli and Wagner (2024), which indicate that institutional investors believe that climate risks have financial implications for their portfolios and actively engage in pricing these risks as well as adjusting their holdings towards low carbon stocks.

Table 4.7: Heterogeneity Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Loan Amount (Ln)		Spreads (Ln)		Loan Amount (Ln)		Spreads (Ln)	
Sample	UNEP	Non-UNEP	UNEP	Non-UNEP	High IO	Low IO	High IO	Low IO
Treat × High Emitter × Post	0.550**	0.241	0.107**	0.042	0.712**	0.093	0.105**	0.081
	(0.190)	(0.174)	(0.048)	(0.039)	(0.341)	(0.085)	(0.049)	(0.065)
Treat × High Emitter	-0.502***	-0.167	-0.082**	-0.023	-0.639*	0.013	-0.058	-0.042
	(0.150)	(0.174)	(0.035)	(0.041)	(0.316)	(0.026)	(0.041)	(0.038)
High Emitter × Post	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Treat × Post	-0.140	0.020	-0.051**	-0.037*	-0.113	-0.000	-0.056***	-0.004
	(0.107)	(0.038)	(0.022)	(0.020)	(0.093)	(0.040)	(0.019)	(0.028)
High Emitter	-0.257	-0.333	-0.415**	-0.335*	-0.098	-1.170***	-0.391**	-0.100
	(0.503)	(0.532)	(0.150)	(0.177)	(0.559)	(0.222)	(0.164)	(0.103)
Post	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	644	1037	636	1037	881	806	867	806
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	No	No	No	No	No	No	No
Year FE	No	No	No	No	No	No	No	No
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table explores whether banks that have signed the UN Principles for Responsible Banking and banks with a high proportion of institutional investors affect the magnitudes of the key coefficients of interest. Columns 1 to 4 show the differences between signatories and non-signatories of the UN Principles for Sustainable Banking affecting lending behavior towards their brown borrowers. Columns 5 to 8 show the effect of the banks' institutional ownership (IO) on their lending behavior to brown borrowers. $Treated$ is a dummy taking on the value 1 if a bank participates in climate stress tests, and 0 otherwise. *High Emitter* is a dummy variable equal to 1 if the average carbon emissions of borrower f before 2020 is above the median, and 0 otherwise. $Post$ is a dummy variable equal to 1 for the period after the French climate stress test (2020Q3 onwards), 0 otherwise. Standard errors are double-clustered at the bank and borrower level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.6.5 Climate Pilot Exercise and Information Production

We characterize the climate pilot exercise as an information production exercise, and this new information is crucial to updating beliefs of banks and their borrowers about how climate change affects borrower business models with corresponding knock-on effects on the banks' lending activities.

This subsection provides empirical support to back up this claim. To this end, we retrieve information about conference calls by all banks in our sample from S&P Capital IQ between 2015 and 2023 and manually review 2,322 transcripts. Next, we collapse the data into the bank-quarter level to reflect that large banks often have multiple conference calls per quarter. This results in 1,125 observations. We construct three variables to capture the banks' awareness for the climate pilot exercise. First, our variable *Mention climate stress test* takes on the value of one if the transcript mentions the term 'climate stress test' or 'climate pilot exercise' (0 otherwise). Analogously, the dummy variables for *Borrower Communication on Transition Risk* take on the value of one (0 otherwise) if the transcripts highlight that the bank mentions discussions with their borrower concerning issues related to transition risk. The variable *Discussion about Emissions* is the number of times that banks discussed about carbon emissions in their earning calls. Section 4.9.5 of Appendix C provides some examples of texts that we collected from conference calls related to these issues.

Our results in Table 4.8 reinforce this view. The participating banks display a significantly higher probability of mentioning the climate stress test after the exercise took place. They are also significantly more likely to communicate with their borrowers about transition risks. Finally, they discuss carbon emissions more often in their earning calls.

Table 4.8: Information Production During the Climate Pilot Exercise

	Mentioning Climate Stress Tests (1)	Communication with Borrowers on Transition Risk (2)	Discussion about Emissions (3)
Post	0.031	0.018	-0.073
	(0.028)	(0.026)	(0.060)
Treat × Post	0.045 **	0.097 ***	0.513 *
	(0.022)	(0.017)	(0.281)
Observations	1,125	1,125	1,125
Bank FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adjusted R^2	0.035	0.053	0.041

Note: This table reports how the climate pilot exercise affects the probability of banks discussing issues related to transition risk such as scenarios of climate stress tests, communication with borrowers on transition risk, the number of times that banks discussed about carbon emissions in their earning calls. Data on discussions of these issues are hand-collected from conference calls of all banks in our sample. $Treated$ is a dummy taking on the value 1 if a bank participates in climate stress tests, and 0 otherwise. $Post$ is a dummy variable equal to 1 for the period after the French climate stress test (2020Q3 onwards), 0 otherwise. Standard errors are clustered at the bank level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.6.6 The Firms' Environmental Performance

Our final set of analyses homes in on the question of whether borrowers whose banks changed loan volumes and spreads following the climate pilot exercise changed their behavior in terms of adjusting environmentally relevant dimensions. Table 4.9 reports the results from estimating Equation 4.3. Our tests on the borrower level use annual data, reflecting the frequency of the availability of borrower information related to environmental performance.

We find our coefficient of interest, β_1, is significant and positive for all short-term adjustments. After getting a loan from a participating bank, higher carbon emitters are 25 percentage points more likely to have eco-friendly products, their ESG scores improve by 12 points, environmental scores improve by 15 points, and emission scores improve by 20 points. They are also 23 percentage points more likely to have emission policies and 40 percentage points more likely to have targets for carbon-emission reduction.

Table 4.9: Short-Term Adjustments: Environmental Performance

	(1) Eco-Friendly Product	(2) ESG Score	(3) Env. Score	(4) Emission Score	(5) Emission Policies	(6) Target Emissions
Treat × Post × High Emitter	0.251 *	0.129 **	0.155 *	0.202 **	0.232 **	0.404 **
	(0.145)	(0.059)	(0.091)	(0.094)	(0.097)	(0.200)
Treat × Post	-0.059	0.016	0.046	0.061 **	-0.035	-0.008
	(0.075)	(0.028)	(0.028)	(0.028)	(0.038)	(0.088)
Treat × High Emitter	-0.203	0.009	0.086	0.105	0.120	-0.183
	(0.196)	(0.072)	(0.092)	(0.112)	(0.107)	(0.191)
Post × High Emitter	-0.196 *	-0.110 **	-0.108	-0.157 *	-0.158 *	-0.348 *
	(0.106)	(0.047)	(0.081)	(0.082)	(0.087)	(0.182)
Treat	0.264 **	0.031	0.031	0.022	0.065	0.037
	(0.108)	(0.040)	(0.050)	(0.050)	(0.046)	(0.106)
High Emitter	0.114	-0.040	-0.094	-0.112	-0.147	0.196
	(0.156)	(0.063)	(0.081)	(0.099)	(0.101)	(0.153)
Observations	943	943	943	943	943	943
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.305	0.140	0.221	0.151	0.595	0.244
Number of Firms	151	151	151	151	151	151
Clustering	Firm	Firm	Firm	Firm	Firm	Firm

Note: This table reports regression results for whether a borrower with loans from banks participating in the climate pilot exercise changes environmental performance from a short-term perspective. Short-term adjustments in the borrowers' environmental profiles include *Eco-Friendly Products*, *ESG Score*, *Environmental Score*, *Emission Score*, *Emission Policies*, and *Target Emissions*. The analysis of the borrowers' environmental performance uses annual frequency of the data. $Treated$ is a dummy taking on the value 1 if a borrower has at least one loan from a bank participating in the climate pilot exercise from 2021 onwards, and 0 otherwise. *High Emitter* is a dummy variable equal to 1 if the average carbon emissions of firms f before 2020 is above the median, and 0 otherwise. $Post$ is a dummy variable equal to 1 for the period after the French climate pilot exercise (2021 onwards), 0 otherwise. Standard errors are clustered at the firm level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In contrast, Table 4.10 does not show convincing signs of improvement in dimensions that require longer-term transitioning towards becoming more environmentally friendly or sourcing environmentally friendly materials. While borrowers of participating banks increase 15 percentage points in their usage of renewable energy, they show little or no signs of reducing their total or direct emission growth. Nor do they terminate supply-chain links to environmentally unfriendly suppliers or try to source more environmentally friendly raw materials. A potential explanation for this result is that these dimensions take greater effort and, therefore, take a longer time to show in the data.

Table 4.10: Long-Term Adjustments: Environmental Performance

	(1) Renewable Energy (%)	(2) Total Emission Growth	(3) Direct Emission Growth	(4) Supply Chain Policy	(5) Termination of Env. Unf. Suppliers	(6) Materials Sourcing Criteria
Treat × Post × High Emitter	0.158 *	-2.941	5.172	-0.022	0.076	0.241
	(0.084)	(13.116)	(9.987)	(0.097)	(0.149)	(0.155)
Treat × Post	-0.014	2.091	-2.674	0.142 ***	0.231 **	-0.075
	(0.042)	(6.627)	(3.732)	(0.050)	(0.091)	(0.085)
Treat × High Emitter	0.029	-2.314	-5.117	0.176	0.145	-0.064
	(0.073)	(6.523)	(6.554)	(0.113)	(0.221)	(0.185)
Post × High Emitter	-0.066	-1.843	-8.287	0.027	-0.003	-0.142
	(0.065)	(10.719)	(8.031)	(0.085)	(0.117)	(0.130)
Treat	-0.006	4.142	4.297	0.040	-0.094	0.298 ***
	(0.045)	(3.967)	(2.660)	(0.070)	(0.117)	(0.106)
High Emitter	-0.059	-0.226	3.985	-0.115	-0.104	0.070
	(0.062)	(4.760)	(5.068)	(0.093)	(0.171)	(0.155)
Observations	943	943	943	943	943	943
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.150	0.024	0.011	0.458	0.153	0.330
Number of Firms	151	151	151	151	151	151
Clustering	Firm	Firm	Firm	Firm	Firm	Firm

Note: This table reports regression results for whether a borrower with loans from a bank participating in the climate pilot exercise changes environmental performance from a long-term perspective. Long-term adjustments in the borrowers' environmental profiles include *Renewable Energy Investments (%)*, *Total Emissions Growth*, *Direct Emissions Growth*, *Supply Chain Environmental Policies*, *Termination of Environmentally Unfriendly Suppliers*, and *Materials Sourcing Environmental Criteria*. $Treated$ is a dummy taking on the value 1 if a borrower has at least one loan from a bank taking part in the climate pilot exercise from 2021 onwards, and 0 otherwise. *High Emitter* is a dummy variable equal to 1 if the average carbon emissions of firms f before 2020 is above the median, and 0 otherwise. $Post$ is a dummy variable equal to 1 for the period after the French climate pilot exercise (2021 onwards), and 0 otherwise. Standard errors are clustered at the firm level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.7 Robustness Checks

4.7.1 Falsification Tests

We perform falsification tests to establish that the treatment effects are not observable in the absence of participation in the climate pilot exercise. To do so, we randomly assign banks that did not participate in the climate pilot exercise tests to be participants. Columns 1 and 2 of Table 4.11 show that the key coefficient is rendered insignificant.

Table 4.11: Falsification Tests

	Loan Amount (Ln) (1)	Spread (Ln) (2)
Placebo Treat × High Emitter × Post	-0.030	0.007
	(0.106)	(0.066)
Observations	1,673	1,673
Loan Controls	Yes	Yes
Firm Controls	Yes	Yes
Bank FE	Yes	Yes
Loan Type FE	Yes	Yes
Industry × Year FE	Yes	Yes
Adjusted R^2	0.855	0.907

Note: This table explores the effect of the climate pilot exercise on the banks' lending behavior towards brown firms, but on the basis of a sample that comprises randomly assigned participation in the climate pilot exercise (*Placebo Treat*). *Loan Amount (Ln)* and *Spread (Ln)* are dependent variables. *High Emitter* is a dummy variable equal to 1 if the average carbon emissions of firms f before 2020 is above the median, and 0 otherwise. Standard errors are double-clustered at the bank and the borrower level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.7.2 Alternative Measurements of Transition Risk

While carbon emissions are a standard measurement of carbon-transition risk in contemporary literature (Bolton and Kacperczyk 2023), one critique for using carbon emissions to gauge the borrowers' exposure to transition risk would be that our measurement picks up other firm characteristics, such as size, or how forward-looking the firms are in their estimation of carbon emissions (Aswani, Raghunandan and Rajgopal 2024; Zhang 2024). We alleviate this concern by using carbon-emission intensities, which are calculated as the borrowers' carbon emissions divided by the borrowers' total assets to assign high vs. low carbon emitters. We report the results in Columns 1 and 2 of Table 4.12. We continue to find that participating banks increase loan volumes (15%) and spreads (3%) for high carbon emitters.

In Columns 3 and 4, we consider that carbon emissions do not reflect how advanced borrowers are in the transition to a low-carbon economy. Thus, we employ the exposure to climate transition index by Sautner et al. (2023) to capture net opportunities and challenges that firms face related to climate change. We still find that borrowers with higher exposure to climate change get more credit, albeit at higher prices after their banks participated in the climate pilot exercise.

Finally, in Columns 5 and 6, we use the Reprisk Environmental Risk Index to capture the borrowers' transition risk. Previous literature shows that Reprisk is one of the few sources of ESG data that is not subject to green-washing bias because it relies entirely on negative news coverage by external sources (Berger et al. 2020). Our findings remain intact.

Table 4.12: Robustness Check: Alternative Measurements of Transition Risk

	(1)	(2)	(3)	(4)	(5)	(6)
	Emission Intensity		Exposure to Transition Risk		Reprisk Index	
	Loan Amount (Ln)	Spread (Ln)	Loan Amount (Ln)	Spread (Ln)	Loan Amount (Ln)	Spread (Ln)
Treat × Post × High Transition Risk	0.150 ***	0.031 ***	0.375 ***	0.113 ***	0.154 *	0.095 ***
	(0.047)	(0.020)	(0.126)	(0.040)	(0.077)	(0.027)
Observations	1,673	1,673	1,673	1,673	1,673	1,673
Loan Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry -Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Loan Type FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.858	0.906	0.856	0.915	0.855	0.907

Note: This table reports robustness checks using different measurements for the borrowers' exposure to transition risk including the level of carbon-emission intensity, the exposure to transition risk index by Sautner et al. (2023), and the Environmental Risk Index (ERI) from Reprisk. *Loan Amount (Ln)* and *Spread (Ln)* are dependent variables. $Post$ is a dummy variable equal to 1 for the period after the French climate stress test (2020Q3 onwards), 0 otherwise. Standard errors are double-clustered at the bank and borrower level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.7.3 The Borrowers' Financial Constraints

Heider and Inderst (2022) show that financial constraints may limit the borrowers' ability to fund green projects. One concern would be that our high carbon emitters could also be financially constrained borrowers; thus, the results we observe would be due to differences in the borrowers' characteristics, rather than being causally attributable to participation in the climate pilot exercise. We therefore include a measure of borrower financial constraints, the Hadlock and Pierce (2010) size-age (SA) index, into our regressions. Although firms with greater financial constraints receive less credit at higher interest rates, our key inferences in Table 4.13 remain unaffected.

Table 4.13: Robustness Check: Borrowers' Financial Constraints

	(1) Loan Amount (Ln)	(2) Spread (Ln)
Treat × Post × High Emitter	0.226 **	0.118 ***
	(0.103)	(0.038)
SA Index	-1.044 **	1.229 ***
	(0.499)	(0.147)
Observations	1,425	1,425
Loan Controls	Yes	Yes
Firm Controls	Yes	Yes
Bank FE	Yes	Yes
Industry × Year	Yes	Yes
Loan Type FE	Yes	Yes
Adjusted R^2	0.855	0.949

Note: This table explores the effect of the climate pilot exercise on the banks' lending behavior towards brown firms controlling for the borrowers' financial constraints. We measure the borrowers' credit constraints using the SA index (Hadlock and Pierce 2010). *High Emitter* is a dummy variable equal to 1 if the average carbon emissions of firms f before 2020 is above the median, and 0 otherwise. $Post$ is a dummy variable equal to 1 for the period after the French climate stress test (2020Q3 onwards), and 0 otherwise. Standard errors are double-clustered at the bank and the borrower level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.7.4 Bank Characteristics

One may argue that our results are driven by other bank characteristics such as size, capital, and profitability rather than because of participation in the climate pilot exercise. We tackle this issue by gradually introducing sets of bank characteristics into our main regressions and report results in Table 4.14.

In Column 1, we include bank size (natural logarithm of total assets). In Column 2, we further include capital ratios (total equity capital over total assets). In Column 3, we include ROA. In all instances, our inferences remain unaffected.

Table 4.14: Robustness Check: Bank Characteristics

	Loan Amount (Ln)			Spread (Ln)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post × High Emitter	0.373 **	0.357 **	0.359 **	0.086 **	0.073 **	0.073 **
	(0.181)	(0.176)	(0.174)	(0.035)	(0.034)	(0.034)
Lender Size	-0.013	-0.009	-0.008	0.013 ***	0.016 ***	0.016 ***
	(0.019)	(0.015)	(0.015)	(0.004)	(0.006)	(0.006)
Lender Capital		0.002	0.005 *		0.002 **	0.002
		(0.001)	(0.003)		(0.001)	(0.001)
Lender ROA			-0.047 *			0.002
			(0.027)			(0.013)
Observations	1,673	1,673	1,673	1,673	1,673	1673
Loan Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year	Yes	Yes	Yes	Yes	Yes	Yes
Loan Type FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.856	0.856	0.856	0.907	0.907	0.907

Note: This table explores the effect of participation in the climate pilot exercise on the banks' lending behavior towards brown firms controlling for bank characteristics such as size, capital ratio, and ROA. *Loan Amount (Ln)* and *Spread (Ln)* are dependent variables. *High Emitter* is a dummy variable equal to 1 if the average carbon emissions of firms f before 2020 is above the median, and 0 otherwise. $Post$ is a dummy variable equal to 1 for the period after the French climate stress test (2020Q3 onwards), and 0 otherwise. Standard errors are double-clustered at the bank and the borrower level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.7.5 Disentangling Different Climate Stress Tests

We address one final concern relating to anticipation of other climate stress tests. While the climate pilot exercise by the ACPR was the first comprehensive climate stress test, other central banks like the Bank of England and the European Central Bank followed suit with climate stress tests shortly thereafter in 2021 and 2022. In particular, the latter climate stress test poses a potential confounding event, as some of the large French banks (and banks from the control group) that participated in the French climate pilot exercise are also subject to scrutiny by the Single Supervisory Mechanism from the European Central Bank. Given that the ECB climate stress test was conducted in 2022, our coefficients may therefore reflect these banks' anticipation of the ECB climate stress test rather than representing an exclusive reaction to the climate pilot exercise conducted by the ACPR. To avoid an erroneous attribution of our coefficients to the French climate exercise, we replicate in Table 4.15 our main tests, but remove all observations for the years 2022 and 2023. Our inference remains qualitatively unaffected.

Table 4.15: Disentangling the French Climate Pilot Exercise from the ECB Climate Stress Test

	Loan amount (Ln)		Spread (Ln)	
	(1)	(2)	(3)	(4)
Treat × High Emitter × Post	0.345 **	0.333 **	0.096 ***	0.094 **
	(0.170)	(0.165)	(0.035)	(0.036)
Treat × High Emitter	-0.296 *	-0.291 *	-0.055 *	-0.053 *
	(0.168)	(0.165)	(0.029)	(0.030)
High Emitter × Post	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
High Emitter	-0.311	-0.350	-0.383 **	-0.411 **
	(0.499)	(0.493)	(0.173)	(0.176)
Observations	1,277	1,277	1,277	1,277
Loan controls	Yes	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes
Loan Type FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.840	0.841	0.883	0.885

Note: This table shows the effect of the climate pilot exercise on the banks' lending behavior towards brown firms. In these tests, we remove observations from 2022 onwards and only use data for 2015 to 2021 to avoid that our coefficients of interest reflect anticipation of the climate stress test performed by the ECB in 2022. *Loan Amount (Ln)* and *Spread (Ln)* are dependent variables. $Treated$ is a dummy taking on the value 1 if a bank participates in climate stress tests, and 0 otherwise. *High Emitter* is a dummy variable equal to 1 if the average carbon emissions of borrower f before 2020 is above the median, and 0 otherwise. $Post$ is a dummy variable equal to 1 for the period after the French climate stress test (2020Q3 onwards), 0 otherwise. Standard errors are double-clustered at the bank and borrower level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.8 Conclusion

Bank supervisors are pressuring banks to protect themselves from the effects of climate change, and this pressure also affects bank borrowers. We exploit data from the climate pilot exercise conducted by the French prudential regulatory agency that serves as a plausibly exogenous shock to these banks' information-production efforts to understand the effects of climate change, and combine it with data on these banks' borrowers' carbon emissions to capture their exposure to transition risk. This enables us to investigate how supervisory activities that affect the banks' lending policies shape the transition to the carbon-neutral economy and affect the borrowers' actions to decarbonize their business models.

By comparing loan-contract features and environmental performance from borrowers whose banks participate in the French climate pilot exercise with such outcomes from borrowers whose banks do not participate, we can establish the causal effect of the climate pilot exercise on the banks' lending behavior, and, ultimately, on the borrowers' transition paths.

Our work illustrates that climate stress tests can be viewed as a learning exercise for banks. We show that the climate pilot exercise triggers reassessments of the banks' lending policies because it produces new information signals that improve the banks' comprehension of the long-run implications of climate change. Therefore, banks are better able to assess the borrowers' transition risk. In other words, supervision, in the form of climate stress tests, is valuable as an information-collection exercise that has ramifications not just for loan contracting decisions, but also for real outcomes.

Our first novel finding is that banks that take part in the climate pilot exercise increase lending to borrowers despite their higher transition risk. While it is plausible for participating banks to facilitate the transition to the carbon-neutral economy, their support of high transition risk borrowers does not come for free because they raise loan rates at the same time. This result contrasts with banks that do not participate in the climate pilot exercise. These banks reduce credit supply. The latter finding does not seem surprising. Non-participating banks are not required to spend time and effort collecting information about the borrowers' transition plans to assess the long-term effects of climate change, and therefore evaluate transition risk with a short-term perspective.

What is surprising and, importantly, also encouraging is our result that participation in climate stress tests reverses the banks' assessment of borrowing firms that are considered more exposed to climate-transition risk. These banks update their beliefs about borrowers because of the information acquired during the climate stress tests, and this is also reflected in our second novel insight that participating banks significantly increase their focus on climate change in conference calls and information-collection efforts related to the borrowers' carbon emissions. Rather than reducing credit, the participating banks' deeper comprehension of the transition process results in a greater willingness to commit funds to borrowers and support their transition to the carbon-neutral economy.

Our third set of novel findings further reinforces this view. The tests of the borrowers' environmental performance show that brown borrowers of participating banks are more likely to have emission policies and set carbon emission targets, and use more renewable energy. The probability that these borrowers have eco-friendly products is also higher. ESG scores, environmental scores, and emission scores of these borrowers also improved after receiving loans from climate stress-tested banks. These positive developments, however, need to be considered in light of other findings concerning environmental dimensions that are more difficult to adjust in the long run. We neither observe reductions in direct carbon emissions, nor do borrowers terminate contracts with suppliers that are flagged as environmentally unfriendly.

Taken together, our results illustrate the role of climate stress tests beyond their primary objective of identifying vulnerabilities in the financial system related to climate change. Climate stress tests are valuable because they reduce information asymmetries between banks and borrowers with regard to how to measure the effect of climate change, and therefore can also be justified on the grounds that they support the transition to a carbon-neutral economy. They boost the banks' understanding of transition risk to engage in 'greener' lending and facilitate the borrowers' efforts in the process of making their businesses more resilient towards climate change. To that extent, our research helps advance the understanding of the role of banking supervision in the context of climate change.

4.9 Appendix C

4.9.1 French Climate Pilot Exercise

Preceding the actual climate exercise, the preparatory phase of the pilot published in April 2019 was based on questionnaires. In total, 15 insurance and nine banking groups got involved. Institutions participated as part of a system-wide exercise where scenarios and assumptions were provided by the authorities, a classical bottom-up approach in stress-testing. The nine banking groups that we focus on cover 85 percent of French banks' total assets illustrating high added value of the sector and underlining the representative nature of results as these groups represent a very significant part of the banking activity in France. Due to the complex interactions with economic and social systems involved, there are several modifications in contrast to standard stress-testing procedures.[6]

First, the exercise adds a forward-looking view of risks over a long-term horizon conditional on the implementation of several alternative scenarios. In particular, the exercise looks at a 30-year horizon ranging from 2020-2050 containing three transition scenarios.[7] Different from the 3-5 years that are considered in traditional stress testing scenarios this period is sufficiently long to integrate the effects of climate change. However, the long time horizon requires a revision of the static balance-sheet assumption. Therefore, the pilot exercise combines two assumptions: First a "static balance sheet" assumption until 2025, following a "dynamic balance sheet" from 2025-2050 to analyse the strategies of financial institutions and the actions implemented to mitigate the effects of climate change allowing financial institutions to take new risks into consideration and assess corrective actions. Second, geographical and sectoral scopes are expanded. Due to the fact that the activities of institutions have international impact climate-related risks have to be considered differently based on the geographical areas. Additionally, aggregated asset classes are split into 55 activity sectors allowing for a more granular analysis.

The baseline transition scenario corresponding to an orderly transition is consistent with the narrative of the SNBC, France's roadmap for fulfilling commitments made under the Paris Agreement. It includes a significant increase in the price of carbon where financial institutions face different CO2 emission trajectories. To compare to the

[6]See https://acpr.banque-france.fr/sites/default/files/medias/documents/20200717_main_assumptions_and_scenarios_of_the_acpr_climate_pilot_exercise.pdf

[7]The network of central banks and supervisors for greening the financial sector serves as a guideline on the construction of climate change scenarios and serves as a basis for two of the scenarios published by the NGFS in June 2020. The third one is a physical risk scenario.

baseline, there are two disorderly transition scenarios. The first one is referred to as "late transition". It relies on the assumption that the target for reducing greenhouse gas emissions is not met by 2030 assuming that carbon sequestration technologies are less efficient than expected.

This scenario replicates the aggregate level of emission, carbon price and GDP trajectories of the representative scenario for a "disorderly" transition. It is based on a very high increase in the carbon price in 2030 to maintain carbon neutrality target in 2050 (in particular it rises from 14$ to 704$ per ton of CO2). The second scenario is called the "sudden transition" scenario and combines a sharp increase in the price of carbon that reaches 917$ per ton of CO2 in 2050 and a less favourable evolution of productivity than in the baseline scenario from 2025 onwards. Moreover, renewable-energy technologies are less efficient than expected, implying even higher energy prices and additional investment. It is important to note that contrary to usual stress-testing exercises the scenarios on CO2 emission trajectories do not trigger an economic downturn by 2050 but slower economic growth combining different assumptions in terms of carbon tax trajectories and total productivity levels.

The scenarios on CO2 emission trajectories are based on a set of assumptions modelling the interactions between socioeconomic systems and the climate. The three scenarios combine assumptions in terms of trajectory on carbon tax and total productivity levels. The main objective is to measure the consequences of these scenarios that materialise via transition risk on bank balance sheets.

Among the variety of risk categories, they chose to focus on two important financial risks: credit and market risk. For credit risk projections, the banking groups were asked to measure the impact of the various transition scenarios on expected credit losses. They approximate the annual cost of credit risk.[8] In general, institutions were requested to perform credit risk projections on three portfolios: (i) a corporate portfolio including SMEs; (ii) a retail portfolio; (iii) and a sovereign portfolio using benchmark probabilities of default provided by the ACPR.

Market risk focuses on analysing the impact of financial shocks caused by the implementation of energy transition policies. Specifically, institutions looked at (i) the fair value revaluation of the trading book following an instantaneous market shock induced by the valuation of assets under adverse transition scenarios; and (ii) the impact of market shocks on the counterparty risk in the most sensitive sectors.

[8]Expressed in basis points and calculated by dividing the total annualised provisioning flows for each time interval by the average exposure over the same time interval.

Counterparty risk was measured by using the impact of default of the two largest counterparties of the institution. This is especially useful for identifying substantial market positions on carbon intensive counterparties.

4.9.2 Climate Pilot Exercise Participants

Table C.1: French Climate Pilot Exercise Participants

Number	Bank name
1	AGENCE FRANÇAISE DE DÉVELOPPEMENT
2	BNP PARIBAS
3	BPCE
4	CAISSE DES DÉPÔTS
5	CREDIT AGRICOLE
6	CREDIT MUTUEL
7	LA BANQUE POSTALE
8	SOCIÉTÉ GÉNÉRALE
9	SOCIÉTÉ DE FINANCEMENT LOCALE

Note: This table shows an overview about the nine banking groups that participated in the French climate pilot exercise in 2020.

4.9.3 Variable Descriptions

Table C.2: Variable Definitions

Variable	Description	Source
Loan Amount (Ln)	Natural log of loan amount	Dealscan
Loan Spreads	Spread in basis points over Libor	Dealscan
Loan Maturity (Years)	Loan maturity in years	Dealscan
Green Loan	Dummy that equals 1 if a loan is a sustainability-linked loan or for green purposes, 0 otherwise	Dealscan
Treated	Dummy that equals 1 if a bank participated in the French climate pilot exercise, 0 otherwise	Authors' Collection
Post	Dummy that equals 1 if after 2020Q3, 0 otherwise	Authors' Collection
High Emitter	Dummy that is equal to one if firms' emissions between 2015 and 2019 is above mean and zero otherwise	Refinitiv
Borrower Size	Natural log of borrowers' total assets	Compustat
Borrower Leverage	Ratio of borrowers' total debts over total assets	Compustat
Borrower ROA	Borrowers' returns on total assets	Compustat
SA Index	Size-Age Index	Compustat
Lender Size	Natural log of borrowers' total assets	Compustat
Lender Capital	Ratio of lenders' equity capital over total assets	Compustat
Lender Deposits	Ratio of lenders' deposits over total assets	Compustat
Lender ROA	Lenders' returns on total assets	Compustat
Green Lender	Dummy that equals 1 if a bank joined the United Nations Environment Programme (UNEP) before 2020, 0 otherwise	Authors' Collection
Institutional Ownership	Percentage of institutional ownership	Bloomberg
Eco-Friendly Product	Dummy that equals 1 if a firm produces eco-friendly products, 0 otherwise	Refinitiv
ESG Scores	ESG scores	Refinitiv
Environmental Scores	Environmental scores	Refinitiv
Emission Policies	Dummy that equals 1 if a firm has emission policies, 0 otherwise	Refinitiv

Continues next page

Table C.2: Variable Definitions (Continued)

Variable	Description	Source
Target Emissions	Dummy that equals 1 if a firm has target emissions, 0 otherwise	Refinitiv
Renewable Energy	Dummy that equals 1 if a firm invests in renewable energy technologies, 0 otherwise	Refinitiv
Emissions Score	Emission scores	Refinitiv
Total Emissions Growth	Growth in total emissions	Refinitiv
Direct Emissions Growth	Growth in scope 1 emissions	Refinitiv
Termination of Environmentally Unfriendly Suppliers	Dummy that equals 1 if a firm terminates contracts with suppliers who are environmentally unfriendly, 0 otherwise	Refinitiv
Materials Sourcing Criteria	Dummy that equals 1 if a firm claims to use environmental criteria to source material, 0 otherwise	Refinitiv
Mentioning Climate Stress Test	Dummy that equal 1 if a bank mentions climate stress tests during their conference calls	S&P Capital IQ
Communication with Borrowers on Transition Risk	Dummy that equal 1 if a bank mentions that the bank communicates with borrowers on transition risk during their conference calls	S&P Capital IQ
Discussion about Emission Data	The number of times that a bank discusses information about emissions	S&P Capital IQ

4.9.4 Heckman Selection Model

Table C.3: Heckman Selection Model

	Second Stage Result		First Stage Result	
	(1)	(2)		(3)
Dependant Variable	Loan Amount (Ln)	Spreads (Ln)		P(Treat)
Treat × Post × High Emitter	0.350*	0.079**	Green Lender	0.503*
	(0.175)	(0.038)		(0.293)
Treat × High Emitter	-0.339*	-0.038	Lender Size	-0.247
	(0.184)	(0.038)		(0.377)
Treat × Post	0.185	-0.074	Lender Capital	-0.119
	(0.165)	(0.052)		(0.182)
High Emitter × Post	0.000	0.000	Lender Deposit	-0.031
	(0.000)	(0.000)		(0.040)
High Emitter	-0.394	-0.298**		
	(0.504)	(0.132)		
Post	0.000	0.000		
	(0.000)	(0.000)		
Inverse Mills Ratio	-0.073	0.015		
	(0.061)	(0.017)		
Observations	1,696	1,696		698
Loan Controls	Yes	Yes		No
Firm Controls	Yes	Yes		No
Bank FE	Yes	Yes		Yes
Industry Year FE	Yes	Yes		No
Year FE	No	No		Yes
Loan Type FE	Yes	Yes		No
Country FE	No	No		Yes
Adjusted R^2	0.837	0.884		-
Pseudo R^2	-	-		0.849
Clustering	Bank, Firm	Bank, Firm		Bank

Note: This table explores the effect of climate stress tests on the banks' lending and pricing behavior towards brown firms, controlling for possible selection bias. *Loan Amount (Ln)* and *Spread (Ln)* are dependent variables. *Green lender* is a dummy variable that takes on the value of one if a bank signed the UN Principles for Responsible Banking (0 otherwise) prior to the climate pilot exercise before 2020. *High Emitter* is a dummy variable equal to 1 if the average carbon emissions of firms f before 2020 is above the median, and 0 otherwise. $Post$ is a dummy variable equal to 1 for the period after the French climate stress test (2020Q3 onwards), 0 otherwise. Standard errors are clustered at the bank and firm level and reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.9.5 Sample Transcripts from Conference Calls

"So at this stage, for us, it's a way to work on stress test on climate risk. It doesn't lead into additional capital yet. But of course, it prepares us to handle all the upcoming stress tests, which are going to be led by regulators and supervisors. It's part also of our disclosures, and it will prepare us also in terms of segmentation of clients that are information and reporting when we will have to align also with the various regulations."

Société Générale Société Anonyme; Q3 2020 Earnings Call; Nov 05, 2020

"And then on your other question of the Net-Zero Banking, it is what you see, and it will even be further strengthened in all the climate stress tests that are ongoing. So there is – we at BNP Paribas, we really have embarked already since years. The fact that one has to evolve. And of course, you have to be sure that you yourself as a bank are carbon neutral, but you have to accompany the clients for this to happen."

BNP Paribas SA; Q3 2021 Earnings Call; Oct 29, 2021

"We therefore maintain an increasingly close dialogue with our clients on CSR issues with the aim of analyzing and understanding their specific needs, assisting them in their own projects with positive impacts, selecting or structuring appropriate offers, all of this is in compliance with the group's own commitments."

Société Générale Société Anonyme; Shareholder/Analyst Call; May 17, 2022

"Our CO2 emission reduction targets are fed by the constant dialogue we're having with our clients. Meeting the target depends on our working hand-in-hand with those clients on the execution of their transition strategies, which in turn depends on the competence and effectiveness of our teams."

BNP Paribas SA; Q1 2023 Earnings Call; May 03, 2023

Bibliography

Accetturo, Antonio, Giorgia Barboni, Michele Cascarano, Emilia Garcia-Appendini, and Marco Tomasi. 2022. "Credit supply and green investments." Working paper. Available at SSRN 4217890.

Acharya, Viral V, Allen N Berger, and Raluca A Roman. 2018. "Lending implications of US bank stress tests: Costs or benefits?" *Journal of Financial Intermediation*, 34: 58–90.

Acharya, Viral V, Richard Berner, Robert F Engle III, Hyeyoon Jung, Johannes Stroebel, Xuran Zeng, and Yihao Zhao. 2023. "Climate stress testing." *National Bureau of Economic Research*.

Agarwal, Sumit, David Lucca, Amit Seru, and Francesco Trebbi. 2014. "Inconsistent regulators: Evidence from banking." *The Quarterly Journal of Economics*, 129(2): 889–938.

Ahn, Tom, Peter Arcidiacono, Amy Hopson, and James Thomas. 2024. "Equilibrium grading policies with implications for female interest in STEM courses." *Econometrica*, 92(3): 849–880.

Alt, Christian, Daniela Gesell, Sandra Hubert, Katrin Hüsken, Ralf Kuhnke, and Kerstin Lippert. 2017. "DJI-Kinderbetreuungsreport 2017." *Inanspruchnahme und Bedarfe aus Elternperspektive im Bundesländervergleich. München*.

Anger, Christina, and Axel Plünnecke. 2022. "MINT gewinnt. Hohe Löhne in den MINT-Berufen." IW-Kurzbericht 106, Cologne.

Anginer, Deniz, Karel Hrazdil, Jiyuan LI, and Ray Zhang. 2023. "Climate reputation and bank loan contracting." Working paper. Available at SSRN 3723771.

Ashraf, Nava, Oriana Bandiera, Edward Davenport, and Scott S Lee. 2020. "Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services." *American Economic Review*, 110(5): 1355–1394.

Aswani, Jitendra, Aneesh Raghunandan, and Shiva Rajgopal. 2024. "Are carbon emissions associated with stock returns?" *Review of Finance*, 28(1): 75–106.

Attanasio, Orazio P, and Katja M Kaufmann. 2014. "Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender." *Journal of Development Economics*, 109: 203–216.

Bapna, Ravi, Nishtha Langer, Amit Mehra, Ram Gopal, and Alok Gupta. 2013. "Human capital investments and employee performance: An analysis of IT services industry." *Management Science*, 59(3): 641–658.

Bassett, William F, Seung Jung Lee, and Thomas Popeck Spiller. 2015. "Estimating changes in supervisory standards and their economic effects." *Journal of Banking & Finance*, 60: 21–43.

Baudino, Patrizia, and Jean-Philippe Svoronos. 2021. "FSI Insights."

Belenzon, Sharon, and Ulya Tsolmon. 2016. "Market frictions and the competitive advantage of internal labor markets." *Strategic Management Journal*, 37(7): 1280–1303.

Belot, Michele, Philipp Kircher, and Paul Muller. 2022. "How wage announcements affect job search – A field experiment." *American Economic Journal: Macroeconomics*, 14(4): 1–67.

Berger, A.N., S. El Ghoul, O. Guedhami, and R.A. Roman. 2020. "Deregulation and banks' cost of equity capital." *Mimeo*.

Bernanke, Ben. 2006. "Bank regulation and supervision: Balancing benefits and costs." Board of Governors of the Federal Reserve System. No. 238, Speech.

Bettinger, Eric P, and Bridget Terry Long. 2005. "Do faculty serve as role models? The impact of instructor gender on female students." *American Economic Review*, 95(2): 152–157.

Bianchi, Nicola, and Michela Giorcelli. 2020. "Scientific education and innovation: From technical diplomas to university STEM degrees." *Journal of the European Economic Association*, 18(5): 2608–2646.

Blau, Francine D, and Lawrence M Kahn. 2017. "The gender wage gap: Extent, trends, and explanations." *Journal of Economic Literature*, 55(3): 789–865.

Bobba, Matteo, and Verónica Frisancho. 2016. "Learning about oneself: The effects of performance feedback on school choice." Unpublished manuscript, Inter-American Development Bank.

Bolton, Patrick, and Marcin Kacperczyk. 2023. "Global pricing of carbon-transition risk." *The Journal of Finance*, 78(6): 3677–3754.

Bond, Timothy N, George Bulman, Xiaoxiao Li, and Jonathan Smith. 2018. "Updating human capital decisions: Evidence from SAT score shocks and college applications." *Journal of Labor Economics*, 36(3): 807–839.

Borio, Claudio, Stijn Claessens, and Nikola A. Tarashev. 2023. "Finance and climate change risk: Managing expectations." *EconPol Forum*, 24(1): 5–7.

Breda, Thomas, and Clotilde Napp. 2019. "Girls' comparative advantage in reading can largely explain the gender gap in math-related fields." *Proceedings of the National Academy of Sciences*, 116(31): 15435–15440.

Breda, Thomas, Elyès Jouini, and Clotilde Napp. 2018. "Societal inequalities amplify gender gaps in math." *Science*, 359(6381): 1219–1220.

Breda, Thomas, Elyès Jouini, and Clotilde Napp. 2023. "Gender differences in the intention to study math increase with math performance." *Nature Communications*, 14(1): 3664.

Breda, Thomas, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre. 2019. "Female role models: Are they effective at encouraging girls to study science?" Notes IPP, (45).

Breda, Thomas, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre. 2023. "How Effective are Female Role Models in Steering Girls Towards STEM? Evidence from French High Schools." *The Economic Journal*, 133(653): 1773–1809.

Brown, Charles, and James Medoff. 1989. "The employer size-wage effect." *Journal of Political Economy*, 97(5): 1027–1059.

Bruno, Brunella, and Sara Lombini. 2023. "Climate transition risk and bank lending." *Journal of Financial Research*, 46: S59–S106.

Burkacki, Ondrej, Nikolaus Lehmann, and Julia Dragon. 2022. "The semiconductor decade: A trillion-dollar industry." McKinsey and Company. Accessed on June 09, 2024: `https://www.mckinsey.de/industries/semiconductors/our-insights/the-semiconductor-decade-a-trillion-dollar-industry#/`.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *The Review of Economics and Statistics*, 90(3): 414–427.

Canaan, Serena, and Pierre Mouganie. 2023. "The impact of advisor gender on female students' STEM enrollment and persistence." *Journal of Human Resources*, 58(2): 593–632.

Card, David, and A Abigail Payne. 2021. "High school choices and the gender gap in STEM." *Economic Inquiry*, 59(1): 9–28.

Card, David, Fabrizio Colella, and Rafael Lalive. 2024. "Gender preferences in job vacancies and workplace gender diversity." *Review of Economic Studies*. Forthcoming.

Carlana, Michela. 2019. "Implicit stereotypes: Evidence from teachers' gender bias." *The Quarterly Journal of Economics*, 134(3): 1163–1224.

Carnevale, Anthony P, Nicole Smith, and Michelle Melton. 2011. "STEM: Science Technology Engineering Mathematics." Georgetown University Center on Education and the Workforce. Accessed on July 25, 2024: `https://files.eric.ed.gov/fulltext/ED525297.pdf`.

Ceccarelli, Marco, Stefano Ramelli, and Alexander F Wagner. 2024. "Low carbon mutual funds." *Review of Finance*, 28(1): 45–74.

Ceci, Stephen J, Donna K Ginther, Shulamit Kahn, and Wendy M Williams. 2014. "Women in academic science: A changing landscape." *Psychological Science in the Public Interest*, 15(3): 75–141.

Chaly, Stephanie, James Hennessy, Lev Menand, Kevin Stiroh, and Joseph Tracy. 2017. "Misconduct risk, culture, and supervision." *Federal Reserve Bank of New York*.

Chava, Sudheer. 2014. "Environmental externalities and cost of capital." *Management Science*, 60(9): 2223–2247.

Coffman, Katherine B, Manuela R Collis, and Leena Kulkarni. 2024. "Whether to apply." *Management Science*, 70(7): 4649–4669.

Coff, Russell W. 1997. "Human assets and management dilemmas: Coping with hazards on the road to resource-based theory." *Academy of Management Review*, 22(2): 374–402.

Correa, Ricardo, Ai He, Christoph Herpfer, and Ugur Lel. 2022. "The rising tide lifts some interest rates: climate change, natural disasters, and loan pricing." International Finance Discussion Paper (1345).

Correia, Sergio, Paulo Guimarães, and Tom Zylkin. 2020. "Fast Poisson estimation with high-dimensional fixed effects." *The Stata Journal*, 20(1): 95–115.

Cortés, Kristle R, Yuliya Demyanyk, Lei Li, Elena Loutskina, and Philip E Strahan. 2020. "Stress tests and small business lending." *Journal of Financial Economics*, 136(1): 260–279.

Dal Bó, Ernesto, Frederico Finan, and Martín A Rossi. 2013. "Strengthening state capabilities: The role of financial incentives in the call to public service." *The Quarterly Journal of Economics*, 128(3): 1169–1218.

Dang, Tri Vi, Gary Gorton, and Bengt Holmstrom. 2009. "Opacity and the optimality of debt for liquidity provision." *Manuscript Yale University*.

Danisewicz, Piotr, Danny McGowan, Enrico Onali, and Klaus Schaeck. 2018. "Debt priority structure, market discipline and bank conduct." *Review of Financial Studies*, 31(11): 4493–4555.

Dee, Thomas S. 2005. "A teacher like me: Does race, ethnicity, or gender matter?" *American Economic Review*, 95(2): 158–165.

Dee, Thomas S. 2007. "Teachers and the gender gaps in student achievement." *Journal of Human resources*, 42(3): 528–554.

Delaney, Judith M, and Paul J Devereux. 2019. "Understanding gender differences in STEM: Evidence from college applications." *Economics of Education Review*, 72: 219–238.

Del Carpio, Lucia, and Maria Guadalupe. 2022. "More women in tech? Evidence from a field experiment addressing social identity." *Management Science*, 68(5): 3196–3218.

Del Carpio, Lucia, and Thomas Fujiwara. 2023. "Do gender-neutral job ads promote diversity? Experimental evidence from Latin America's tech sector." Working paper, National Bureau of Economic Research.

Delfino, Alexia. 2024. "Breaking gender barriers: Experimental evidence on men in pink-collar jobs." *American Economic Review*, 114(6): 1816–1853.

Delis, Manthos D, and Panagiotis K Staikouras. 2011. "Supervisory effectiveness and bank risk." *Review of Finance*, 15(3): 511–543.

Delis, Manthos D, Kathrin de Greiff, Maria Iosifidi, and Steven Ongena. 2024. "Being stranded with fossil fuel reserves? Climate policy risk and the pricing of bank loans." *Financial Markets, Institutions & Instruments*, 33(3): 239–265.

Deserranno, Erika. 2019. "Financial incentives as signals: Experimental evidence from the recruitment of village promoters in Uganda." *American Economic Journal: Applied Economics*, 11(1): 277–317.

DeVaro, Jed, Oliver Gürtler, Marc Gürtler, and Christian Deutscher. 2024. "Big fish in small (and big) ponds: A study of careers." *The Journal of Law, Economics, and Organization*, 40(1): 76–107.

Diamond, Douglas W. 1984. "Financial intermediation and delegated monitoring." *The Review of Economic Studies*, 51(3): 393–414.

Duan, Tinghua, Frank Weikai Li, and Quan Wen. 2023. "Is carbon risk priced in the cross-section of corporate bond returns?" *Journal of Quantiative and Finanical Analysis*.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya." *American Economic Review*, 101(5): 1739–74.

Ehlers, Torsten, Benoit Mojon, and Frank Packer. 2020. "Green bonds and carbon emissions: exploring the case for a rating system at the firm level." *BIS Quarterly Review, September*.

Eisenbach, Thomas M, David O Lucca, and Robert M Townsend. 2016. "The economics of bank supervision." National Bureau of Economic Research No. w22201.

Elsner, Benjamin, and Ingo E Isphording. 2017. "A big fish in a small pond: Ability rank and human capital investment." *Journal of Labor Economics*, 35(3): 787–828.

Elsner, Benjamin, Ingo E Isphording, and Ulf Zölitz. 2021. "Achievement rank affects performance and major choices in college." *The Economic Journal*, 131(640): 3182–3206.

European Central Bank. 2022. "2022 Climate Risk Stress Test." Accessed on November 22, 2024: `https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.climate_stress_test_report.20220708~2e3cc0999f.en.pdf`.

Flannery, Mark, Beverly Hirtle, and Anna Kovner. 2017. "Evaluating the information in the federal reserve stress tests." *Journal of Financial Intermediation*, 29: 1–18.

Flory, Jeffrey A, Andreas Leibbrandt, and John A List. 2015. "Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions." *The Review of Economic Studies*, 82(1): 122–155.

Flory, Jeffrey A, Andreas Leibbrandt, Christina Rott, and Olga Stoddard. 2021. "Increasing workplace diversity: Evidence from a recruiting experiment at a Fortune 500 company." *Journal of Human Resources*, 56(1): 73–92.

Fox, Jeremy T. 2009. "Firm-size wage gaps, job responsibility, and hierarchical matching." *Journal of Labor Economics*, 27(1): 83–126.

Francesconi, Marco, and Matthias Parey. 2018. "Early gender gaps among university graduates." *European Economic Review*, 109: 63–82.

Fuchs, Larissa, Huyen Ngyuen, Trang Nguyen, and Klaus Schaeck. 2024*a*. "Climate stress tests, bank lending, and the transition to the carbon-neutral economy." IWH Discussion Papers.

Fuchs, Larissa, Matthias Heinz, Pia Pinger, and Max Thon. 2024*b*. "How to attract talents? Field-experimental evidence on emphasizing flexibility and career opportunities in job advertisements." ECONtribute Discussion Paper.

Giannetti, Mariassunta, and Luc Laeven. 2012. "Flight home, flight abroad, and international credit cycles." *American Economic Review*, 102(3): 219–224.

Gill, Andrej, Matthias Heinz, Heiner Schumacher, and Matthias Sutter. 2023. "Social preferences of young professionals and the financial industry." *Management Science*, 69(7): 3905–3919.

Glitz, Albrecht. 2017. "Coworker networks in the labour market." *Labour Economics*, 44: 218–230.

Goldin, Claudia. 2014. "A grand gender convergence: Its last chapter." *American Economic Review*, 104(4): 1091–1119.

Goldsmith-Pinkham, Paul S, Beverly Hirtle, and David O Lucca. 2016. "Parsing the content of bank supervision." FRB of NY Staff Report.

Goldstein, Itay, and Haresh Sapra. 2014. "Should banks' stress test results be disclosed? An analysis of the costs and benefits." *Foundations and Trends® in Finance*, 8(1): 1–54.

Gorton, Gary, and Guillermo Ordonez. 2014. "Collateral crises." *American Economic Review*, 104(2): 343–378.

Goulas, Sofoklis, Silvia Griselda, and Rigissa Megalokonomou. 2022. "Comparative advantage and gender gap in STEM." *Journal of Human Resources*, 0320–10781R2.

Gropp, Reint, Thomas Mosk, Steven Ongena, and Carlo Wix. 2019. "Banks response to higher capital requirements: Evidence from a quasi-natural experiment." *The Review of Financial Studies*, 32(1): 266–299.

Groshen, Erica L. 1991. "Sources of intra-industry wage dispersion: How much do employers matter?" *The Quarterly Journal of Economics*, 106(3): 869–884.

Guiso, Luigi, Ferdinando Monte, Paola Sapienza, and Luigi Zingales. 2008. "Culture, gender, and math." *Science*, 320(5880): 1164–1165.

Gustafson, Matthew T, Ivan T Ivanov, and Ralf R Meisenzahl. 2021. "Bank monitoring: Evidence from syndicated loans." *Journal of Financial Economics*, 139(2): 452–477.

Hadlock, Charles J., and Joshua R. Pierce. 2010. "New evidence on measuring financial constraints: Moving beyond the KZ index." *The Review of Financial Studies*, 23(5): 1909–1940.

Heckman, James J. 1979. "Sample selection bias as a specification error." *Econometrica: Journal of the Econometric Society*, 153–161.

Heider, Florian, and Roman Inderst. 2022. "A corporate finance perspective on environmental policy." CEPR Discussion Paper No. DP16878.

Heinz, Matthias, and Heiner Schumacher. 2017. "Signaling cooperation." *European Economic Review*, 98: 199–216.

Hemkes, Barbara, Kim Maureen Wiesner, Lex Borghans, Philipp Seegers, Jan Bergerhoff, and Stephan Hartmann. 2016. "Studierendenbefragung zur Attraktivität der beruflichen Bildung." *Abschlussbericht des Entwicklungsprojekts*, 3: 306.

Hirtle, Beverly, Anna Kovner, and Matthew Plosser. 2020. "The impact of supervision on bank performance." *The Journal of Finance*, 75(5): 2765–2808.

Hoffman, Mitchell, Lisa B Kahn, and Danielle Li. 2018. "Discretion in hiring." *The Quarterly Journal of Economics*, 133(2): 765–800.

Hong, Fuhai, Tanjim Hossain, and John A List. 2015. "Framing manipulations in contests: A natural field experiment." *Journal of Economic Behavior & Organization*, 118: 372–382.

Idson, Todd L, and Walter Y Oi. 1999. "Workers are more productive in large firms." *American Economic Review*, 89(2): 104–108.

Imbens, Guido W, and Jeffrey M Wooldridge. 2009. "Recent developments in the econometrics of program evaluation." *Journal of Economic Literature*, 47(1): 5–86.

Ivanov, Ivan T, Mathias S Kruttli, and Sumudu W Watugala. 2024. "Banking on carbon: Corporate lending and cap-and-trade policy." *The Review of Financial Studies*, 37(5): 1640–1684.

Ivashina, Victoria. 2009. "Asymmetric information effects on loan spreads." *Journal of Financial Economics*, 92(2): 300–319.

James, Christopher. 1987. "Some evidence on the uniqueness of bank loans." *Journal of Financial Economics*, 19(2): 217–235.

Jann, Ben. 2008. "The Blinder-Oaxaca decomposition for linear regression models." *The Stata Journal*, 8(4): 453–479.

Jensen, Robert. 2010. "The (perceived) returns to education and the demand for schooling." *The Quarterly Journal of Economics*, 125(2): 515–548.

Jobvite. 2019*a*. "Share of Job Applications Worldwide in 2018, by Source." Accessed October 27, 2023: `https://www.statista.com/statistics/881116/recruitment-share-of-job-applications-by-source-worldwide/`.

Jobvite. 2019*b*. "Share of Job Hires Worldwide in 2018, by Source." *Chart*, Accessed October 27, 2023: `https://www.statista.com/statistics/881139/recruitment-share-of-job-hires-by-source-worldwide/`.

Kacperczyk, Marcin T, and José-Luis Peydró. 2022. "Carbon emissions and the bank-lending channel." Working paper. Available at SSRN 3915486.

Kahn, Lisa B. 2010. "The long-term labor market consequences of graduating from college in a bad economy." *Labour Economics*, 17(2): 303–316.

Kasy, Maximilian, and Anja Sautmann. 2021. "Adaptive treatment assignment in experiments for policy choice." *Econometrica*, 89(1): 113–132.

Kiser, Elizabeth K, Robin A Prager, and Jason Scott. 2012. "Supervisor ratings and the contraction of bank lending to small businesses." FEDS Working Paper.

Kleven, Henrik, and Camille Landais. 2017. "Gender inequality and economic development: Fertility, education and norms." *Economica*, 84(334): 180–209.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. "Children and gender inequality: Evidence from Denmark." *American Economic Journal: Applied Economics*, 11(4): 181–209.

Kok, Christoffer, Carola Müller, Steven Ongena, and Cosimo Pancaro. 2023. "The disciplining effect of supervisory scrutiny in the EU-wide stress test." *Journal of Financial Intermediation*, 53: 101015.

Krueger, Philipp, Zacharias Sautner, and Laura T Starks. 2020. "The importance of climate risks for institutional investors." *The Review of Financial Studies*, 33(3): 1067–1111.

Kugler, Adriana D, Catherine H Tinsley, and Olga Ukhaneva. 2021. "Choice of majors: Are women really different from men?" *Economics of Education Review*, 81: 102079.

Kuhn, Peter, and Kailing Shen. 2023. "What happens when employers can no longer discriminate in job ads?" *American Economic Review*, 113(4): 1013–1048.

Lievens, Filip, and Jerel E Slaughter. 2016. "Employer image and employer branding: What we know and what we need to know." *Annual Review of Organizational Psychology and Organizational behavior*, 3(1): 407–440.

Li, Hongyan, and Xing Xia. 2024. "Grades as signals of comparative advantage: How letter grades affect major choices." *Journal of Economic Behavior & Organization*, 227: 106717.

Lummer, Scott L, and John J McConnell. 1989. "Further evidence on the bank lending process and the capital-market response to bank loan agreements." *Journal of Financial Economics*, 25(1): 99–122.

ManpowerGroup. 2024. "2024 Global talent shortage survey." Accessed online on June 9, 2024: `https://go.manpowergroup.com/hubfs/Talent%20Shortage/Talent%20Shortage%202024/MPG_TS_2024_GLOBAL_Infographic.pdf`.

Marianne, Bertrand. 2011. "New perspectives on gender." In *Handbook of Labor Economics*. Vol. 4, 1543–1590. Elsevier.

Marjenko, Artem, Martin Müller, and Stefan Sauer. 2021. "Das KFW-IFO-Fachkräftebarometer: Jedes fünfte deutsche Unternehmen wird derzeit durch Fachkräftemangel beeinträchtigt." *ifo Schnelldienst*, 74(04): 57–59.

Mas, Alexandre, and Amanda Pallais. 2017. "Valuing alternative work arrangements." *American Economic Review*, 107(12): 3722–3759.

Mas, Alexandre, and Amanda Pallais. 2020. "Alternative work arrangements." *Annual Review of Economics*, 12: 631–658.

Meisenzahl, Ralf. 2023. "How climate change shapes bank lending: Evidence from portfolio reallocation." FRB of Chicago Working Paper.

Moretti, Enrico. 2011. "Local labor markets." *Handbook of Labor Economics*, 4: 1237–1313.

Morgan, Donald P, Stavros Peristiani, and Vanessa Savino. 2014. "The information value of the stress test." *Journal of Money, Credit and Banking*, 46(7): 1479–1500.

Mueller, Isabella, and Eleonora Sfrappini. 2022. "Climate change-related regulatory risks and bank lending." ECB Working Paper.

Mueller, Isabella, Huyen Nguyen, and Trang Nguyen. 2022. "Carbon transition risk and corporate loan securitization." Working paper. Available at SSRN 4276781.

Murfin, Justin, and Matthew Spiegel. 2020. "Is the risk of sea level rise capitalized in residential real estate?" *The Review of Financial Studies*, 33(3): 1217–1255.

Murphy, Richard, and Felix Weinhardt. 2020. "Top of the class: The importance of ordinal rank." *The Review of Economic Studies*, 87(6): 2777–2826.

Nekoei, Arash. 2023. "Will Markets Provide Humane Jobs?" Working Paper. Available at SSRN 4523278.

Neumark, David. 1988. "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination." *Journal of Human Resources*, 23: 279–295.

Nguyen, Duc Duy, Steven Ongena, Shusen Qi, and Vathunyoo Sila. 2022. "Climate change risk and the cost of mortgage credit." *Review of Finance*, 26(6): 1509–1549.

Nicoletti, Cheti, Almudena Sevilla, and Valentina Tonei. 2022. "Gender stereotypes in the family." IZA Discussion Paper.

Niederle, Muriel, and Lise Vesterlund. 2010. "Explaining the gender gap in math test scores: The role of competition." *Journal of Economic Perspectives*, 24(2): 129–144.

Nollenberger, Natalia, Núria Rodríguez-Planas, and Almudena Sevilla. 2016. "The math gender gap: The role of culture." *American Economic Review*, 106(5): 257–261.

Nosek, Brian A, Frederick L Smyth, Natarajan Sriram, Nicole M Lindner, Thierry Devos, Alfonso Ayala, Yoav Bar-Anan, Robin Bergh, Huajian Cai, Karen Gonsalkorale, et al. 2009. "National differences in gender-science stereotypes predict national sex differences in science and math achievement." *Proceedings of the National Academy of Sciences*, 106(26): 10593–10597.

Oaxaca, Ronald. 1973. "Male-female wage differentials in urban labor markets." *International Economic Review*, 693–709.

Oaxaca, Ronald L, and Michael R Ransom. 1994. "On discrimination and the decomposition of wage differentials." *Journal of Econometrics*, 61(1): 5–21.

OECD. 2017. *OECD Handbook for Internationally Comparative Education Statistics.*

OECD. 2019*a*. *PISA 2018 Results (Volume I).*

OECD. 2019*b*. *PISA 2018 Results (Volume II).*

OECD. 2024. "The OECD Going Digital Toolkit." Accessed online on July 15, 2024: `https://goingdigital.oecd.org/en/indicator/43`.

Oehmke, Martin, and Marcus M Opp. 2022. "Green capital requirements." Swedish House of Finance Research Paper (22-16).

OpenAI. 2024. "ChatGPT-4o." `https://chat.openai.com`, Generated using ChatGPT, an AI language model by OpenAI.

Opitz, Saskia, Dirk Sliwka, Timo Vogelsang, and Tom Zimmermann. 2024. "The Algorithmic Assignment of Incentive Schemes." *Management Science*, 0(0).

Ouazad, Amine, and Matthew E Kahn. 2022. "Mortgage finance and climate change: Securitization dynamics in the aftermath of natural disasters." *The Review of Financial Studies*, 35(8): 3617–3665.

Owen, Ann L. 2010. "Grades, gender, and encouragement: A regression discontinuity analysis." *The Journal of Economic Education*, 41(3): 217–234.

Peek, Joe, and Eric Rosengren. 1995. "Bank regulation and the credit crunch." *Journal of Banking & Finance*, 19(3-4): 679–692.

Pissarides, Christopher A. 2011. "Equilibrium in the labor market with search frictions." *American Economic Review*, 101(4): 1092–1105.

Porter, Catherine, and Danila Serra. 2020. "Gender differences in the choice of major: The importance of female role models." *American Economic Journal: Applied Economics*, 12(3): 226–254.

Rask, Kevin, and Jill Tiefenthaler. 2008. "The role of grade sensitivity in explaining the gender imbalance in undergraduate economics." *Economics of Education Review*, 27(6): 676–687.

Raven, John C, and John Hugh Court. 1998. *Raven's progressive matrices and vocabulary scales.* Oxford Psychologists Press Oxford.

Rezende, Marcelo, and Jason Wu. 2014. "The effects of supervision on bank performance: Evidence from discontinuous examination frequencies." Midwest Finance Association 2013 Annual Meeting Paper.

Riegle-Crumb, Catherine, Barbara King, Eric Grodsky, and Chandra Muller. 2012. "The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into STEM college majors over time." *American Educational Research Journal*, 49(6): 1048–1073.

Ronen, Simcha. 1994. *An underlying structure of motivational need taxonomies: A cross-cultural confirmation.* Consulting Psychologists Press.

Sautner, Zacharias, Laurence Van Lent, Grigory Vilkov, and Ruishen Zhang. 2023. "Firm-level climate change exposure." *The Journal of Finance*, 78(3): 1449–1498.

Schwert, Michael. 2018. "Bank capital and lending relationships." *The Journal of Finance*, 73(2): 787–830.

Shapiro, Joel, and Jing Zeng. 2024. "Stress testing and bank lending." *The Review of Financial Studies*, 37(4): 1265–1314.

SOEP. 2022. "Own calculation." German Institute for Economic Research Unpublished raw data from Socio-Economic Panel (SOEP) v37. Accessed online on Feb 20, 2024: `https://paneldata.org/soep-core/datasets/biojob/occmove`.

Stinebrickner, Ralph, and Todd R Stinebrickner. 2014. "A major in science? Initial beliefs and final outcomes for college major and dropout." *Review of Economic Studies*, 81(1): 426–472.

Stinebrickner, Todd, and Ralph Stinebrickner. 2012. "Learning about academic ability and the college dropout decision." *Journal of Labor Economics*, 30(4): 707–748.

Tan, Brandon Joel. 2023. "The consequences of letter grades for labor market outcomes and student behavior." *Journal of Labor Economics*, 41(3): 565–588.

Tarullo, Daniel K. 2019. "Financial regulation: Still unsettled a decade after the crisis." *Journal of Economic Perspectives*, 33(1): 61–80.

Terrier, Camille. 2020. "Boys lag behind: How teachers' gender biases affect student achievement." *Economics of Education Review*, 77: 101981.

Vattuone, Giulia. 2024. "Worker Sorting and the Gender Wage Gap." Working paper.

Winters, Marcus A, Robert C Haight, Thomas T Swaim, and Katarzyna A Pickering. 2013. "The effect of same-gender teacher assignment on student achievement in the elementary and secondary grades: Evidence from panel data." *Economics of Education Review*, 34: 69–75.

Wiswall, Matthew, and Basit Zafar. 2015. "Determinants of college major choice: Identification using an information experiment." *The Review of Economic Studies*, 82(2): 791–824.

Wiswall, Matthew, and Basit Zafar. 2018. "Preference for the workplace, investment in human capital, and gender." *The Quarterly Journal of Economics*, 133(1): 457–507.

Wiswall, Matthew, and Basit Zafar. 2021. "New approaches to understanding choice of major." NBER Reporter (2): 18–21.

Zafar, Basit. 2011. "How do college students form expectations?" *Journal of Labor Economics*, 29(2): 301–348.

Zafar, Basit. 2013. "College major choice and the gender gap." *Journal of Human Resources*, 48(3): 545–595.

Zhang, Shaojun. 2024. "Carbon returns across the globe." *The Journal of Finance*. Available at https://doi.org/10.1111/jofi.13402.

Affidavit according to § 8 paragraph 3 of the doctoral degree regulations of February 17, 2015

I hereby affirm in lieu of oath that I have prepared the submitted work independently and without using any other than the specified aids. Statements, data and concepts taken directly or indirectly from other sources are identified with reference to the source. In the selection and evaluation of the following material, the persons listed below helped me in the manner described, for a fee / free of charge (underline as appropriate): /

Other people, in addition to the co-authors listed in the introduction to the work, were not involved in the content-related preparation of the present work. In particular, I did not make use of the paid help from mediation or advisory services. Nobody has received direct or indirect monetary benefits from me for work that is related to the content of the submitted dissertation.

The thesis has not yet been submitted to another examination authority in the same or a similar form, either in Germany or abroad.

I assure you that to the best of my knowledge, I have told the pure truth and have not concealed anything.

I confirm that the submitted electronic version corresponds in full to the submitted printed version.

I am aware of the criminal liability of a false affidavit, namely the threat of punishment according to § 156 StGB with an imprisonment of up to three years or with a fine in the case of intentional commission of the act or according to § 161 Paragraph 1 StGB up to one year imprisonment or fine in the case of negligent commission.

Köln, 30.11.2024