The London School of Economics and Political Science

Essays on Firms, Innovation, and Culture

Kieu-Trang Nguyen

A thesis submitted to the Department of Economics of the London School of Economics and Political Science for the degree of Doctor of Philosophy. London, January 2019.

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that this thesis consists of approximately 70,000 words.

Statement of conjoint work

I confirm that Chapter 2 is jointly co-authored with Antoine Dechezleprêtre, Elias Einiö, Ralf Martin, and Professor John Van Reenen, and I contributed 50% of this work.

Abstract

This thesis examines innovation and culture within the firm. The first chapter provides evidence on the effect of trust on innovation within firms. I build a new matched CEO-firm-patent dataset, and exploit variations in (i) generalized trust across the countries of CEOs' ancestry, inferred from their last names, and (ii) CEOs' bilateral trust towards inventors' countries ancestry or R&D labs, both yielding an effect of around 6% more future patents for a standard deviation increase in trust, controlling for stringent fixed effects. Trust-induced innovation is driven by higher-quality patents, consistent with a model in which CEO's trust incentivizes researchers to undertake high-risk explorative R&D. Finally, CEO's generalized trust is strongly correlated with broader corporate culture of trust, measured from online employee reviews. The evidence provides a micro-foundation for the well-known macro relationship between trust and growth.

The second chapter presents evidence of causal impacts of R&D tax incentives on innovation and technological spillovers using administrative data. Our Regression Discontinuity Design exploits a change in the size threshold that determines eligibility for R&D tax subsidies, and uncovers their large effects on R&D and patenting up to 7 years after the change. R&D tax price elasticities are large (lower bound of 1.1), as treated firms are smaller, and more likely financially constrained. Neighboring firms in small technology class with treated firms enjoy positive spillovers.

The third chapter shows strong positive spillovers of privatization on firms' total factor productivity through backward linkages in Vietnam. 10% more market share of privatized firms in downstream industries is associated with 4 percentage points increase in TFP. The effect is driven by privatization in local markets, is stronger (weaker) in upstream industries facing more import (export), and in provinces with higher entry costs. It likely works through elevated pressure from privatized client firms.

Acknowledgments

Throughout my long but exciting PhD journey, I am continually inspired by and greatly indebted to many amazing individuals: my supervisors, my colleagues, my friends, and my family, without whom this thesis would not have come into existence.

I would like to thank Oriana Bandiera, Catherine Thomas, and John Van Reenen for their extensive advice and guidance, as well as their never-ending support and encouragement. I am also grateful to Philippe Aghion, Pierre Azoulay, Clare Balboni, Andres Barrios Fernandez, Robin Burgess, Filipe Campante, Antoine Dechezleprêtre, Quoc-Anh Do, Andreas Ek, Benjamin Enke, Daniel Ferreira, Torsten Figueiredo Walter, Robert Gibbons, Juanita Gonzàlez-Uribe, Jesus Gorrin, Moqi Groen-Xu, Daniel Gross, Chao He, Emeric Henry, Dana Kassem, Danielle Li, Rocco Macchiavello, Alan Manning, Steve Pischke, Yona Rubinstein, Raffaella Sadun, Scott Stern, Donald Sull, Claudia Steinwender, Eric Van den Steen, Anh N. Tran, Luigi Zingales, and seminar participants at the LSE, MIT, and other institutions for their valuable help and comments. I am particularly thankful to Donald Sull and Anh N. Tran for having generously shared their data. In addition, I thank my classmates at the LSE, Clare Balboni, Matteo Benetton, Carlo Cabrera, Thomas Drechsel, Andreas Ek, Torsten Figueiredo Walter, Dana Kassem, Sevim Kosem, Tsogsag Nyamvadaa, Arthur Seibold, Guo Xu, and many others, for being the best colleagues and friends one could have asked for.

I also much appreciate the help of the staff at the HMRC Datalab and the administrative and academic support staff at the LSE, especially Mark Wilbor, Linda Cleavely, Jane Dickson, Paul Horsler, Manpreet Kheera, Mary Yacoob, and Yee Wan Yau, who have all made things easier than they should have been.

Last, but by no means least, I am deeply indebted to Quoc-Anh Do and our son Quoc-Thanh for being my constant source of inspiration and support, and to our mothers and family for always being there for me.

Contents

List of Tables

List of Figures

Chapter 1

Trust and Innovation within the Firm: Evidence from Matched CEO-Firm Data

This chapter provides evidence on the effect of trust on innovation within firms. I build a new matched CEO-firm-patent dataset covering 5,753 CEOs in 3,598 large US public firms and 700,000 patents during 2000-2011. To identify the effect of CEO's trust, I exploit variation in generalized trust across the countries of CEOs' ancestry, inferred from their last names using de-anonymized historical censuses, as well as variation in CEOs' bilateral trust towards inventors. First, one standard deviation increase in CEO's generalized trust following a CEO turnover is associated with over 6% increase in firm's future patents. Second, changes in CEO's bilateral trust towards inventors in different countries (i.e., different R&D labs within multinational firms) or from different ethnic origins in the same firm have comparable effects on inventors' patenting, controlling for CEO and other stringent fixed effects. Trust-induced improvements in innovation are driven entirely by higher-quality patents, consistent with a model in which CEO's trust incentivizes researchers to undertake high-risk explorative R&D. Finally, I show that across and within firms, CEO's generalized trust is strongly correlated with a broader corporate culture of trust, as measured from the text analysis of one million online employee reviews. The evidence provides a micro-foundation for the well-known macro relationship between trust and growth.

"Virtually every commercial transaction has within itself an element of trust."

—Arrow (1972)

1.1 Introduction

Arrow (1974, c.1, p.23) emphasized trust as "an important lubricant of a social system," as it is impossible to fully contract upon all possible states of nature.[1] This insight is especially relevant in the context of research and innovation, in which the inherent uncertainty of research makes contracts necessarily incomplete (Arrow, 1962). It is thus essential to understand the relationship between trust and innovation, in order to better understand how to incentivize innovation, an important driver of growth.

This paper studies the role of trust on innovation within the firm. I develop a theoretical framework in which the CEO's trust in good researchers encourages their risk-taking, thereby increasing innovation.[2] To test this effect, I assemble a large matched CEO-firm dataset covering 5,753 US CEOs in 3,598 US public firms between 2000 and 2011, associated with 700,000 patents and inventors. I infer a CEO's ethnic origins from her last names using de-anonymized historical US censuses, then measure CEO's inherited generalized trust as ethnic-averaged trust in US samples. Similarly, I compute CEO's bilateral trust towards researchers in different ethnic groups within the same firm based on average bilateral trust between countries (highlighted in Guiso et al., 2009). First, I exploit CEO turnovers to estimate the effect of CEO's trust on patent counts and future citations. Second, I estimate how a CEO's bilateral trust towards different ethnic groups affects patents filed by inventors in different overseas R&D labs within multinational firms, and by inventors of different ethnic origins in the same US firm, in each case controlling for firm by year, CEO, and inventor country fixed effects. I further examine trust's effect on the distribution of patent quality to differentiate risk-taking from possible alternative mechanisms.

I model the process of research based on Arrow's (1962) insight that research is inherently uncertain, and by nature difficult to observe and contract on researchers' behaviors. In a simple two-period principal-agent model between a CEO and a researcher, the researcher's type and actions are private information, and only outcomes

[1] Arrow's general view on trust has received ample macroeconomic empirical support on the association of trust and development and growth, as surveyed by Algan and Cahuc (2013, 2014). Knack and Keefer (1997), La Porta et al. (1997), Guiso et al. (2004, 2006, 2009), Tabellini (2010), and Algan and Cahuc (2010), among others, provide evidence that trust is a deep-root determinant of development and growth, through its channels of influence on the accumulation and allocation of factors of production (such as investments, loans, allocation of capital). This economic literature has built on seminal work by sociologists and political scientists on trust and development, including Banfield (1958), Gambetta (1988), Coleman (1990), Putnam et al. (1993), Putnam (2000), Fukuyama (1995), and others.

[2] While in principle all relations involving researchers inside the firm may matter to its innovative activities, it has been suggested that, in practice, the chief executive may wield significant influence on the firm's culture (e.g., Guiso et al., 2015; DeBacker et al., 2015; Liu, 2016, among others), therefore their trust is of first order importance in studying the firm's trust towards researchers.

are observable. In each period, a "good" researcher faces the choice between (i) exploration, a high-risk high-return project that can result in innovation or failure with probability known only to him, and (ii) exploitation, a risk-free low-return common path that surely signals a good type from a bad one.[3] On the other hand, failure means either that the researcher's exploration is unsuccessful, or that he is a bad type.[4] Considering those elements, the CEO, who inherits a pool of existing researchers in period 1, decides to whether rehire or fire the researcher in period 2 based on his period 1's outcome.

In this setting, the CEO's trust in the researcher is modeled as her prior belief about his type.[5] A more trusting CEO is more likely to interpret observed failure as being due to bad luck rather than bad type, therefore more likely to tolerate failure. In anticipation, a good researcher will be more likely to undertake exploration, thereby producing more innovation. The model thus predicts that higher trust induces more innovation through encouraging exploration (versus exploitation). These results resonate with Manso (2011) and Aghion et al. (2013), whose models also imply that tolerating failure and reducing career risk help induce risky innovation. However, while Manso (2011) suggests that this objective could be achieved with long-term incentives and Aghion et al. (2013) with monitoring, my model instead emphasizes the enabling role of trust.[6] It also highlights the possible suboptimality of excessive trust due to too much retention of bad researchers,[7] and the role of trust as substitute for the CEO's commitment capacity.[8]

I turn to matched CEO-firm and patent data in the US to study the empirical relationship between CEO's trust and innovation. Innovation outputs, measured by patent counts,[9] are extracted from PATSTAT, a dataset covering close to the universe of patents ever filed from 1900 up to 2016 with 70 million patent documents from over 60 major patent offices all over the world, including the US Patent and Trademark Office

[3]March (1991) first emphasized the trade-off between exploration and exploitation in the context of research and innovation. I follow Manso (2011) in modelling research as the choice between exploration and exploitation. Unlike Manso (2011), who studies the implementation of either path, I focus on how the CEO's prior belief on the researcher's type, i.e., trust, affects innovation outcomes.

[4]In this setting, a bad research is understood as someone who lacks in ability or willingness to undertake appropriate courses of actions. By normalization, I assume that the bad type always fails.

[5]My choice to model trust as a belief reflects Gambetta's (1988) definition of trust as "the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action." Similar approach has been used in Guiso et al. (2008) and Bloom et al. (2012), among others.

[6]More broadly, this paper relates to the literature on contractual and financial arrangements to incentivize innovation (surveyed by Ederer and Manso, 2011), which includes for example Lambert's (1986) consideration of incentivizing an agent's risk taking, and Azoulay et al.'s (2011), Ederer and Manso's (2013), and González-Uribe and Groen-Xu's (2017) evidence of the effects of Manso-type contractual incentives (i.e., tolerance of failure and long-term incentives) on innovation.

[7]This result complements Butler et al. (2016) finding on the "right amount of trust," which suggests that highly trusting individuals tend to assume too much social risk while individuals with overly pessimistic beliefs can give up profitable opportunities.

[8]That is, when commitment to tolerance of failure is not possible, trust helps implement it. This result formalizes the intuition on the reliance between trust and commitment studied in the literature of sociology of organization, such as Klein Woolthuis et al. (2005).

[9]The measure of innovative outputs by patents, correcting for quality or not, is not perfect (Hall et al., 2014). However, to the extent that the use of patents to protect notable innovations is common within an industry, my focus on patents is unlikely subject to a serious bias if I only consider within-firm or within-industry variations. At worst, it likely underestimates the effect of trust on innovation.

(USPTO) and its counterparts in Europe (EPO) and Japan (JPO).[10] In addition to patent counts, I also refine data on patent citations, technological class, family, and inventors' names and addresses to further investigate the mechanism at work.

Detailed data on the background of CEOs and top officers of US public firms are provided by BoardEx. My focus on CEOs is motivated by a growing body of empirical evidence that individual CEOs matter for firm decisions and performance (e.g., Bertrand and Schoar, 2003; Bennedsen et al., 2010; Smith et al., 2017, and the Bertrand, 2009 survey). It proposes a new factor, i.e., trust, that fits the description of manager styles as coined by Bertrand and Schoar (2003) and that contributes to the strand of literature studying how differences in CEO traits relate to differences in firm performance.[11]

Building on the literature on transmitted and inherited cultural values that highlights the role of cultural origin in shaping an individual's cultural traits,[12] I measure a CEO's trust based on her ethnic origins as inferred from her last name and measures of inherited trust among descendants of US immigrants. First, I construct a probabilistic mapping between CEO's last names and ethnic origins from four de-anonymized US censuses.[13] Second, I compute an ethnic-specific measure of trust for 36 different ethnic origins most common in the US using responses to the trust question in the US General Social Survey (GSS).[14] I only select survey answers from GSS respondents in highly prestigious occupations similar to the CEO sample. Each CEO's inherited trust measure is the weighted average of ethnic-specific trust based on her likely ethnic composition.

To assess the role of generalized trust, the first empirical strategy uses firm fixed effects to exploit changes in CEOs and subsequent changes in patenting within the same firm over time, controlling for CEO observable characteristics such as age, education, and tenure in the firm. The identifying condition is supported by the empirical evidence that both timing of CEO change and the new CEO's trust are not related to the firm's past patenting activities. I find that one standard deviation in CEO's inherited generalized trust, equivalent to the shift from Greek to English, is associated with 6% increase in the number of annual patents filed. This result is robust to a large set of controls for country of origin characteristics and ethnic of origin socioeconomic

[10] The PATSTAT dataset is thus much more general and suitable for studies with a cross-country perspective than the usual USPTO dataset. PATSTAT data have since recently been used in research on innovation, such as Dechezleprêtre et al. (2018).

[11] Recent studies have started to explore a broad range of CEO characteristics (Malmendier and Tate, 2005, 2009; Kaplan et al., 2012; Kaplan and Sørensen, 2017; Gow et al., 2016) and practices (Bandiera et al., 2015, 2017). In particular, this literature has also considered certain aspects of CEO cultural background such as corruption culture and firms' misconducts (DeBacker et al., 2015; Liu, 2016).

[12] E.g., theoretical foundation by Bisin and Verdier (2000), Bisin and Verdier (2001), Tabellini (2008), Guiso et al. (2016); empirical evidence by Giuliano (2007), Fernández and Fogli (2009), Algan and Cahuc (2010), among others.

[13] The four US censuses from 1910 to 1940 contain 80 million individuals with foreign birth places or ancestry, sharing among them five million unique last names, out of which 75,000 last names appear for at least 100 times each. 83% of CEO last names are among those 75,000 sufficiently common last names. The inference of ethnic origin from last names was pioneered by Kerr and Lincoln (2010).

[14] This approach follows Guiso et al. (2006), Algan and Cahuc (2010), and the related literature. I also perform a robustness check with data from the World Value Survey (WVS). Those two surveys cover most of cross-country research on the economics of trust since Knack and Keefer (1997) and La Porta et al. (1997).

conditions and cultural traits, suggesting that it is not driven by other ethnic-related characteristics.

To separate the role of trust from other CEO's unobservable characteristics such as management style or ability, the second empirical strategy exploits within-CEO variation in CEO's bilateral trust towards different ethnic groups and patents by inventors from those different ethnicities, which allows for a full set of stringent, including CEO, fixed effects. Bilateral trust measures are calculated between CEO's inferred ethnic origin and countries of inventors using Eurobarometer data.[15] Patent inventors' countries of origin are obtained from either their addresses (for inventors in overseas R&D labs of multinational firms) or their last names (for US-based inventors). Under the same CEO, one standard deviation increase in bilateral trust towards an inventor country of origin is associated with 3-5% more patents by inventors from the corresponding R&D lab or corresponding ethnicity, controlling for a broad range of time-variant fixed effects at the firm by year, CEO, and inventor country levels. These results remain stable even in the presence of firm by inventor country fixed effects (i.e., using variation in changes in bilateral trust following CEO changes), and after accounting for possible alternative explanations such as favoritism or better information flows between CEOs and researchers.

To distinguish the proposed mechanism that trust induces innovation via encouraging risk-taking and exploration from other mechanisms in which trust induces more effort by researchers,[16] I develop a formal framework to identify the mechanisms via their different implications on trust's effect on the distribution of patent quality. Using future citation counts and other patent quality measures, I show that, consistent with the risk-taking mechanism, trust increases only high-quality patents, but not low-quality ones, thereby increasing average patent quality as measured by citations per patent by 4%. In addition, I find that trust is most effective in inducing innovation in firms with likely high researcher quality.

These results on the effect of CEO's trust on firm innovation provide a possible micro-foundation for the macro relationship between long-term economic outcomes and trust, as previously evidenced in Guiso et al. (2006), Tabellini (2010), and Algan and Cahuc (2010), among others.[17] As it shows that trust can spur innovations by solving contractual shortcomings, a high-trust society possesses not only the advantage

[15] These bilateral trust measures have been exploited by Guiso et al. (2009) in the context of international trade, Bloom et al. (2012) in delegation to subsidiaries, Giannetti and Yafeh (2012) in syndicated bank loan interests, Ahern et al. (2015) in mergers and acquisitions, and Bottazzi et al. (2016) in venture capital flows.

[16] For example, there is a large literature on delegation in organization since Aghion and Tirole (1997), such as Acemoglu et al. (2007) and Bloom et al. (2012). In the context of innovation, trust, understood as the preference congruence between the principal and the agent, would lead to more delegation to researchers, which then induce them to put in more effort, thereby producing more innovation.

[17] The larger literature on the cultural origins of long-term economic development has discussed the role of religion (Barro and McCleary, 2003, 2018), work ethic (Becker and Woessmann, 2009), individualism (Gorodnichenko and Roland, 2017), and others as surveyed by Nunn (2012). The macro correlation between trust and innovation has been briefly suggested by Hall and Jones (1999) on TFP and Algan and Cahuc (2014) on R&D and patents.

of higher investment and accumulation of factors of production (or even better allocative efficiency), but also the potential to invent more and thus grow productivity faster in the long run.[18] This mechanism thus helps explain the macroeconomic differences not only in development levels but also in growth rates across countries.[19] Separately, this paper extends the empirical literature of more traditional determinants of R&D and patents, such as tax credit and grant (e.g., Howell, 2017; Dechezleprêtre et al., 2018), as surveyed by Cohen (2010).

I further link CEO's culture to firm's culture, measured from text analysis of almost one million employee reviews on Glassdoor.com, one of the largest career intelligence websites worldwide (from Sull, 2018). The dataset covers many dimensions of employees' sentiments based on O'Reilly et al. (1991, 2014) across 500 large US public firms between 2008 and 2017 (similar to Grennan, 2014). In different specifications with CEO controls, industry fixed effects, and even firm fixed effects (i.e., using variation in changes in CEO's trust following CEO changes), CEO's inherited trust is associated with stronger corporate trust culture. In that regard, this paper also provides new findings supporting the role of corporate culture in determining corporate outcomes.[20] Furthermore, it shows a channel through which corporate culture can be influenced: by an injection of culture from the top (as suggested by Van den Steen, 2010).

Beyond the economics literature, the interplay between management and trust and other cultural traits has been examined in sociology of organization and management, e.g., in classic studies by O'Reilly et al. (1991, 2014), and other work on organization culture such as Schein (1985) or Hofstede et al. (1991).[21] My results broaden this literature with a large-scale sample of firms, and with inherited trust computed systematically from surveys of opinions.

The rest of the paper is organized as follows. Section 1.2 discusses the model of trust and innovation. Section 1.3 provides descriptions of the data. Sections 1.4 and 1.5 describe the within-firm and within-CEO empirical strategies and the corresponding empirical results. Section 1.6 studies the mechanism through risk-taking and exploration. Section 1.7 provides further discussions and interpretations, and section 1.8 concludes.

[18] This statement holds in the large class of endogenous growth model à la Aghion and Howitt (1992) in which sustained innovation maintains long-term growth.

[19] From a macro perspective, Doepke and Zilibotti (2014) summarizes theories on the relationship between cultural traits (such as risk attitude, patience, and trust), entrepreneurship, and growth. Reviews by Durlauf et al. (2005) and Caselli (2005) provide evaluations of the crucial roles of productivity growth in explaining cross-country differences in growth and income level, respectively.

[20] E.g., Guiso et al. (2015), Grennan (2014), Gibbons and Kaplan (2015); Martinez et al. (2015), Graham et al. (2018).

[21] Notably, the management literature has considered the culture of trust in firms as crucial to innovation (Nooteboom and Stam, 2008), and can be substitute for or complement to formal control (Knights et al., 2001; Klein Woolthuis et al., 2005).

1.2 Theoretical framework

This section models how CEO's trust could affect researchers' choices and consequently innovation outcomes. As "trust is an important lubricant of a social system" (Arrow, 1974), it is likely to also impact innovation through other different mechanisms. Therefore, it should be noted that my choice to focus on the CEO's trust (instead of the researcher's trust) and this particular model is guided by the empirical evidence presented in the latter part of the paper.

1.2.1 A model of trust and innovation

Set up. My starting point is a two-period principal agent game with asymmetric information in which the principal is the CEO and the agent is the researcher.

Researcher. The researcher could be good type with probability θ or bad type with probability $1-\theta$. In this setting, a bad researcher is understood as someone who lacks ability or willingness to take the appropriate courses of actions. The CEO, who is not an expert in science, knows neither the researcher's type nor θ. In each period, a bad researcher always shirks and produces s^L, while a good researcher chooses between exploitation and exploration. Exploitation is a low-cost, safe R&D project that requires no effort cost and produces s^M with certainty. Exploration is a high-cost, risky R&D project that requires effort cost c and produces s^H (innovation) with probability π and s^L (failure) with probability $1-\pi$.[22] π is independently drawn from the unit uniform distribution in each period and privately observed by the good researcher before choosing which project to pursue. The CEO does not know what project is chosen and only observes the outcome produced by the researcher at the end of each period.[23]

CEO. The CEO asks the researcher to carry out R&D at the beginning of period 1 without knowing his type. Simultaneously, she decides on an outcome-contingent contract that maps period 1's potential outcome s^i to (b_1^i, D^i) $(i \in \{L, M, H\})$ where b^i is a bonus on top of fixed wage w for the researcher and $D^i \in \{0, 1\}$ denotes whether the CEO would fire ($D^i = 0$) or rehire ($D^i = 1$) the researcher after period 1. If the researcher is rehired, the game continues to period 2, in which the CEO specifies contract (b_2^i) and the researcher chooses from the same action set as described. The game ends after period 2's outcome and payment are realized. In the baseline model, I assume that the CEO can credibly commit to the contracts specified at the beginning of each period.[24]

[22]The trade-off between the exploitation of well-known approaches and the exploration of new untested approaches was first emphasized by March (1991) and has since then been widely studied both theoretically and empirically (see survey by Ederer and Manso, 2011).

[23]The values of s^L, s^M, s^H, c and π's uniform distribution are common knowledge and satisfy $s^H - c > s^M > 0 > s^L$. The key results of the model remain under more general assumptions about the distribution of π.

[24]Alternatively, the contract can be designed as a mapping (specified at the beginning of period 1) between each potential outcome of the game (i, j) to (b^{ij}, D^i) where $i, j \in \{L, M, H\}$ denote the game's outcomes in periods 1 and 2 respectively. This set up is equivalent to the baseline under the assumption of credible commitment. I consider relaxing this assumption in subsection 1.2.2.

Trust. Although the CEO does not observe θ, she has her own prior subjective belief that the researcher is good with probability θ^P, which reflects her trust level towards the agent. A more trusting CEO would have a higher subjective θ^P than a less trusting one.[25] This model focuses on studying how this key parameter of CEO's trust affects her and the researcher's strategies in the game and its outcome.[26]

Payoffs and restrictions. After each period $t \in \{1, 2\}$ with realized outcome $i \in \{L, M, H\}$, the researcher gets $w + b_t^i - c$ (if he chooses exploration) or $w + b_t^i$ (otherwise) and the CEO gets $s^i - b_t^i$.[27] If the researcher is fired at the end of period 1, both players' payoffs in period 2 is zero.[28] The researcher has limited liability and $b_t^i \geqslant 0 \; \forall \, i, t$. That is, the CEO can reward the researcher for good performance but cannot financially punish him for bad outcome. I also restrict the parameters to satisfy the players' participation constraint, which implies that the CEO's expected payoff from hiring a good researcher is positive. It then follows that $D^i = 1$ for $i \in \{H, M\}$, as these outcomes fully reveal that the research is the good type. However, if period 1's outcome is L, the CEO cannot tell if the researcher is bad or if he is good but unlucky. Her choice of whether to tolerate period 1's failure D^L depends on her assessment of which is more likely to be the case, and this assessment depends on her prior subjective belief θ^P.

Solution outline. I first consider two separate cases in which $D^L = 1$ and $D^L = 0$, then compare the CEO's expected payoffs under these two cases to solve for her optimal choice of D^L. Note that a good researcher's choice in period 2 (conditional on its happening) is independent of period 1's outcome and therefore is the same under both cases. Thus, let V_2^P denote the CEO's period-2 expected payoff from hiring a good researcher and let V_2^A denote a good researcher's period-2 expected payoff, both are positive under the participation constraints discussed earlier.[29]

Researcher's project choice: explore versus exploit. As a bad researcher always shirks, the meaningful action to analyze is a good researcher's choice between exploration and

[25]This concept of trust resonates with Gambetta's (1988) definition of trust as "the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action." Similarly, Guiso et al. (2008) also model trust as a subjective belief about being cheated by the counterpart in a financial transaction.

[26]As the CEO's prior belief, or trust, affects R&D outcomes via influencing the researcher's choice, the latter's perception of the former's belief is as important as the belief itself. This is especially true in real life settings where CEOs' influence on firm's policies takes time to materialize and credible commitment to such policies is unlikely. In these settings, the researcher's perception of the CEO's beliefs and preferences is likely based on the collective reputation of the group to which the CEO belongs in addition to the CEO's own reputation based on her past actions.

[27]w is a fixed wage that is set exogenously. For notation simplicity, s^L, s^M, and s^H represent R&D project's returns after fixed wage payment, so w does not enter the CEO's payoff. I assume that both players are risk neutral and do not discount future payoff. Introducing risk aversion or time discounting does not affect the model's key insights.

[28]The assumption is that the CEO and the researcher cannot immediately find new matches in period 2.

[29]It can be shown that period 2's subgame has a unique Nash equilibrium in which the CEO chooses $(b_2^H, b_2^M, b_2^L) = (b_2^*, 0, 0)$ and the good researcher chooses to explore when $\pi_2 > c/b_2^*$. V_2^P and V_2^A are functions of s^L, s^M, s^H, c, w.

exploitation in period 1 given D^L. To reduce notations, I omit the outcome superscript L from D^L and the period subscript 1 from π^1, b_i^1 for the rest of this subsection.

A good researcher chooses exploration over exploitation when it yields higher expected payoff:

$$\begin{aligned} \pi(w + b^H + V_2^A) + (1-\pi)(w + b^L + DV_2^A) - c > w + b^M + V_2^A \\ \iff \pi > \bar{\pi}(D).^{30} \end{aligned} \tag{1.1}$$

The above condition implies that in both cases ($D = 1$ and $D = 0$), the researcher follows a cutoff strategy and chooses exploration when the realized probability of success π is above threshold $\bar{\pi}(D)$.

Given the good researcher's strategy, the CEO indirectly chooses $\bar{\pi}(D)$ via setting the bonuses to maximize her expected payoff from hiring a good researcher. It is optimal for her to set b^L and b^M to zero and only vary b^H to achieve her desired $\bar{\pi}(D)$ threshold.[31] For each value of D, the CEO's maximization problem then has a unique solution $b^*(D)$ that induces the good researcher to explore when π is above threshold $\bar{\pi}^*(D) = \frac{c+(1-D)V_2^A}{b^*(D)+(1-D)V_2^A}$. This leads us to the following proposition:

Proposition 1. *For a given set of parameters, $\bar{\pi}^*(1) < \bar{\pi}^*(0)$. That is, tolerance of failure induces more exploration and innovation.*

The proof is detailed in appendix A.1.1. The intuition is that period 1's exploration threshold $\bar{\pi}^*(.)$ is increasing in $(1-D)V_2^A$ and $(1-D)V_2^P$, which represent a good researcher's and the CEO's foregone period-2 payoffs after a bad outcome in period 1.[32] When failure is not tolerated and termination implies a large loss in future payoff, a good researcher requires higher probability of success to undertake exploration. Similarly, the CEO also prefers a good researcher to take less exploration risk in period 1 for the same fear of losing her future payoff from a relationship with such good researcher when exploration fails. Put differently, tolerance of failure enables a good researcher to take more risk and explore more, which then produces more instances of successful innovation.[33] This result resonates with Manso's (2011) insights that the optimal incentive scheme to motivate exploration exhibits tolerance for early failure and reward for long-term success.[34]

[31] That is, she chooses b^H to maximize:

$$\int_0^{\bar{\pi}(D)} \left[s^M + V_2^P\right] d\pi + \int_{\bar{\pi}(D)}^1 \left\{\pi\left[s^H - b^H + V_2^P\right] + (1-\pi)\left[s^L + DV_2^P\right]\right\} d\pi,$$

where $\bar{\pi}(D) = \frac{c+(1-D)V_2^A}{b^H+(1-D)V_2^A}$ is also a function of b^H.

[32] When $D = 1$, these losses are zero. When $D = 0$, these losses depend on V_2^A and V_2^P.

[33] $\bar{\pi}^*(1) < \bar{\pi}^*(0) \implies \int_{\bar{\pi}^*(1)}^1 \pi d\pi > \int_{\bar{\pi}^*(0)}^1 \pi d\pi$. Furthermore, $\bar{\pi}^*(1)$ represents the optimal level of exploration for the CEO in a single-period game.

[34] Azoulay et al. (2011) and Tian and Wang (2014) among others provide empirical evidence that tolerance of failure induces more innovation in different settings.

CEO's tolerance of failure: rehire versus fire. Let $V_1^P(D)$ denotes the CEO's period-1 expected payoff from hiring a good researcher under policy $D \in \{0, 1\}$. It can be shown that $V_1^P(1) > V_1^P(0) > 0$.[35] The CEO chooses to tolerate failure (i.e., $D^L = 1$) if it maximizes her total expected payoff:

$$\theta^P \left[V_1^P(1) + V_2^P \right] + (1 - \theta^P) s^L > \theta^P \left\{ V_1^P(0) + \left[1 - \frac{(1 - \bar{\pi}^*(0))^2}{2} \right] V_2^P \right\}$$
$$\iff \theta^P > \bar{\theta}.^{36} \qquad (1.2)$$

Proposition 2. *The manager chooses $D^L = 1$ iff $\theta^P > \bar{\theta}$. That is, she chooses to tolerate failure when her trust towards the researcher is high enough.*

This is a direct result from inequality 1.2. Intuitively, when observing a bad outcome, a more trusting CEO ascribes more weight to the researcher's being unlucky than him being of the bad type. As the benefits of incentivizing optimal exploration then outweighs the benefits of screening out bad researchers, she chooses to tolerate failure to avoid mistakenly screening out good researchers in period 2 and also to induce more exploration in period 1.

Combining Propositions 1 and 2 yields the prediction that a more trusting CEO induces more innovation, which is the focus of this paper's empirical investigation.

1.2.2 Model extensions

First, I relax the assumption that the CEO can credibly commit to being tolerant of failure. Appendix A.1.2 shows that in this setting tolerance of failure is a unique equilibrium only when $\theta^P > \bar{\theta}_{post} > \bar{\theta}$ where $\bar{\theta}_{post}$ is an unique cutoff based on the game's parameters.[37] Furthermore, for $\theta^P \in (\bar{\theta}, \bar{\theta}_{post})$ there always exists an equilibrium in which the CEO does not tolerate period 1's bad outcome, even though it is *ex ante* optimal for her to do so. This equilibrium is even the unique one in some cases. Such problem can be avoided if the CEO can *ex ante* credibly commit to the being tolerant of failure as in the baseline model, or if she is high trusting with $\theta^P > \bar{\theta}_{post}$. In other words, trust acts as a substitute for commitment.

Next, I allow a bad researcher to also be able to produce exploitation outcome with some luck (i.e., with probability q). In this setting, as only innovation outcome (i.e., successful exploration) fully reveals a researcher's type, would a good researcher explores *more* under a less trusting CEO in order to separate himself from the bad ones, even when it is risky to do so? I find that this is not the case unless q is large, for

[35] Under $D^L = 0$, the good researcher is less willing to choose exploration than what is optimal for the CEO. In addition, the CEO also has to provide additional exploration incentive for the good researcher through bonuses (i.e., $b^{H*}(0) > b^{H*}(1)$). As a result, $V_1^P(1) > V_1^P(0)$. Note that $V_1^P(D)$ is function of s^L, s^M, s^H, c, w, and D.

[37] In this setting, the CEO's decision whether to tolerate failure is based on her updated belief at the end of period 1 with the aim to maximize her period-2 payoff. As a result, she does not internalize the gain from optimal exploration in period 1 under tolerance of failure and therefore is less likely to tolerate period 1's bad outcome (i.e., $\bar{\theta}_{post} > \bar{\theta}$).

exploitation still provides signaling value for a good researcher when a bad researcher is not too likely to produce the same outcome by luck. Therefore, a less trusting CEO induces more exploitation and less exploration and vice versa, as in the baseline model.[38]

Third, I extend the model to three periods to study if a longer horizon strengthens the CEO's incentive to screen out bad researchers in earlier periods and induces her to adopt a different strategy. The key intuitions of the two-period game go through in this three-period game. A high-trust CEO always rehires the researcher after a bad outcome; an average-trust CEO tolerates first time failure but not the second time; and a low-trust CEO fires the researcher after first time failure in period 1. A good researcher chooses to explore at the optimal level under a high-trust CEO but undertakes less exploration when the termination threat worsens the downside of failure. As in the baseline model, higher trust maps into higher tolerance of failure and induces more exploration and innovation.[39] The results from this three-period game suggest that the findings extend to in longer-horizon settings.

Finally, how does a CEO with subjective prior θ^P compare to one knowing the true quality of the researcher pool θ? The model implies higher trust *always* induces more innovation, but also more failure. As a result, when the researcher pool is generally bad but the CEO is too trusting, tolerance of failure leads to costly excessive innovation. On the other hand, when the researcher pool is generally good but the CEO is too distrusting, intolerance of failure leads to inefficiently low level of innovation. Furthermore, when the CEO cannot credibly commit to her policies, a more trusting CEO still can outperform an objective one, as then trust helps substitute for commitment. Subsection 1.6.3 provides evidence consistent with these implications that CEO's trust effects on both innovation and firm's performance are larger among firms with likely better researcher quality.

1.3 Data and measurement

1.3.1 Patents as a measure of innovation

I follow the innovation literature in using patent and citation counts as measures for innovation (e.g., Trajtenberg, 1990; Bloom and Van Reenen, 2002; Hall et al., 2005).[40] My patent data come from PATSTAT, the largest available international patent database which covers close to the population of all worldwide patents since the 1900s up to 2016. It brings together nearly 70 million patent documents from over 60 patent offices, including the United States Patent and Trademark office (USPTO) and all other major

[38]The proof for this is available upon request.

[39]The proof for this is available upon request.

[40]As previously mentioned, the measure of innovative outputs by patents, correcting for quality or not, is not perfect (Hall et al., 2014). However, to the extent that the use of patents to protect notable innovations is common within an industry, my focus on patents is unlikely subject to a serious bias if I only consider within-firm or within-industry variations. At worst, it likely underestimates the effect of trust on innovation.

offices such as the European Patent Office (EPO) and the Japan Patent Office (JPO). I assign patents to firms using the matching procedure implemented by the OECD and made available via Bureau van Dijk's ORBIS platform.[41]

The dataset contains comprehensive information from the patent record, including application and publication dates, backward and forward citations, technology classification, and patent family. These data allow me to construct various measures of patent quality besides forward citation counts, such as backward citations to scientific literature, patent scope, generality index, originality index, etc. (details in appendix A.2.2). In addition, PATSTAT also provides information on the inventors of each patent, including their names and addresses, as are available on the patent record. This further enables me to link patents to their inventors' countries of residence (based on their addresses) or countries of origin (based on their last names) to construct patent counts at the firm by inventor country level (details in subsection 1.5.1).

I consider only patents that are classified as "patent of invention" in PATSTAT (equivalent to USPTO's utility patents). To avoid double-counting inventions, I classify patents in the same patent family (i.e., a set of patents protecting the same invention across several jurisdictions) as one single patent, and assign the patent to the year of its earliest application date. Finally, PATSTAT's patent data are more comprehensive for the years before 2012, as it takes up to 1.5 years for a patent application to be published and on average 5 years for a patent to gain 50% of its lifetime citations (Squicciarini et al., 2013). As a result, I focus only on patents filed before 2012.

1.3.2 CEO's inherited trust measure

I obtain information on firms' CEOs, senior executives, and board directors of US publicly listed firms from BoardEx. The dataset spans from 2000 to 2016, covers almost all US publicly listed firms in this period, and includes rich information on the executives' background, employment history, and compensation. Among these variables, the executives' names are essential for the measurement of inherited trust, as explained below. In addition, I also use information on the timing of their positions, gender, education, employment history, and compensation (details in appendix A.2.3).

Measuring CEO's inherited trust. I measure a CEO's inherited generalized trust based on her ethnic origins inferred from her last name and measures of inherited trust among descendants of US immigrants. That is,

$$trust_d = \sum_e w_{de} \times ethtrust_e \tag{1.3}$$

[41]ORBIS also provides information on firm's ownership structure, which allows one to identify and include patents filed by a firm's subsidiaries.

where $ethtrust_e$ is the average trust measure among all descendants of US immigrants from country e and w_{de} is the probability that CEO d is a descendant of US immigrants from that country.[42]

I follow the literature on inherited trust (e.g., Guiso et al., 2006; Algan and Cahuc, 2010) in computing $ethtrust_e$ using individual-level data on trust attitude and ethnic origins from the US General Social Survey (GSS), a representative survey of social attitudes among US residents conducted between 1972 and 2014, covering a total of 60,000 respondents. A respondent's trust attitude is measured by the standard generalized trust question *"Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?"*.[43] His ethnic origin is captured by the question *"From what countries or part of the world did your ancestors come?"*, which covers 36 most common ethnic origins in the US, including almost all European countries in addition to Canada, Mexico, China, and India.[44] The baseline $ethtrust_e$ measure is then the average trust attitude of GSS respondents whose self-reported ethnic origin is e (see Table A.1). I only consider respondents in highly prestigious occupations (by GSS' classification), in order to better match the CEO sample.[45] In addition, I also construct an alternative trust measure that takes into account demographic characteristics such as gender, education, age, and birth cohort.

Next, I construct a probabilistic mapping between a CEO's last name and different ethnic origins (i.e., $\mathbf{W}_d$) using historical de-anonymized US censuses from 1910 to 1940 (e.g., Kerr and Lincoln, 2010; Liu, 2016).[46] These data contain individual-level data on birthplace and ancestry of the whole US population during 1910-1940, merged with information on individual names obtained from the Minnesota Population Center. Across four censuses there are 80 million individuals with foreign birthplaces or ancestry, sharing among them five million unique last names. I only consider 75,000 last names with at least 100 occurrences and allow each of them to be mapped to multiple ethnic origins with probabilities equal to their shares of occurrences. Separately, I also compile lists of most common last names in 50 different countries from various sources

[42]I exclude CEOs who are not US citizens. They comprise only 4.8% of the 54% of CEOs for whom BoardEx contains nationality information. A quick check reveals that the other 46% represent cases in which the CEOs are obviously US citizens, so that the firm's website does not state their nationality. They are thus counted as US citizens.

[43]Following the literature, I construct a trust indicator equal to 1 if the answer to is "Most people can be trusted," and 0 if the answer is "Can't be too careful" or "Other, depends." This grouping makes a clear separation between high trusting individuals as opposed to moderate or low trusting ones (Algan and Cahuc, 2010).

[44]37% of respondents report two or three countries of origin, in which case I select the one to which they feel the closest to. I also exclude 5 ethnic categories: "American Indian," "American only," "Other European," "Other Asian," and "Other," which together comprise only 9% of total respondents.

[45]Specifically, I only consider respondents whose GSS occupation prestige score is in the top 25% (i.e., 50 or above), which cover most respondents in management occupations. The correlation between $ethtrust_e$ computed from this sample and that computed from all GSS respondents is 0.85 (0.75).

[46]The US Census Bureau is allowed to release de-anonymized individual census records after 72 years.

and use these lists to cross-check and supplement the baseline census-based mapping (details in appendix A.2.4).[47]

83% of the CEO sample are mapped to their ethnic origins based on their last names. Panel A of Table A.3 shows that, there ar no significant differences between these name-matched 83% and the remaining non-matched 17% across all observable characteristics. Three most common ethnic origins among CEOs are Irish, German, and English, which together account for about half of the CEO sample (see Table A.2). The average CEO's inherited trust measure is 0.56, considerably higher than the average GSS trust measure of 0.38 but comparable to the average trust measure of 0.51 among respondents in highly prestigious occupations. Despite the high total shares of three most common ethnic origins among CEOs, Figure 1.1 shows that there remains meaningful variation in their inherited generalized trust measure.

Validity of inherited trust measure. There is a growing literature in economics that highlights the role of cultural origin in shaping individual trust and other cultural traits. Studies by Bisin and Verdier (2000, 2001), Tabellini (2008), and Guiso et al. (2016) provide theoretical mechanisms for cultural transmission of preferences and beliefs from parents to children. Empirically, a large body of evidence shows that trust attitude and other values among descendants of US immigrants are strongly correlated with related traits, behaviors, and outcomes of those in their home countries, consistent with intergenerational cultural transmission among US immigrants.[48] Following this literature, I verify the existence of trust transmission by comparing the measure of inherited trust among US immigrants, calculated from the GSS, with an alternative measure based on average trust attitude among the populations of the countries of origin, available from the World Value Survey (WVS). The correlation between the GSS- and WVS-based trust measures is above 0.5 at country level, consistent with the view that immigrants in the US inherit a large part of their cultural traits from their countries of origin, such as shown in Giuliano (2007).[49]

Ideally, one would like to observe each CEO's individual trust attitude, yet this latent variable is incredibly challenging to measure. Even if one could administer a trust survey or a trust game among CEOs, the resulting measure would still be affected by measurement error.[50] The inherited trust measure misses (i) the individual-specific

[47]One concern is this last name-based mapping only captures an individual's patrilineage. However, given the documented high level of ethnic segregation in the US during the 1940s (Eriksson and Ward, 2018) and high intra-ethnic marriage rates during this period, this is unlikely to be a first order concern. (Note that the majority of the CEOs in my sample were born in the 1940s or 1950s.) Separately, Pan et al. (2018) employ similar approach to identify CEOs' ethnic origins and find that the uncertainty avoidance indices constructed from CEOs' last names and from their mothers' maiden names are highly correlated, which further supports the mentioned observations. Finally, as 98% of my CEO sample are male, name changing due to marriage is not a concern.

[48]See, for example, surveys by Algan and Cahuc (2013, 2014) and Fernández (2011).

[49]That the correlation is not perfect possibly reflects the fact that immigrants to the US (i) have been non-randomly selected from the original population, and (ii) have somewhat assimilated to the host culture.

[50]Falk et al. (2016) find that the correlation between trust measures of the same individual elicited from trust games conducted one week apart is 0.6, suggesting a considerable amount of measurement

component of trust but also helps smooth out (ii) these measurement errors. In appendix A.3.1, I develop a framework to assess the relative sizes of (i) and (ii) using parameters from the literature (e.g., Glaeser et al., 2000). The results suggest that the baseline inherited trust measure is better than an individual-level survey-based trust measure and about 80% as precise as an individual-level game-based measure.[51] Furthermore, using the inherited trust measure does not introduce attenuation bias as in the case of classical measurement errors but produces unbiased estimates of the true effect (details in appendix A.3.2).[52]

Finally, as remarked in subsection 1.2.1, a CEO's trust attitude likely affects her firm's R&D outcome via its influence on firm's policies and consequently researchers' choices. In real life settings, CEOs' influence on firm's policies takes time to materialize and credible commitment to such policies is unlikely. As a result, researchers' perception of their CEO's trust are as important to their choices as the CEO's actual trust attitude. In large firms, this perception are most likely based on the collective reputation of the group to which the CEO belongs (e.g., Tirole, 1996; Macchiavello, 2010; Xu, 2015), most notably her ethnic group as it is a salient feature of her identity. The inherited trust measure precisely captures the CEO's ethnic group's collective reputation of trust attitude and therefore is the key explanatory variable on its own under this interpretation of the mechanism.

Measuring CEO's bilateral trust. Similar to her inherited generalized trust measure, CEO d's bilateral trust towards individuals from country c is calculated as

$$bitrust_{dc} = \sum_e w_e \times ethbitrust_{ec} \tag{1.4}$$

where $ethbitrust_{ec}$ is a measure for how much a person from country of origin e trusts a person from country of origin c. This country-level bilateral trust measure ($ethbitrust_{ec}$) comes from the Eurobarometer, a series of surveys conducted for the European Commission in which individuals in each country are asked the following question *"I would like to ask you a question about how much trust you have in people from various countries. For each, please tell me whether you have a lot of trust, some trust, not very much trust, or no trust at all."*,[53] The relevant Eurobarometer surveys cover respondents from 16 EU countries and ask about their trust attitude towards 28 countries, including a

errors. Results from other studies on the stability of experimental and survey measures of preferences are consistent with this finding (see survey by Chuang and Schechter, 2015).

[51] Of course, if one can administer *many* trust surveys or games on the same individual, one can average out much more precisely individual trust. However, this is highly infeasible.

[52] In essence, using the inherited trust measure is similar to using the cell-average of the right hand side variable as a new regressor, a very helpful procedure when one only observe cell averages (see Angrist and Pischke, 2009, c. 2).

[53] Following the literature, I recode the answers to 1 (no trust at all), 2 (not very much trust), 3 (some trust), and 4 (a lot of trust) before averaging them by country pair to derive $ethbitrust_{ec}$.

number of non-EU countries such as Russia, Japan, and China.[54] Existing studies using the same measure have shown that bilateral trust matters to a wide range of economic activities, from cross-country trade (Guiso et al., 2009) to venture capital investment (Bottazzi et al., 2016) to within-firm internal organization (Bloom et al., 2012). The CEO's bilateral trust measure $bitrust_{dc}$ is available for CEOs whose ethnic origins are among the 16 surveyed countries, which comprise 45% of the CEO name-matched sample (details in appendix A.2.4).

1.3.3 Baseline sample

To construct the baseline sample, I combine patent data from PATSTAT and CEO data from BoardEx with US public firms' performance data from Compustat, excluding firms in the financial sector and those whose headquarters are outside of the US. For practical purpose, I only consider firms with at least one name-matched CEO[55] and further restrict the sample to firms and CEOs for which all key variables are non missing. This results in a final baseline sample of 3,598 US public firms and corresponding 5,753 name-matched CEOs, with 29,384 firm by year by CEO observations during the period between 2000 and 2011 (see Table A.3). About 60% of these firms are R&D performing firms and patenting firms, sharing among them 700,000 patent applications between 2001 and 2012. Separately, about two thirds of the firms have more than one CEOs during this 12-year period, with an average of 1.7 name-matched CEOs each firm. 98% of the CEOs are male, each CEO has an average tenure of 7 years, and very few CEOs are the chief executive of more than one Compustat firm.

1.4 Within-firm effect of CEO's generalized trust

1.4.1 Within-firm empirical strategy

I first consider a difference-in-differences specification with firm fixed effects:

$$\text{asinh}(pat_{fd,t+1}) = \beta_1 trust_{fdt} + \mathbf{X}_{ft} + \mathbf{Z}_{dt} + \theta_f + \omega_t + \varepsilon_{fdt}. \tag{1.5}$$

Each observation represents a firm f in a year t with its current CEO d. $pat_{fd,t+1}$ is firm f's forward patent application counts in year $t+1$.[56] As patent distribution is skewed, I use the inverse hyperbolic sine transformation $\text{asinh}(pat_{fd,t+1})$ as the main outcome

[54]Unlike the GSS, Eurobarometer surveys are conducted among residences of European countries. However, given the discussed evidence of intergenerational transmission of trust attitude, it seems reasonable to use the Eurobarometer-based bilateral trust measure as a proxy for the bilateral trust among descendants of US immigrants.

[55]These firms comprise 92% of the firm sample and are mechanically larger than the remaining 8%. However, as my key estimates are within-firm estimates, any firm-level selection only poses threat to the results' external validity. This is unlikely to be a first order concern given that selected firms cover a large share of the firm sample.

[56]Results are robust to using further-forward patent application counts in year $t+2$ or $t+3$.

variable instead of raw patent counts (following Card and DellaVigna, 2017).[57, 58] The main explanatory variable $trust_{fdt}$ is the time-invariant measure of CEO d's inherited trust attitude, which also corresponds to researchers' perception of CEO d's trust attitude (details in subsection 1.3.2). To facilitate interpretation, $trust_{fdt}$ is standardized by its standard deviation at ethnic level.[59] The specification includes a full set of firm fixed effects θ_f, which helps control for all firm-level time-invariant characteristics that are correlated with either firms' innovation capability or selection of CEO. In addition, equation 1.5 also includes controls for firm's time-variant characteristics $\mathbf{X}_{ft}$ (i.e., firm's age, log(assets), log(sale)), CEO's time-variant characteristics $\mathbf{Z}_{dt}$ (i.e., CEO's age, gender, education dummies, tenure in firm), and a set of year fixed effects ω_t that accounts for macro-level cyclicality in innovation. Standard errors are clustered by CEO's main ethnic origin in case there are idiosyncratic factors that are specific to an ethnicity.[60] Alternative specifications that (i) further include controls for employment and R&D stocks or flows, (ii) employ additional industry-by-year fixed effects, or (iii) apply two-way clustering by CEO's main ethnic origin and firm all yield quantitatively similar results.

The coefficient of interest β_1 estimates the effect of CEO's trust on firm's patents. With the inclusion of firm and year fixed effects, equation 1.5 identifies β_1 from changes in CEOs and subsequent changes in patenting within the same firm over time. The difference-in-differences identifying assumption requires that the trend in potential outcomes be mean-independent from changes in CEO's trust, conditional on covariates. Under this identifying assumption of common trends, β_1 can be interpreted as the causal effect of CEO d on firm f's patents.[61] That is, the effect captured by β_1 is unlikely to be the result of reverse causality or confounded by firm f's time-variant unobservable characteristics that affect both the firm's choice of CEO and its innovation outputs (e.g., changes in firm's strategy driven by the board).

To formally test for common trend, I regress the change in CEO's trust in each CEO transition event on firm's patent application counts in different years before the transition, controlling for pre-change firm's and CEO's characteristics. The resulting coefficients are all small and not statistically different from zero, indicating that there is no association between firm's pre-change patenting and subsequent change in CEO's

[57]The inverse hyperbolic sine transformation $\operatorname{asinh}(x) = \ln(x + \sqrt{1+x^2})$ takes value 0 at $x = 0$ and approximates $\ln x + \ln 2 + O(\frac{1}{x \ln x})$ for large x. It has been promoted as a substitute for $\ln(x+1)$ by David Card, because one can still interpret changes in $\operatorname{asinh}(x)$ as close approximates of percentage changes in x for sufficiently large x thanks to its similarity with $\ln x$, while the function's behavior around $x = 0$ approximates $\ln(1+x) + O(x^2)$.

[58]Results are robust to (i) using $\log(1 + pat_{f,t+1})$, winsorized, or raw $pat_{f,t+1}$ as the outcome variable, (ii) estimating a semi-log Poisson count model with $pat_{f,t+1}$ as the outcome variable, instead of OLS.

[59]Inherited trust's standard deviation at ethnic level is 0.11. This equals the difference between Greek and English inherited trust levels.

[60]CEO d's main ethnic origin is $e^*_d = \operatorname{argmax}_e(w_{de})$. Average weight of the main ethnic origin (i.e., average w_{de^*}) among CEOs is 71%.

[61]In a more general setting with $\operatorname{asinh}(pat) = f(trust, \mathbf{X}, \mathbf{Z}, \eta)$, β_1 estimates an average causal effect $\mathbb{E}[\frac{\partial f(trust, \mathbf{X}, \mathbf{Z}, \eta)}{\partial trust}]$.

trust (Figure 1.2).[62] In addition, Figure 1.3 plots the average patent application counts by the number of years before the transition for the full sample of firms.[63] The flat pre-trend in patents suggests that the timing of the CEO transition is not driven by a trend in patenting, which implies that reverse causation is unlikely to be a concern in this setting.

1.4.2 Baseline effect of CEO's trust on firm's patents

Figure 1.4 presents the paper's key empirical finding visually with an event-study plot of firms' patent application counts by year with respect to CEO transition year (i.e., year 0). The solid blue line groups together all CEO transitions in which the new CEOs are *more* trusting than their predecessors (i.e., trust-increasing transitions), and the dotted red line corresponds to those in which the new CEOs are *less* trusting (i.e, trust-decreasing transitions).[64] The two lines exhibit similar pre-trends in the years before CEO transitions, but diverge visibly post-CEO change.[65] Firms that experience an increase in CEO's trust after the transition also experience increases in patenting in post-transition years (i.e., the upward-sloping solid line) and vice versa (i.e., the downward-sloping dotted line). In addition, Table A.5 shows that while the difference in average pre-transition patents between these trust-increasing and trust-decreasing CEO transitions is small and not statistically different from zero (by matching), the difference in their post-transition patents is large, positive, and statistically significant at 5% level. Figure 1.4 suggests that CEO's trust does have a considerable effect on firm's innovation.

Table 1.1 then estimates equation 1.5, which exploits changes in CEOs and subsequent changes in patenting within the same firm over time, using the full baseline sample described in subsection 1.3.3. Given evidence of common trend (see subsection 1.4.1), the coefficient on CEO's trust captures its effect on firm's forward patent count. I first report two basic specifications without any firm or CEO controls (column 1) or without firm controls that could also be outcomes of CEO's trust, such as assets and

[62]Figure 1.2 plots the coefficients $\hat{\gamma}_k$ for $k \in [-6,-1]$ from estimating: $\Delta trust_{fdt} = \sum_{k=-7}^{-1} \gamma_k(\text{asinh}(pat_{fdt}) \times event_{t-k}) + \beta trust_{fdt} + \mathbf{X}_{ft} + \mathbf{Z}_{dt} + \omega_t + \varepsilon_{fdt}$, in which (i) $\Delta trust_{fdt}$ is the difference between CEO d's and her successor's trust measures, (ii) $event_{t-k}$ is an indicator equal to 1 if the transition happens in year $t-k$, and (iii) $\mathbf{X}_{ft}$ additionally includes a full set of firm's 3-digit industry dummies.

[63]Figure 1.3 plots the coefficients $\hat{\gamma}_k$ for $k \in [-7,-2]$ relative to $\hat{\gamma}_{-1}$ from estimating: $\text{asinh}(pat_{fdt}) = \sum_{k=-7}^{-1} \gamma_k event_{t-k} + \mathbf{X}_{ft} + \mathbf{Z}_{dt} + \theta_f + \omega_t + \varepsilon_{fdt}$, in which $event_{t-k}$ is an indicator equal to 1 if the next CEO transition happens in year $t-k$.

[64]To plot Figure 1.4, I first (i) partial out the covariates by regressing patent application counts on firm's and CEO's controls with firm's industry and year fixed effects, then (ii) average the residuals by year separately for each group of CEO transitions, and finally (iii) normalize these annual averages to their respective group's pre-transition mean. I restrict the sample to CEO transitions in which both predecessor's and successor's tenures are at least 5 years, so that the plotted patent trends are not driven by changes in firm composition. Furthermore, to address possible mean reversion, each trust-increasing transition is matched to a trust-decreasing transition based on their average pre-transition residual patent counts.

[65]Similar event-study plot using all CEO transitions that meet the CEO tenure restriction (Figure A.1) also exhibits the same pattern. This provides further evidence in support of the common trend identification assumption discussed in subsection 1.4.1, as the patent pre-trends in this figure are not guaranteed to coincide by construction.

sale (column 2). Column 3 then presents the baseline specification that includes the full set of controls for firm's age, size and CEO's age, gender, education, and tenure in addition to firm and year fixed effects. Finally, column 4 further adds industry-by-year fixed effects to account for industry-level patenting cyclicality. The resulting CEO's trust estimates are almost identical across these four columns and imply that one standard deviation increase in CEO's trust is associated with 6.3% increase in firm's patent filing, statistically significant at 1% level.[66] This equals 1.1 additional patents annually for the average baseline sample firm with value equal to $3 million in additional R&D,[67] suggesting that CEO's trust also has a substantial impact from an economic perspective.

Figure A.2 plots the CEO's trust estimates as a function of the change in CEO's trust (i.e., difference between trust measures of new and old CEOs) and shows that the effect is driven by both positive and negative changes in CEO's trust, similar to the pattern shown in event study Figure 1.4.[68] Additionally, Table A.6 describes and reports results from further robustness checks, some of which are also mentioned in subsection 1.4.1.

Columns 5 to 7 turn to alternative measures of trust. Column 5 further refines the trust measure with a machine-learning procedure using all variables commonly available in both BoardEx and GSS besides ethnic origin, including age, gender, education, and birth cohort, which demographic characteristics have been shown to predict individual trust attitude.[69] This measure yields a slightly larger CEO's trust coefficient, suggesting that the baseline trust measure captures most of the meaningful variations in individual trust across observable demographic characteristics. Column 6's trust measure uses the trust answers of all GSS respondents, not just those in highly prestigious occupations, and column 7's uses the World Value Survey's (WVS) trust answers collected from each ethnicity's home country, instead of the GSS's. As the baseline trust measure is closer to the US CEO population, one would expect smaller effect using the full-GSS-based measure, and even smaller effect using the WVS-based measure, as shown in columns 6 and 7.[70]

Table A.7 focuses on a special subsample of CEO transitions for which the common trend condition is better warranted: transitions following CEO retirements or deaths (e.g., Fee et al., 2013; Bennedsen et al., 2010). As the need to replace the existing CEO arises exogenously, the timing of the subsequent transition is likely exogenous to

[66]The inclusion of controls does not affect the magnitude of CEO's trust estimates but helps improve their precision.

[67]Dechezleprêtre et al. (2018) estimate that a patent costs 1.8 million in 2007 British pounds.

[68]The graph represents the effect of trust on patent counts as a function of change in CEO's trust, namely $\frac{\partial \text{asinh}(pat_{fd,t+1})}{\partial trust_{fdt}}(\Delta trust_E)$. Each point estimate is obtained from the benchmark regression, weighted by a kernel function around that value of $\Delta trust_E$ (see Do et al., 2017's appendix for details of this method).

[69]I first fit a LASSO model (Tibshirani, 1996) to predict trust attitude from these demographic characteristics and their interactions with ethnic origin, using individual-level data from the GSS. I then use the LASSO-selected model to predict CEOs' trust attitude. It is worth noting that all ethnic origin dummies are retained in the selected model.

[70]The fact that the magnitude and precision of the CEO's trust coefficient increase with the quality of the CEO's inherited trust measure is reassuring, as it is difficult to specify an omitted variable that is always more precisely measured when trust is better measured.

firm's other decisions. Even though the replacement CEO choice is endogenous, under the assumption that firm's underlying characteristics do not change in the event of an exogenous transition, firm fixed effects sufficiently control for firm's new CEO selection. I focus on CEO natural retirements around the age of 65 in columns 1 to 3.[71] Column 6 further includes CEO deaths within one year of leaving the office; however, there are only very few such events in my data. The CEO's trust estimates are large, positive, and statistically significant across these subsamples.[72] Although these results should be taken with caution (they are estimated from small subsamples of special events), overall, they suggest that CEO's trust does have impact on firm's innovation. This resonates with findings from existing literature that CEOs matter to firm performance (see survey by Bertrand, 2009).

However, as the same literature has shown that various CEO's and firm's decisions are influenced by CEO's other characteristics, one would be concerned that some may correlate with CEO's trust and directly affect innovation at the same time. I take two approaches to address this concern: first, I control for potential confounding factors related to CEO's origins, and second, I exploit within-CEO variation to control for all CEO observable and unobservable characteristics with CEO fixed effects in the next section (subsection 1.5). Even though the latter provides better identification of the effect of CEO's trust, the former sheds some light on other factors by CEO's origins that may also have an effect on firm's innovation.

Given firms' inclination to trade with, have business in, or hire from their CEOs' home countries and the possible spillovers from these linkages, Panel A of Table 1.2 controls for a range of macroeconomic variables that measure the CEO's home countries' level of development and technological capabilities. These controls include country-by-year-level (i) GDP, population and GDP growth (column 1), (ii) high school graduation rate (column 2), (iii) governance quality index, (iv) total trade volume with the US (column 4), and (v) total patent applications (column 5), all of which have been shown to be related to country-level trust measure (see surveys by Algan and Cahuc, 2013, 2014).[73] Among these factors, only GDP growth and trade volume with the US seem to have a relationship with firm's patenting (column 6). More importantly, the magnitude and statistical significance of the CEO's trust effect is not affected by the inclusion of these controls across Panel A.

In Panel B, I turn to examine if the observed effect is driven by other cultural traits instead of trust attitude. First, a CEO's ethnic groups' socioeconomic characteristics could impact her skill accumulation, both directly via investments in human capital and

[71] 65 is the official Social Security retirement age and the traditional retirement age used in the related literature (e.g, Fee et al., 2013). In the data, I also observe a spike in CEO's leaving executive positions for good around 65.

[72] The estimates' magnitude suggests that CEO's trust is more important to innovation in times of greater uncertainty.

[73] For each home-country variable h, the control variable h_{fdt} is calculated as $h_{fdt} = \sum_e w_{de} \times h_{et}$ where h_{et} is the value of h in the home country e in year t and w_{de} is as described in subsection 1.3.2.

indirectly via exposure (e.g., Bell et al., 2018). As there are strong correlations among self-perceived class, occupational prestige, earnings, and education, I use an ethnicity's share of college graduates as a summary statistics for its socioeconomic characteristics in column 1.[74] The related literatures on culture and on CEOs have also pointed to some salient cultural values that have economic significance at both macro and individual levels. These include: (i) Protestant work ethic, which has been discussed since Weber (1905) and shown to influence individual's choices of incentive contract and total work hours (e.g., Liu, 2013; Spenkuch, 2017), and (ii) risk preference, which affects national saving behaviors at the macro level and firm's financing decisions at the micro level (e.g., Pan et al., 2017). In column 2, I measure the Protestant work ethic, which promotes the intrinsic value of work, using answers to the GSS question on the relative importance of work versus luck as the means to get ahead.[75] In column 3, risk preference is inferred from the shares of GSS respondents having stock market or mutual fund investments.[76] Finally, column 4 pools together all the cultural trait controls and column 5 adds the home country controls that are statistically significant in Panel A.[77] The inclusions of these variables does not affect the magnitude and statistical significance of the coefficient on CEO's trust very much, implying that this effect is unlikely to be confounded by other factors related to the CEO's ethnic origins.

1.5 Within-CEO effect of CEO's bilateral trust

1.5.1 Within-CEO empirical strategy

Despite the evidence discussed so far, there remain many other CEO personal characteristics that one cannot observe, measure, or directly control for, such as ability, management style, or preference for innovation. These characteristics can have direct effects on firm's innovation and be correlated with her trust attitude at the same time. As a result, equation 1.5's β_1 captures not only the effect of CEO's trust but also the effects of those other characteristics. To address this concern, I exploit within-CEO variation in bilateral trust towards different groups of inventors and corresponding variation in patenting among those different inventor groups. Such within-CEO variation allows me

[74] These variables are derived from corresponding GSS questions similarly to the trust measure (see subsection 1.3.2).

[75] The question reads *"Some people say that people get ahead by their own hard work; others say that lucky breaks or help from other people are more important. Which do you think is most important?"*. Relatedly, individuals' answer to the same question has been shown to be correlated with their preference for redistribution (Alesina and Angeletos, 2005; Giuliano and Spilimbergo, 2013).

[76] Campbell (2006) shows that risk averse individuals participate less in the stock market.

[77] Appendix A.8 explores alternative measures for socioeconomic status, work ethic, and risk preference derived from the GSS. Appendix A.9 presents additional tests using trust and other cultural trait measures derived from the Global Preference Survey (Falk et al., 2018), including risk preference, time preference, positive reciprocity, negative reciprocity, and altruism. All results are quantitatively similar to those reported in Table 1.2.

to include a full set of CEO fixed effects in the following equation:

$$\text{asinh}(pat_{fdc,t+1}) = \beta_2 bitrust_{fdct} + \theta_{ft} + \kappa_c + \omega_d + \varepsilon_{fdct}. \quad (1.6)$$

Each observation is a combination of firm f by year t by its CEO d by country c. The outcome variable $pat_{fdc,t+1}$ is firm f's total patent counts by inventors from country c in year $t+1$ (details below). The main explanatory variable $bitrust_{fdct}$ measures how much CEO d trusts individuals from country c, which aligns with researchers from country c's perception of CEO d's trust attitude towards them (details in subsection 1.3.2). Besides the crucial CEO fixed effects (ω_d), equation 1.6 also controls for inventor country's baseline characteristics (e.g., development level, institution quality, technological comparative advantage, inventor pool quality) with inventor country fixed effects (κ_c), and for firms' time-variant characteristics with firm by year fixed effects (θ_{ft}). Given these stringent sets of fixed effects, any remaining variations would have to be at the firm-inventor country pair or CEO-inventor country pair levels. To control for potential confounding factors by those variations, I additionally include firm-by-inventor country fixed effects or an array of CEO-inventor country pairwise controls in the robustness checks. The coefficient β_2 then captures the effect of CEO's bilateral trust toward individuals in a country on corresponding patent counts by inventors from that country. Standard errors are clustered by CEO's main ethnicity-inventor country pair.[78]

To construct the outcome variable pat_{fct}, I use information on patent inventors' addresses or last names to allocate patents to different inventor countries or origins (similar to Foley and Kerr, 2013, details in appendix A.2.2). In the address-based approach, a patent is assigned to the country where its inventors reside, which is also likely to be where the invention is created.[79] About 30% of my patent sample are by non-US R&D labs of US-based multinational firms, most of these labs are in Europe, Japan, and China. For the remaining US-based patents, I infer their inventors' ethnic origins from the inventors' last names (see subsection 1.3.2), then assign the corresponding patents to their inventors' countries of origin accordingly (I refer to this as the last name-based approach).[80] The variable pat_{fct} is the sum of all patents filed by firm f in year t that are allocated to country c.

The sample used to estimate equation 1.6 includes all firm f-inventor country c pairs such that firm f has patents by inventors from country c in at least one year during the study period.[81] If a firm-inventor country pair satisfies this condition, then it is included in the sample even in the years when the pair has zero patent to avoid biases arising

[78]Results are robust to two-way clustering by CEO's main ethnicity-inventor country pair and firm.

[79]Compared to patent office's location, inventors' location is a better proxy for where the invention is created, as an invention can be filed for protection in many different jurisdictions. In the few cases in which a patent has multiple inventors living in different countries, I allocate a proportional fraction of the patent to each of those countries.

[80]Similar to CEOs, over 80% of all inventors are mapped to their ethnic origins based on their last names.

[81]That is, $pat_{fct} > 0$ for some $t \in [2000, 2012]$. In addition, the corresponding CEO d's ethnic origins and country c are among the countries surveyed and/or covered by the Eurobarometer, so that $bitrust_{dc}$ is non-missing.

from selection into patenting over time.[82] Therefore, β_2 captures both the intensive and the extensive margins of CEO's bilateral trust effect on inventors' patenting.[83] Using the address-based approach to identify inventors' countries results in a sample of 3,481 firm by country dyads, covering 730 firms with R&D labs in 27 countries (outside of the US) and 960 CEOs. Additionally employing the last name-based approach gives a larger sample of 8,554 firm by country dyads, covering 1,263 firms with inventors from 27 countries and 1,654 CEOs (see Table A.4). Figure 1.5 shows the distribution of the CEO's bilateral trust measure in these subsamples.

1.5.2 Effect of CEO's bilateral trust on inventors' patents

Table 1.3 reports the effect of CEO's bilateral trust towards individuals in a country on corresponding patent counts by inventors from that country, estimated using equation 1.6. Columns 1 to 3 consider the bilateral trust sample constructed from only non-US-based inventors and columns 4 to 6 use both non-US- and US-based inventors.[84] The baseline bilateral-trust specification (equation 1.6) is reported in columns 1 and 4. In addition, I fully interact the sets of firm, CEO, and inventor country dummies with year dummies in columns 2 and 5. Columns 3 and 6 further include firm-by-inventor country fixed effects to control for specific characteristics of each R&D lab or group within in the firm that are not already captured by firm or inventor country fixed effects.[85, 86] The CEO's trust coefficient in column 1, which estimates the effect of CEO's bilateral trust on patents after controlling for firm's time-variant and CEO's characteristics, implies that one standard deviation increase in CEO's bilateral trust towards a country is associated with 5% increase in patents by the R&D lab in that country (statistically significant at 5% level). This effect is similar in magnitude to the baseline CEO's trust effect of 6% reported in Table 1.1, and is robust to adding even more stringent fixed effects in columns 2 and 3.

Columns 4 to 6 exhibit similar pattern across the different specifications. One would expect that the effect of CEO's bilateral trust is smaller in this subsample (3% compared to 5% in the other subsample) for a couple of reasons. First, as the differences among US-based inventors from different home countries are less salient, CEOs' bilateral trust

[82] Alternatively, one could estimate equation 1.6 using the full sample of all firm-inventor country pairs. However, the inclusion of never-patenting firm-inventor country pairs adds considerable computational burden while offering no additional meaningful within-firm or within-CEO variation in patenting.

[83] I discuss approaches to separate these two effect margins in subsection 1.5.2.

[84] Columns 1 and 2 of Table A.11 show that the within-in firm effects of CEO's generalized trust on firm's patents are also positive and statistically significant among these subsamples.

[85] However, it is difficult to specify what the potential confounding factors at firm-inventor country level are. An example is that firm f has a large group of inventors in or from country c for firm-country specific reasons that are not already explained by firm-level and country-level characteristics, and firm f is inclined to select CEOs with high bilateral trust towards country c for the same reasons.

[86] This specification mirrors equation 1.5's within-firm specification, but is at the within-R&D lab/group level instead. That is, it exploits the change in patenting by the same R&D lab or group following a change in CEO, relative to that of other R&D labs or groups in the same firm under the same CEO. Figure A.3 reports evidence of the common trend identification condition that pre-change patents at R&D lab or group level do not predict the change in CEO's bilateral trust towards the corresponding R&D lab or group.

towards these inventors and these inventors' perception of the CEOs' trust towards them are less heterogenous. Second, from an organizational perspective, CEOs would be more likely to implement differentiating policies towards R&D labs in different countries than towards different groups of US-based. On the other hand, the combined subsample covers a much larger share of the patent pool, which result in a larger estimation sample and more precisely-estimated coefficients.[87] Together, results from both subsamples are complement and they both imply that CEO's trust positively impacts firm's innovation.

Furthermore, to understand whether this impact comes from existing inventors in the firm (i.e., the intensive margin), or from new R&D labs or groups that arrive with the CEO (i.e., the extensive margin), I focus on smaller subsamples of firm f-inventor country c pairs such that firm f has patents by inventors from country c before the corresponding CEO assumes position (Table A.11, columns 5 and 6). Even though the CEO's trust coefficients are less precisely estimated in these subsamples, they are of similar magnitude to the combined bilateral trust effects. This implies that CEO's trust does work through the intensive margin by improving the innovation outputs of existing researchers, as suggested by the model in Section 1.2.

Table 1.4 controls for potential confounding factors at CEO-inventor country pair level. An immediate concern is that CEOs may differentially favor inventors in their home countries or from the same ethnic groups (e.g., Do et al., 2017). To address this, I exclude all CEO-inventor country pairs such that the inventor country is the same as the CEO's main home country (columns 1 and 5), and control for the geographical distance between CEO's and inventor's home countries to account for potential "favoritism spillovers" (columns 2 and 6). Next, bilateral trust is correlated with cultural proximity, which could have a direct impact on R&D outputs thanks to better information flows between CEOs and researchers (e.g., better screening of researchers, better working relationship between researchers and CEOs). In columns 3, 4, 7 and 8, I include CEO's-inventors' home countries pairwise linguistic and genetic distances as proxies for their ease of interaction and cultural proximity (Spolaore and Wacziarg, 2016). The CEO's bilateral trust coefficients remain statistically significant across all of these robustness checks in both bilateral trust subsamples. More importantly, they are similar to Table 1.3's estimates in magnitude, suggesting that the reported CEO's bilateral trust effect is not spuriously driven by favoritism or other confounding factors.

Furthermore, as trust and trustworthiness are correlated, one may concern that the CEO's trust effect instead captures the impact of her trustworthiness. I exploit the differences between (i) the baseline $bitrust_{dc}$ that measures CEO d's bilateral trust towards inventors from country c, and (ii) a new variable $invbitrust_{cd}$ that measures

[87]Consistent with the view that the bilateral trust effect is expectedly weaker towards US-based inventors, Table A.10, which estimates equation 1.6 using an alternative bilateral trust sample constructed from only US-based inventors, yields a smaller CEO's bilateral trust coefficient of 2%, statistically significant at 10% level.

inventors from country c's trust towards CEO's d,[88] and find that in a specification based on equation 1.6 that includes both directions of bilateral trust as explanatory variables, the coefficient on $bitrust_{dc}$ is large and statistically significant while the coefficient on $invbitrust_{cd}$ is zero (Table A.12, column 3). Given the variables' high correlation, I further employ specifications in which I partition the sample into deciles of $invbitrust_{cd}$ ($bitrust_{dc}$) and compare observations within the same decile by introducing a full set of decile dummies. That is, I estimate the effect of $bitrust_{dc}$ ($invbitrust_{cd}$) among observations with very similar $invbitrust_{cd}$ ($bitrust_{dc}$). The resulting coefficients show that $bitrust_{dc}$ is still strongly associated with patent outcome (column 4) while the same is not true for $invbitrust_{cd}$ (column 5). The evidence suggests that it is the CEO's trust toward the inventors that is the main driver of the CEO's trust effect.

1.6 Evidence of mechanism

1.6.1 Framework for separating different mechanisms

Alternative mechanisms. Section 1.2 presents a model in which a CEO's trust in a researcher's type improves the latter's incentives to undertake high-risk explorations, which results in more innovation. Yet there exist other competing mechanisms that can also explain this relationship between CEO's trust and firm's innovation. First, trust could lead to greater delegation by the CEO, which induces more effort from the researcher and therefore improves R&D outcomes (Aghion and Tirole, 1997; Acemoglu et al., 2007; Bloom et al., 2012).[89] Second, trust, as catalyst for cooperation (Putnam et al., 1993; Fukuyama, 1995), could also have an essential role in sustaining informal relational contracts (Baker et al., 1999, 2002). That is, when the CEO cannot credibly commit to her policies, the researcher is more likely to cooperate and exert effort if he trusts that the CEO will honor the promised rewards for success.[90, 91] In both of these frameworks, greater effort by the researcher improves the expected outcome of all R&D projects and therefore increases all types of innovation. On the other hand, greater risk taking likely produces more high-quality patents but not necessarily more low-quality

[88] $invbitrust_{cd} = \sum_e w_{de} \times ethbitrust_{ce}$ where $ethbitrust_{ce}$ is the bilateral trust measure for how much a person from country of origin c trusts a person from country of origin e. Note that $invbitrust_{cd}$ measures inventors from country c's perception of CEO d's trustworthiness, not their perception of CEO d's trust towards them. The latter, as discussed in subsection 1.3.2, aligns more closely with $bitrust_{dc}$ instead.

[89] Aghion and Tirole (1997) and Acemoglu et al. (2007) model higher trust as greater preference congruence between the principal and the agent, which leads to greater delegation by the principal. Separately, Bloom et al. (2012) consider trust as the principal's belief in the agent to behave in the "correct" way, and find that trust is empirically associated with both greater decentralization and better firm's performance.

[90] Therefore, it is the researcher's trust towards the CEO, or relatedly the CEO's trustworthiness, that matters for cooperation and the game's outcomes. This is inconsistent with the evidence presented in subsection 1.5.2 that CEO's trust towards the researcher, not the other direction of trust, is the main driver of CEO's trust effect on innovation.

[91] Additionally, Aghion et al. (2013) show both theoretically and empirically that greater monitoring enables more innovation, also through reducing career risks and allowing more risk taking. However, as trust reduces monitoring incentive, they are more likely substitutes than competing explanation of the other's effect on innovation.

ones. This suggests that it is possible to distinguish between the different mechanisms by examining the patent quality distribution, as detailed below.

Estimation of mechanism. I develop a formal framework to compare section 1.2's risk-taking mechanism that CEO's trust increases innovation by encouraging able researchers to choose risky projects (i.e., explore instead of exploit), against alternative mechanisms of trust inducing more effort by those researchers. The framework identifies the different mechanisms by exploiting their potentially different predictions regarding the outcome *quality* distribution of R&D projects, using patent citation counts and the likes as measures of quality.

First, I establish that the effect of CEO's trust does not work through increasing the scale of R&D, but rather the choice of projects and their eventual outcomes. This is suggested by the regression of (future) R&D expenditure on CEO's trust specified as in equation 1.5, which consistently yields statistically insignificant estimates close to zero (Table A.15). That is, the number of independent R&D projects (i.e., N) the firm runs is not influenced by CEO's trust.

Let us assume that a project's outcome quality x, measured by a patent's forward citation counts, follow a distribution $F_T(\cdot)$ indexed by the CEO's trust T. Note that x is observable only when the project is patented, that is, when $x \geq 0$. It further assume that better quality patents are always rarer (i.e., $F'_T(x)$ is decreasing on $[0, \infty)$ $\forall T$).[92]

I parameterize this family of distributions as $F_T(x) = F_0(\frac{x-\bar{x}-b(T)}{a(T)} + \bar{x}) \overset{def}{\equiv} F(\frac{x-\tilde{b}(T)}{a(T)})$, in which $a(T)$ represents the change in project's outcome quality variance and $b(T)$ the shift in project's outcome quality mean induced by CEO's trust.[93] The risk-taking mechanism suggests that CEO's trust increases patented innovations through $a(T)$ (i.e., $a'(T) \geq 0$), while alternative mechanisms work through $b(T)$ (i.e., $b'(T) \geq 0$). As higher T implies higher number of observed patent counts $N(1 - F_T(0)) = N(1 - F(\frac{-\tilde{b}(T)}{a(T)}))$ under both types of mechanisms, it is not possible to distinguish between the two by just examining the effect of CEO's trust on total patent count.

The solution to this problem comes from considering patents within a specific low quality range $[c_1, c_2]$. Given the assumptions regarding $F_T(.)$, it follows that:

Proposition 3. *Higher $b(T)$ increases the count of patents within the low quality range $[c_1, c_2] \subset [0, \infty)$.*

That is, alternative mechanisms that work through $b(T)$ (i.e., mean shifting) increase not only the total patent counts but also the number of patents within any arbitrary patent quality range (see appendix A.4.1 for detailed proof). The same prediction does not hold for the baseline risk-taking mechanism that works through $a(T)$ instead. On the contrary, under certain mild conditions, it can be shown that higher $a(T)$ *decreases*

[92]This is consistent with the empirical patent quality distribution, as measured by forward citations.

[93]That is, $\tilde{b}(T) \overset{def}{\equiv} \bar{x} + b(T)$ and $F(.) \overset{def}{\equiv} F_0(.)$. Note that $a(0) = 1$, $b(0) = 0$, and $\bar{x} = \mathbb{E}_{F_0}(x)$.

the count of patents within the quality range $[c_1, c_2] \subset [0, c]$ for small enough c (see appendix A.4.2 for details).[94]

These results imply that it is possible to identify the two types of mechanisms by examining patent counts in low patent quality ranges. Specifically, when one consider patents in increasing brackets of quality, the effect of CEO's trust on the corresponding patent counts increases from negative/zero to positive under the risk-taking mechanism. In contrast, the alternative mechanisms predict similar effects of CEO's trust on patent counts in different quality brackets. This will serve as a simple test of the mechanism.[95]

1.6.2 Effect on patent quality distribution via exploration

I apply subsection 1.6.1's methodology to identify between the risk-taking mechanism, by which CEO's trust increases innovation through encouraging able researchers to choose risky instead of safe R&D projects (i.e., exploration instead of exploitation), versus suggested alternative mechanisms of delegation and/or cooperation. The method considers trust effects patents in low quality quantiles.

I follow the literature on innovation and patenting in measuring patents' quality by their forward citation counts.[96] As forward citations take time to accumulate and vary by technology field, I first compute each patent's citation decile with respect to the universe of patents in its same application technology field-by-year cohort, then sum up the number of patents in each quality decile at the firm by year level. The resulting variable pat^q_{ft} counts the number of patents in quality decile q filed by firm f in year t, for $q \in [1, 10]$.[97]

Figure 1.7 documents how CEO's trust effect vary by patent quality decile by plotting the coefficients estimated from equation 1.5 using $\text{asinh}(pat^1_{ft})$ to $\text{asinh}(pat^{10}_{ft})$ as the outcome variables. The upward-sloping pattern indicates that CEO's trust has larger positive effect on higher-quality patents. On the other hand, its effect on patents below-median in quality is not statistically different from zero. As discussed in subsection 1.6.1, these results are consistent with the exploration channel, and reject alternative

[94]Figure 1.6 illustrates these results by showing a baseline distribution (in dotted red line) with its mean-preserving spread counterpart (in solid blue line) in the top figure and its mean-shifting counterpart (in solid green line) in the bottom one. The solid vertical line at zero represents the patent quality threshold and the dashed vertical line corresponds to a quality threshold c. One can only observe patented projects in the half-plane to the right of the patent threshold. Proposition 3 implies that the area between $[0, c]$ is always higher under the higher-mean distribution (see bottom figure) compared to the baseline distribution, while the same is not necessarily true under the higher-risk distribution (see top figure).

[95]Furthermore, one can also identify separately $a(T)$ and $b(T)$ from considering the trust effect on different patent quality quantiles. It is thus possible to structurally estimate the effects of trust via the two types of mechanisms (e.g., by assuming $a(T) = aT$, $b(T) = bT$, and $F \sim \mathcal{N}(\bar{x}, \sigma)$). This is a topic that I will return to in future research.

[96]Hall et al. (2005) show that one more citation per patent (around the median) is associated with 3% higher in market value for the firm. Trajtenberg (1990), Harhoff et al. (1999), and Moser et al. (2015) also find that patent's forward citation counts is correlated with patent quality.

[97]The bottom three deciles contain mostly patents with zero forward citations.

channels such as delegation or cooperation.[98] Furthermore, using Table 1.3's bilateral-trust specification (equation 1.6) also gives similar result that CEO's trust effect is the largest on patents in the top quality quartile, while its effect on the bottom quartile is either small or zero (see Table A.13).[99]

Table 1.5 considers various other patent quality measures and estimates equation 1.5 using quality-weighted patent counts as the outcome variable (Squicciarini et al., 2013). Column 1 uses standard forward citation counts as the quality weights. Column 2 uses the number of backward citations to scientific literature. Columns 3 to 5 use patent scope, generality index, and originality index, which measure the range of technology fields covered by the patent, its forward citations, and its backward citations respectively.[100] Column 6 considers only granted patents. Across these different quality-weighted patent counts, the coefficients on CEO's trust are positive and statistically significant, with magnitudes larger than or similar to the baseline effect. The large effect on citations to scientific literature in column 2 is especially interesting, as it suggests that CEO's trust encourages researchers to explore directions that are closer to scientific frontiers and therefore could result in inventions of significantly higher quality (Cassiman et al., 2008; Branstetter, 2005). Furthermore, column 7 directly shows that CEO's trust improves not only the absolute forward citation counts, but also the average citation counts per patent, and this effect is of sizable magnitude (4.4%, statistically significant at 5% level).[101] The same results also hold in both bilateral trust subsamples, which similarly report largest effects on forward citations, citations to scientific literature, and patent scope (see Table A.14). Together, Tables 1.5 and A.14 provide further evidence that CEO's trust increases both the quantity and the quality of innovation, as is expected under more exploration.

1.6.3 Effect increases with researcher quality pool

Section 1.2's model predicts that CEO's trust always increases total innovation. However, as trust induces innovation through encouraging good researchers to explore, its effect is expected to be larger among firms with better researcher pool quality and vice versa. As data on the full sample of researchers in each firm are not available, I construct a proxy for research quality as the residuals from regressing patents on observable firm and CEO characteristics, controlling for industry and year fixed effects. That is, if there are two firms in the same industry and time space with similar observable characteristics

[98] This resonates with Azoulay et al.'s (2011) finding that scientists at Howard Hughes Medical Institute (HHMI) produce high-impact papers at a higher rate than their NIH-funded peers, as HHMI's policies are better at tolerating early failure and rewarding long-term success.

[99] As patents are already divided into smaller cells of firm by country by year in this specification, I only further classify them by quality into 4 quartiles instead of 10 deciles. Similar results hold in both bilateral trust subsamples.

[100] Trajtenberg et al. (1997) first proposed the generality and originality indices, arguing that a patent is likely more general purposed if it benefits different fields and more original if it relies on different knowledge sources.

[101] The dependent variable in column 7 is the inverses hyperbolic sine of average forward citation counts of patents filed by firm f in year $t+1$, which is set to zero if firm f files zero patents in year $t+1$.

(including R&D expenditure and CEO's trust), and one firm produces more patents than the other, then it is likely that the former has better researchers than the latter.

Table 1.6 interacts CEO's trust measure with firm-level proxy for researcher pool quality during the pre-transition period and finds that consistent with the model's implication, CEO's trust effect significantly increases with researcher quality. This finding is robust to averaging the quality proxy over different pre-transition windows (columns 1 to 3) and using different level of industry fixed effects in computing this proxy. Column 4 further shows that CEO's trust effect is sizable and statistically significant only among firms whose existing researcher pool quality is in the top two quintiles, consistent with the pattern plotted in Figure A.4. Furthermore, the same pattern holds for CEO's trust effect on firm's R&D efficiency, as measured by patent output over R&D expenditure (column 6), as well as firm's future performance, as measured by future sales, employment, and total factor productivity (TFP) (Table A.16). These results suggest that trust is effective to not only innovation but also real performance only when it is not grossly misplaced.

1.6.4 Evidence of effect on "corporate trust culture"

One concern is how a CEO's personal characteristics could affect a researcher's choices in large public firms, given the likely multiple layers between them.[102] On one hand, direct interactions between the CEO and the researcher are not necessary for the mechanism, as the former's trust impacts the latter's choices through her influence on policies, which are observed and even anticipated by the researcher (based on his perception of her trust attitude). On the other hand, in reality it is unlikely that the CEO makes direct decisions regarding a specific researcher, so one would expect her trust to also have an effect on the beliefs and choices of those in below levels in order for it to influence the choices of the researcher. In this subsection, I explore whether CEO's trust attitude is transmitted within the firm.

To measure "firm's trust attitude," I use Sull's (2018) dataset of employee sentiments covering over 500 US large public firms. This dataset is constructed from the text analysis of almost one million online employee reviews on Glassdoor.com, one of the largest career intelligence sites worldwide, between 2008 and 2017.[103] It covers a large set of topics related to corporate culture and contains the number of instances each topic appears in a review with positive or negative sentiment.[104] I am most interested in the

[102]Despite this concern, González-Uribe and Groen-Xu (2017) also find that longer CEO employment contract is associated with more patenting in a sample of US public firms, which result suggests that CEOs can influence innovation activities even firms where they are hierarchically distant from the inventors.

[103]Grennan (2014) uses similar approach of text analyzing online employee reviews to measure corporate culture along the 7 dimensions proposed by O'Reilly et al. (1991, 2014). She finds that changes in corporate governance lead to changes in corporate culture.

[104]These topics are also selected based on O'Reilly et al.'s (1991, 2014) 7 dimensions of corporate culture. For example, "corporate trust culture" is a topic under the integrity dimension. The sentiment count is positive if the topic is mentioned positively and vice versa, and is zero if the topic is not at all mentioned

topic that measures the extent to which employees trust one another, which for ease of exposition I will call "corporate trust culture."

I extend my firm and CEO data to 2016 and match them with review-level sentiment data by firm and year. I then aggregate each topic's sentiment measure by CEO term and standardize the resulting measures by their standard deviations.[105] This results in Table 1.7's sample of 393 observations at firm by CEO term level, covering 277 firms. To examine the relationship between CEO's trust attitude and corporate trust culture, I regress the described aggregated corporate-trust-culture sentiment measure on CEO's inherited trust, controlling for period-average firm's and CEO's characteristics and 3-digit industry fixed effects (column 2) or even firm fixed effects (column 3). The coefficients reports a strong association between CEO's trust and corporate trust culture, which is unlikely to be driven by firm-CEO matching as suggested by column (3). This relationship is robust to adding additional controls for CEO's approval rate by the same sets of reviewers (column 4) and CEO's other cultural traits as studied in Table 1.2 (columns 5 and 6). Even though these results should be taken with great caution,[106] they do provide suggestive evidence that CEO's trust attitude affects how those in below levels view and work with one another.[107]

1.7 Interpretation and discussions

1.7.1 Innovating or patenting?

Using patents to measure innovation raises the concern that a CEO's trust attitude may be correlated with her preference for patenting instead of having a true impact on innovation. The direction this correlation is ambiguous. On one hand, a more trusting CEO may count on her employees to keep trade secrets and therefore chooses to patent less. If so, the baseline estimates do not capture the full extent of CEO's trust effect on innovation. On the other hand, a more trusting CEO could have better confidence in the patent system and thus higher propensity to patent. To assess if the observed effect is driven by this, I construct measures of CEO's confidence in the government and in the scientific community as proxies for her confidence in the patent system.[108] Controlling for either of these measures in equation 1.5 does not significantly alter the

in the review. To avoid overweighing long reviews, I recode positive sentiment counts to 1 and negative sentiment counts to -1.

[105]I only keep the CEO terms for which there are at least 120 reviews, to avoid the results being driven by a few idiosyncratic reviews.

[106]The caveats include: (i) the firm by CEO term subsample used in this analysis is small and only partly overlaps with the baseline sample, and (ii) the corporate-trust-culture and other culture measures are constructed from online reviews by only a subsample of firms' employees.

[107]This resonates with Graham et al.'s (2018) finding that corporate culture is primarily set by the current CEO.

[108]These measures are calculated in the same way as CEO's inherited trust measure using answers to relevant questions in the GSS. Confidence in the government is the average of confidence in the federal government, US Supreme Court, and Congress. On their own, only CEO's trust in the scientific community is positively and significantly associated with patenting.

CEO's trust coefficient, suggesting that the latter is unlikely to be confounded by CEO's confidence in the patent system and relevant institutions (Table A.8, columns 5 and 6).[109] Furthermore, results from the bilateral trust subsamples (Table 1.3) provide another strong piece of evidence, for in this specification CEO's propensity to patent is fully absorbed by CEO fixed effects.[110] Together, the evidence indicate that the estimates from using patents as the outcome variable do capture the effect of CEO's trust on true innovation, not just its effect on patenting propensity.

1.7.2 Timing of CEO's trust effect

Does the effect of CEO's trust increase or decrease over time? Again, there is not a clear prediction of how the effect should evolve. On one hand, it takes time for a new CEO to implement hard policies or transmit soft culture, and for R&D to materialize into innovation, so one should expect larger effect over time. On the other hand, the theory suggests that CEO's prior belief, or inherited trust, should become less important as she updates her belief after each period. Therefore, whether CEO's trust effect becomes larger or smaller remains an empirical question. Figure A.5 plots CEO's trust coefficients by the duration of CEO's tenure in the firm.[111] It may seem surprising that there is a rather immediate effect, but note that the patent outcome variable is one-year forward. That is, CEO's trust has largest effect on patents filed in her third year in the firm. As researchers could anticipate future policy and culture changes following a CEO transition, it is likely that they adjust their project choices accordingly immediately after the transition', even before those changes materialize. In addition, since Hall et al. (1986) it has been shown that patent applications are often timed quite closely to R&D, and in a few exceptions such as pharmaceuticals, I observe no effect of CEO's trust on firm's patents (Table 1.9, column 5). Finally, the declining trend is suggestive of the presence of CEO's belief updating over time as implied by the model, even though the coefficients are not statistically different from one another.

1.7.3 Heterogenous effects by CEO and firm

Table 1.8 investigates how CEO's trust effect varies with CEO background. Column 1 interacts the CEO's inherited trust with her highest education level.[112] While both

[109] The resulting CEO's trust estimate (standard error) is 0.069 (0.017) with control for CEO's confidence in the government, and is 0.063 (0.018) with control for CEO's confidence in the scientific community. I do not include both controls in the same regression as they are highly correlated (correlation of 0.76).

[110] Even if CEOs may have different levels of confidence in the patent systems of different countries, it is unlikely to be an issue for Table 1.3's results. This is because the countries in which firms would file for patent protection are not necessarily the home countries of inventors, which is especially true for the subsample of US-based inventors in which patents are mostly file in the US.

[111] Specifically, Figure A.5 plots the coefficients $\hat{\beta_k}$ for $k \in [1,9]$ from estimating: $\text{asinh}(pat_{fd,t+1}) = \sum_{k=1}^{9} (trust_{fdt} \times tenure_{dk} \times successor_d) + \mathbf{X}_{ft} + \mathbf{Z}_{dt} + \theta_E + \omega_t + \varepsilon_{fdt}$ using the transition-event sample, in which (i) $tenure_{dk}$ is an indicator equal to 1 if the CEO d starts working in firm f in year $t-k+1$, and (ii) $sucessor_d$ is an indicator equal to 1 if CEO d is the successor in transition E.

[112] Education level below college degree is coded to -1, college degree – 0, master degrees – 1, and doctorate – 2.

CEO's inherited trust and education level positively impact innovation, their interaction term is negative, indicating that inherited trust and education are substitutes. To better understand what kind of knowledge reduces the effect of prior belief, column 2 to 4 interacts CEO's trust with dummies indicating if the CEO has at least a master degree (column 2), a doctorate (column 3), and a postgraduate degree that is not an MBA (column 4). The interaction terms in columns 2 and 3 suggest that CEO's trust effect is halved if the CEO has some postgraduate education and eliminated if the CEO has a doctorate. Most interestingly, the negative interaction term is largest in magnitude and statistically significant in column 4, suggesting that it is technical knowledge that reduces the effect of trust. Similarly, column 6 reports a negative interaction term between CEO's trust and a dummy indicating if she has prior R&D experience. These results imply that trust is a substitute for knowledge of R&D processes and that prior belief becomes less important with exposure and experience.[113]

Table 1.9 reports some additional results on heterogeneous effects by firm size and industry. First, interaction terms between CEO's trust and second-order polynomial of firm's size decile (with respect to its 3-digit industry) suggest that the effect is largest among median-size firms. This possibly reflects the observation that it is more difficult for CEOs to have considerable impact on researchers in very large firm (while R&D and innovation may be less relevant for very small ones). Second, CEO's trust effect is considerable larger in ICT and electronic sectors (10.5%, statistically significant at 5% level). This is consistent with the argument that CEO's trust effect should be more visible where the lag between R&D and patents is shorter, and where the firm is smaller.[114] This effect in the remaining sectors, however, is also positive and statistically significant, although of smaller magnitude (4.7%, statistically significant at 5% level). Put different, the effect of CEO's trust on innovation is ubiquitous across different industry sectors.

1.7.4 CEO's practice: qualitative insights

This paper's ensemble of quantitative evidence on the effect of CEO trust on patents may lead to questions on how exactly CEOs influence innovation processes within the firm. Channels highlighted in my model include the management of researchers in terms of recruitment, retention, and incentives. While it remains a major challenge to run large-scale surveys with quantitative questions on top managers' practices,[115] one can still get some qualitative insights from a recent survey on leadership and innovation

[113] In the context of cross-country venture capital investment, Bottazzi et al. (2016) similar finds that education and work experience reduce the effect of bilateral trust on investment.

[114] The average firm in ICT is about half the size (as measured by employment) of the average firms in the remaining sectors. Yet they file about over half of all the patents in the sample.

[115] An exception is Bandiera et al.'s (2017) data on CEO's time use and practices. Even in the recent literature on management surveys since Bloom and Van Reenen (2007), in most cases one can only obtain information from lower-level staffs.

among senior managers worldwide.[116] It is interesting to note that many managers echo support for my model's assumptions. The leadership team are often directly involved in personnel decisions regarding innovation, and other big-picture decisions, but do not "have a lot of control over the innovation process," especially in measuring efforts towards innovation. Similar to my model's setting, while the leadership actively picks and retains innovators, their actions and efforts cannot be observed or monitored.[117]

Top managers' views also corroborate this paper's insights. When asked about processes of great impact on improving innovation performance, the second most agreed answer is about promoting risk taking that encourages innovation.[118] Shortermism and fear of failure are also ranked among the top of inhibitors of innovation. While almost all respondents agree that people and corporate culture are the most important determinants of innovation, top managers' top worry is not having the right talent, yet employees are most concerned about firm's culture. In that regard, trust and engagement are seen as the most important for a strong performance in innovation.

To summarize the qualitative insights from the survey, Barsh et al. (2008) recommends fostering an "innovation culture based on trust," in which people trust that it is safe to pursue risky ideas and paths. While one should be cautious of the methodological rigor of a qualitative, open-end survey, the qualitative insights suggest that trust has in practice been considered as an important driver of corporate innovation.

1.8 Concluding remarks

Let us recall a well-known tale among generations of employees at IBM, one of the most innovative corporations in the 20$^{\text{th}}$ century that has relied on its culture of tolerance of failure to encourage exploration and innovation. Thomas Watson Sr., IBM's founder, was once discussing a ten million dollar mistake one of his executives had just made. "I guess you want my resignation," said the executive. Watson replied, "You can't be serious. We have just spent ten million dollars educating you."[119] The anecdote highlights this paper's message on the role of such trusting CEOs in inculcating a corporate culture of tolerance of failure, which can lead to more innovation.

More generally, this paper provides a broad range of empirical evidence on the association between CEO's trust and firm innovation, measured by patent quantity and quality. I measure a CEO's inherited trust based on her ethnic origins as inferred from her last name. Using within-firm changes in CEOs, I find that one standard deviation

[116]The survey was conducted by *The McKinsey Quarterly* (Barsh et al., 2007, 2008) in September 2007, covering 722 executives at the senior vice president level and above and 736 lower-level executives around the world.

[117]Regarding personnel decisions, there is a broad range of variation, as my model predicted: Innovators are only "protected" in about a third to a half of the surveyed firms, and across firms tolerance of failure in innovation varies greatly, with failure in innovation ranging from an opportunity to learn to a significant threat to one's career.

[118]The only slightly more popular answer is "Making innovation a core part of the leadership agenda."

[119]The anecdote is recounted in Ederer and Manso (2013), based on Bennis and Nanus (1997).

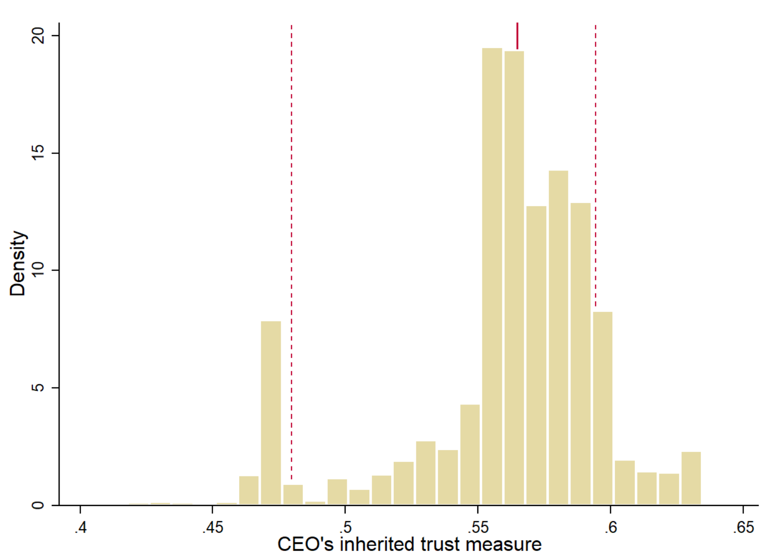
increase in CEO's generalized trust is associated with over 6% increase in annual patent counts and 4% increase in average patent quality. I further use CEO's bilateral trust towards researchers in or from different countries to show that this increase is generated by inventors towards whom the CEO's bilateral trust is stronger, in a specification that controls for firm by year, CEO, and inventor country fixed effects. Given the presence of stringent fixed effects and rigorous controls, these results are unlikely to be driven by confounding factors such as CEO's country of origin or individual characteristics.

These empirical findings are best understood in a simple principal-agent model of exploration versus exploitation with unobserved researcher's type, in which the CEO's trust encourages a good researcher to undertake high-risk explorative R&D thorugh her tolerance of failure. The model further predicts that (i) CEO trust's effect on innovation is driven by high-quality patents, and that (ii) it is larger among firms with better researcher quality, both of which are confirmed in the data. I also find that CEO's inherited trust robustly predicts corporate trust culture in a sample of close to 400 major US firms for which measures of corporate culture are constructed from text analysis of online employee reviews.

The paper's results fit in a crucial gap in the recent literature on long-term development and trust, in showing micro-evidence of how trust may affect innovation, the indispensable determinant of productivity growth in the long run. The paper also broadens our understanding of the impact of CEO's traits on firm's decisions, performance, and culture (Bertrand and Schoar, 2003).

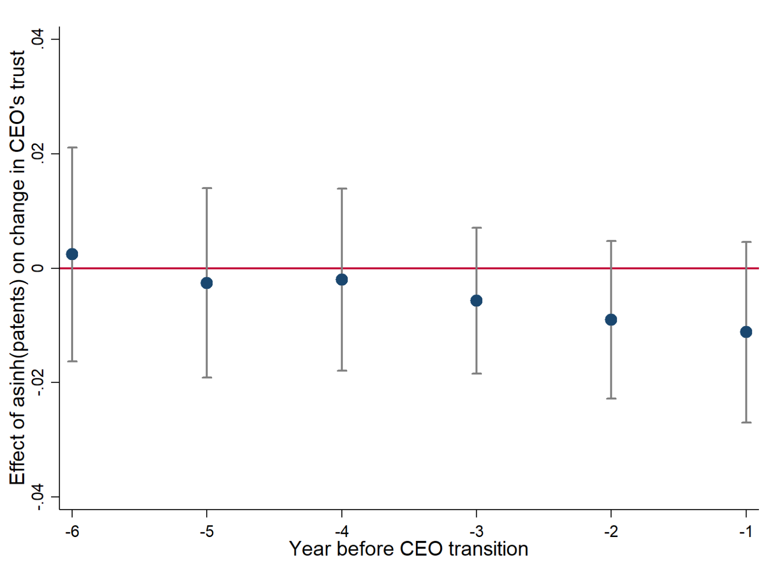
The current set of results leave out a few important questions to future work. First, while the main mechanism works through tolerance of failure, we do not have a direct measure of tolerance. Second, it may be fruitful to model the match between CEOs and firms, and control for it explicitly in trying to estimate the effect of CEO's trust on firms. Third, it would be interesting to better understand in which context does excessive trust lead to suboptimal innovation and performance.

Figure 1.1: Distribution of CEO's inherited trust measure



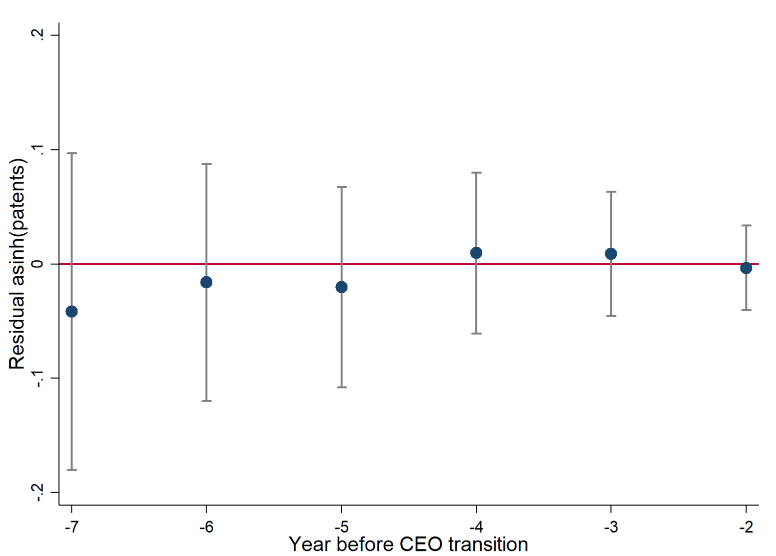
Notes: This figure shows the 1st-99th percentile distribution of CEO's GSS-based inherited generalized trust measure as described in subsection 1.3.2 for 5,753 CEOs in the baseline sample. The solid vertical line corresponds to the 50th percentile of the distribution. The dashed vertical lines correspond to the 10th and 90th percentiles of the distribution.

Figure 1.2: Pre-change patents and change in CEO's trust



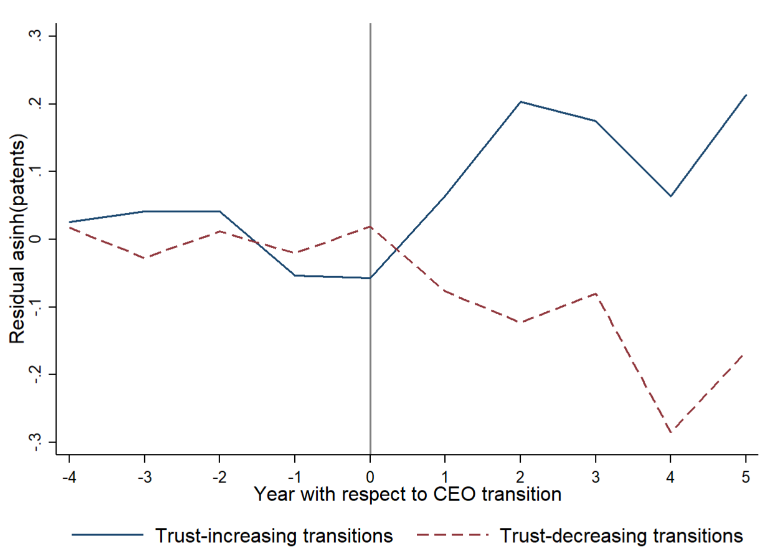
Notes: This figure plots the coefficients $\hat{\gamma}_k$ for $k \in [-6, -1]$ from estimating: $\Delta trust_{fdt} = \sum_{k=-7}^{-1} \gamma_k (\text{asinh}(pat_{fdt}) \times event_{t-k}) + \beta trust_{fdt} + \mathbf{X}_{ft} + \mathbf{Z}_{dt} + \omega_t + \varepsilon_{fdt}$, in which (i) $\Delta trust_{fdt}$ is the difference between CEO d's and her successor's trust measures, (ii) $event_{t-k}$ is an indicator equal to 1 if the transition happens in year $t-k$, and (iii) $\mathbf{X}_{ft}$ additionally includes a full set of firm's SIC3 industry dummies. Estimates are shown with their 95% confidence intervals. Standard errors are clustered by SIC3 industry.

Figure 1.3: Pre-trend in patents



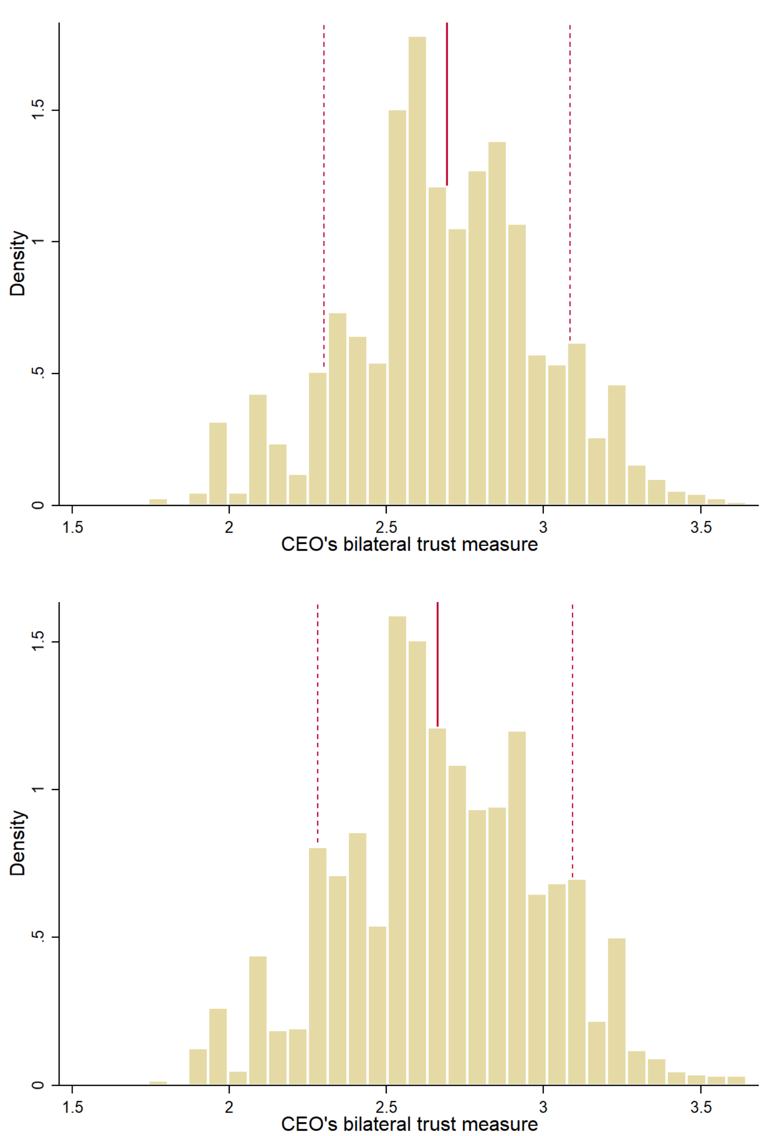
Notes: This figure plots the coefficients $\hat{\gamma}_k$ for $k \in [-7, -2]$ relative to $\hat{\gamma}_{-1}$ from estimating: $\text{asinh}(pat_{fdt}) = \sum_{k=-7}^{-1} \gamma_k event_{t-k} + \mathbf{X}_{ft} + \mathbf{Z}_{dt} + \theta_f + \omega_t + \varepsilon_{fdt}$, in which $event_{t-k}$ is an indicator equal to 1 if the next CEO transition happens in year $t-k$. Estimates are shown with their 95% confidence intervals. Standard errors are clustered by firm.

Figure 1.4: Patents by change in CEO's trust (matched sample)



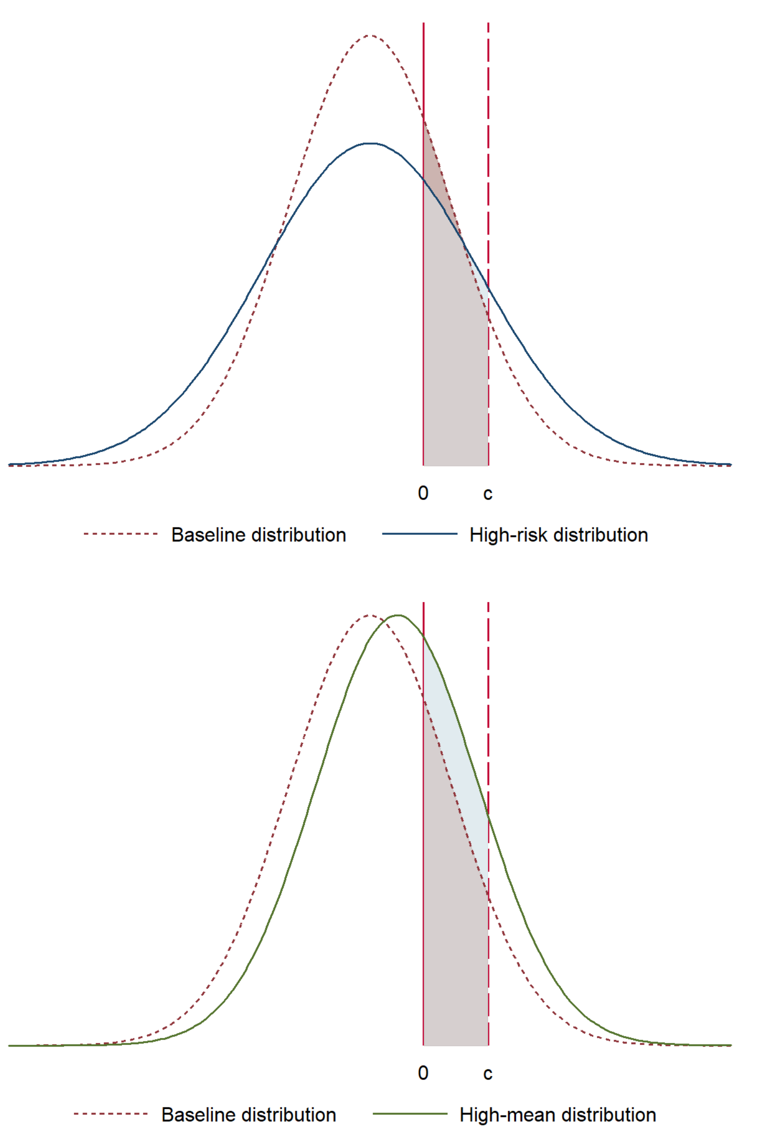
Notes: This figure plots firms' average residual patent application counts (after partialling out the covariates) by year with respect to CEO transition year (i.e., year 0). Among the sample of CEO transitions in which both predecessor's and successor's tenures are at least 5 years, each transition in which the new CEOs are *more* trusting than their predecessors (i.e., trust-increasing transition) is matched to a transition in which the new CEOs are *less* trusting than their predecessors (i.e., trust-decreasing transition) based on their average pre-transition residual patent counts. The solid blue line groups together all trust-increasing CEO transitions and the dotted red line corresponds to their matched trust-decreasing CEO transitions. Each group's annual average residual patent counts are plotted relative to the group's pre-transition mean, which is normalized to 0.

Figure 1.5: Distribution of CEO's bilateral trust towards researchers



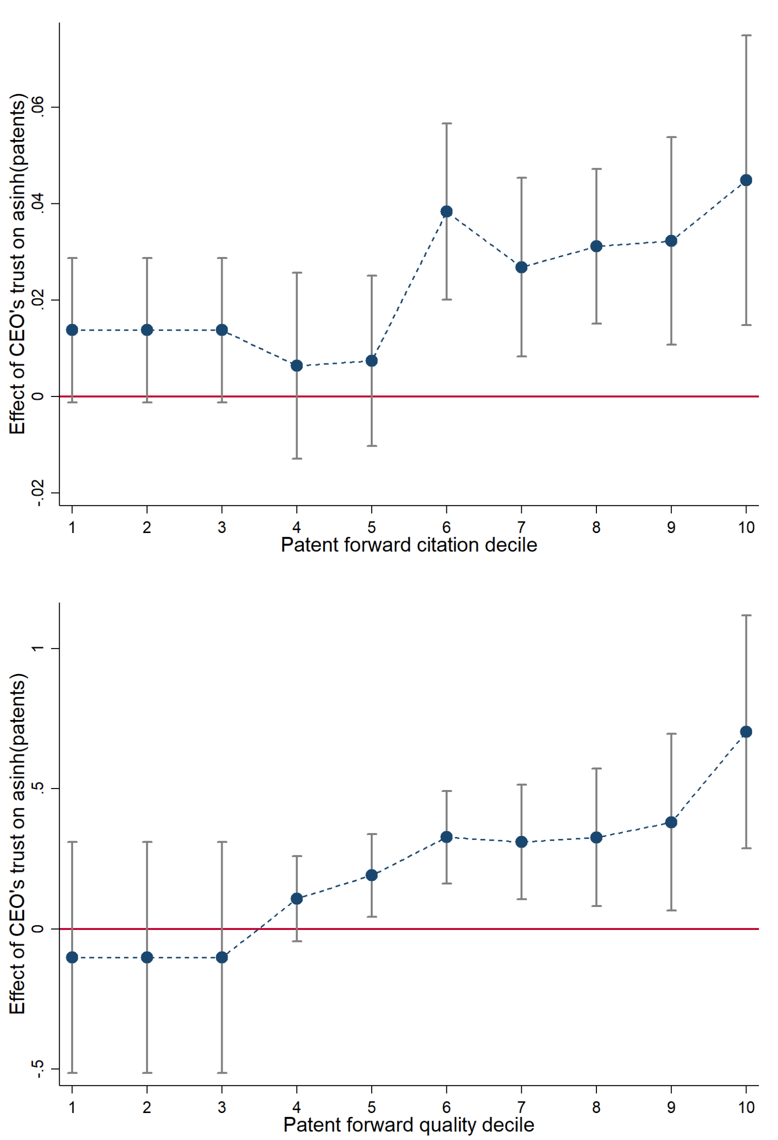
Notes: This figure shows the distribution of CEO's bilateral trust measure as described in subsection 1.3.2 for CEO-inventor country pairs in the baseline bilateral samples. The upper plot corresponds to the bilateral trust sample in which an inventor's country is inferred from his patent-listed address for non-US-based inventors. The lower plot corresponds to the bilateral trust sample in which an inventor's country is additionally inferred from his last name for US-based inventors. The solid vertical line corresponds to the 50th percentile of the distribution. The dashed vertical lines correspond to the 10th and 90th percentiles of the distribution.

Figure 1.6: Project outcome distributions under different mechanisms



Notes: The top figure illustrates the spread of project outcome distributions under more exploration; the bottom figure the rightward shift under greater effort. The dotted red line (both figures) corresponds to the baseline project outcome distribution. The solid blue line (top figure) corresponds to the same distribution under high-risk exploration, which is a mean preserving spread of the baseline. The solid green line (bottom figure) corresponds to the same distribution under greater effort, which is rightward shift of the baseline. The solid vertical line at 0 represents the quality threshold above which projects get patented and become observable. The dashed vertical line corresponds to a quality threshold c.

Figure 1.7: CEO's trust effect by patent quality decile



Notes: This figure plots the effects of CEO's trust on firm's patent counts in different quality deciles estimated using equation (1.5), using as the dependent variable the inverse hyperbolic sine of patent counts in the upper plot and winsorized patent counts in the lower plot. A patent's quality decile is computed based on its forward citation counts with respect to its technology field $\times$ year cohort. Estimates are shown with their 95% confidence intervals. Standard errors are clustered by CEO's main ethnicity.

Table 1.1: Baseline effect of CEO's trust on firm's patents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	**asinh(Future patent applications)**						
CEO's trust (baseline)	0.059***	0.063***	0.063***	0.066***			
	(0.018)	(0.017)	(0.019)	(0.019)			
CEO's trust (LASSO)					0.067***		
					(0.021)		
CEO's trust (full GSS)						0.041**	
						(0.018)	
CEO's trust (WVS)							0.031**
							(0.013)
Firm & Year FEs	X	X	X	X	X	X	X
CEO controls		X	X	X	X	X	X
Firm controls			X	X	X	X	X
Industry $\times$ Year FEs				X			
Observations	29,384	29,384	29,384	29,384	29,384	29,384	29,384
Firms	3,598	3,598	3,598	3,598	3,598	3,598	3,598

Notes: This table reports the baseline effect of CEO's inherited trust on firm's patents using equation (1.5). Baseline sample includes all observations of firm f $\times$ year t $\times$ its current CEO d. The dependent variable is the inverse hyperbolic sine of firm f's patent application counts in year $t+1$. The explanatory variable is CEO d's trust measures, constructed from: (i) CEO's ethnic origins and GSS trust question, considering only respondents in highly prestigious occupations (subsection 1.3.2) (columns 1-4); (ii) all commonly observable characteristics of CEOs and GSS respondents, including ethnic origin, age, gender, education, and birth cohort, using LASSO (column 5); (iii) CEO's ethnic origins and GSS trust question, considering the full sample of respondents (column 6); and (iv) CEO's ethnic origins and WVS trust question (column 7). All trust measures are standardized by the standard deviation of GSS-based inherited trust measure at ethnicity level. Baseline controls include (i) firm's age, age squared, ln(total assets), ln(sale), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. Standard errors are clustered by CEO's main ethnicity.

*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table 1.2: Controlling for confounding variables

Panel A. Controlling for home countries' level of development

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**asinh(Future patent applications)**					
CEO's trust	0.053***	0.061***	0.058***	0.060***	0.067***	0.045**
	(0.017)	(0.019)	(0.019)	(0.019)	(0.019)	(0.020)
ln(GDP)	0.003					-0.015
	(0.013)					(0.027)
ln(Population)	-0.013					-0.006
	(0.011)					(0.020)
GDP growth (%)	0.005**					0.004*
	(0.002)					(0.002)
High school grad (%)		-0.001				-0.000
		(0.000)				(0.001)
Governance quality (pct)			0.022			0.008
			(0.055)			(0.102)
ln(US trade volume)				0.005		0.020*
				(0.008)		(0.010)
ln(Patent applications)					-0.007	-0.003
					(0.005)	(0.012)
Firm & Year FEs	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X
Observations	29,384	29,384	29,384	29,384	29,384	29,384
Firms	3,598	3,598	3,598	3,598	3,598	3,598

Notes: This panel controls for CEO's home countries' macroeconomic characteristics using equation (1.5). Baseline sample includes all observations of firm f $\times$ year t $\times$ its current CEO d. The dependent variable is the inverse hyperbolic sine of firm f's patent application counts in year $t+1$. The explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2). Baseline controls include (i) firm's age, age squared, ln(total assets), ln(sale), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. CEO's home country controls include year t's ln(GDP), ln(population), and GDP growth rate (column 1), population share of high school graduates (column 2), average percentile ranking of World Bank governance indices (column 3), ln(US exports + US imports) (column 4), and ln(total patent applications filed at the country's patent office) (column 5). Column (6) controls for all those variables. Standard errors are clustered by CEO's main ethnicity.
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Panel B. Controlling for other cultural traits

	(1)	(2)	(3)	(4)	(5)
Dependent variable:			**asinh(Future patent applications)**		
CEO's trust	0.063***	0.052**	0.068***	0.053**	0.053**
	(0.018)	(0.019)	(0.019)	(0.021)	(0.023)
College grad (%)	0.073			0.198**	0.143
	(0.100)			(0.093)	(0.115)
Work ethic (z-score)		0.017		0.022	0.023
		(0.015)		(0.016)	(0.015)
Risk preference (z-score)			-0.038	-0.033	-0.033
			(0.023)	(0.025)	(0.029)
GDP growth (%)					0.005*
					(0.003)
ln(US trade volume)					0.001
					(0.009)
Firm & Year FEs	X	X	X	X	X
Baseline controls	X	X	X	X	X
Observations	29,384	29,384	29,384	29,384	29,384
Firms	3,598	3,598	3,598	3,598	3,598

Notes: This panel controls for CEO's other inherited cultural traits using equation (1.5). Baseline sample includes all observations of firm f $\times$ year t $\times$ its current CEO d. The dependent variable is the inverse hyperbolic sine of firm f's patent application counts in year $t+1$. The explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2). Baseline controls include (i) firm's age, age squared, ln(total assets), ln(sale), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. Column (1) controls for the share of college graduates in CEO's ethnic groups. Column (2) controls for CEO's inherited work ethic, derived from the GSS question: *"Some people say that people get ahead by their own hard work; others say that lucky breaks or help from other people are more important. Which do you think is most important?"*. Column (3) controls for CEO's inherited risk preference, proxied by the share of GSS respondents in CEO's ethnic groups who have stock market or mutual fund investments. Column (4) controls for all of those variables. Column (5) further controls for CEO's home countries' GDP growth rate and ln(US exports + US imports). Standard errors are clustered by CEO's main ethnicity.
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table 1.3: CEO's trust effect in bilateral trust samples

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:			**asinh(Future patent applications)**			
Sample: Based on inventors'	Non-US addresses			Addresses and last names		
CEO's bilateral trust	0.052**	0.051**	0.034*	0.026**	0.025**	0.013**
	(0.023)	(0.024)	(0.021)	(0.011)	(0.011)	(0.006)
Firm $\times$ Year FEs	X	X		X	X	
CEO FEs	X		X	X		X
Inventor country FEs	X			X		
CEO $\times$ Year FEs		X			X	
Inv. country $\times$ Year FEs		X			X	
Firm $\times$ Inv. country FEs			X			X
Year FEs			X			X
Observations	23,284	23,284	23,284	56,942	56,942	56,942
Firm $\times$ Inv. country's	3,481	3,481	3,481	8,554	8,554	8,554
Firms	730	730	730	1,263	1,263	1,263

Notes: This table reports the effect of CEO's bilateral trust towards a country on patents by inventors from that country using equation (1.6). Samples include all observations of firm f $\times$ year t $\times$ its current CEO d $\times$ country c such that firm f has patents by inventors from country c during 2000-2012. An inventor's country is inferred from his patent-listed address for non-US-based inventors in columns (1)-(3), and additionally from his last name for US-based inventors in columns (4)-(6). The explanatory variable is CEO d's bilateral trust towards individuals from country c, standardized by its standard deviation at country pair level. The dependent variable is firm f's total patent application counts by inventors from country c in year $t+1$. Standard errors are clustered by CEO's main ethnicity $\times$ inventor country.
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table 1.4: Bilateral trust effect with country pairwise controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	asinh(Future patent applications)							
Sample: Based on inventors'	Non-US addresses				Addresses and last names			
CEO's bilateral trust	0.047*	0.042*	0.052**	0.052**	0.016	0.024*	0.023**	0.027**
	(0.029)	(0.025)	(0.024)	(0.024)	(0.013)	(0.013)	(0.011)	(0.011)
Common language dummy		0.057				0.011		
		(0.047)				(0.025)		
Geographical distance (1000km)			-0.002				-0.011	
			(0.019)				(0.011)	
Genetic distance (z-score)				0.018				-0.019
				(0.084)				(0.053)
Excl. same-country pairs	X				X			
Firm × Year FEs	X	X	X	X	X	X	X	X
CEO FEs	X	X	X	X	X	X	X	X
Inventor country FEs	X	X	X	X	X	X	X	X
Observations	20,878	23,284	23,284	22,881	51,936	56,942	56,942	55,444
Firm × Inv. country's	3,145	3,481	3,481	3,421	7,932	8,554	8,554	8,323
Firms	496	730	730	728	1,020	1,263	1,263	1,263

Notes: This table controls for other CEO-inventor country pairwise characteristics besides bilateral trust using equation (1.6). Samples include all observations of firm f × year t × its current CEO d × country c such that firm f has patents by inventors from country c during 2000-2012. An inventor's country is inferred from his patent-listed address for non-US-based inventors in columns (1)-(4), and additionally from his last name for US-based inventors in columns (5)-(8). The explanatory variable is CEO d's bilateral trust towards individuals from country c, standardized by its standard deviation at country pair level. The dependent variable is firm f's total patent application counts by inventors from country c in year $t+1$. Columns (1) and (5) exclude same-country CEO-inventor country pairs. Columns (2) to (6) control for CEO-inventor country pairwise distances, including: (i) whether the countries share a common language (columns 2 and 6), (ii) weighted geographical distance between the countries (columns 3 and 7), and (iii) weighted genetic distance between the countries' populations (columns 5 and 8) (Spolaore and Wacziarg, 2016). Standard errors are clustered by CEO's main ethnicity × inventor country.
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table 1.5: CEO's trust effect on quality-weighted patents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent var:	asinh(Future quality-weighted patents)							
Quality measure:	Forward cites	Backward NPL cites	Tech. scope	Generality	Originality	Granted all	Granted USPTO	Average cites
CEO's trust	0.100***	0.103***	0.061**	0.053***	0.050***	0.049***	0.056***	0.044**
	(0.031)	(0.031)	(0.025)	(0.015)	(0.015)	(0.016)	(0.014)	(0.020)
Firm & Year FEs	X	X	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X	X	X
Observations	29,384	29,384	29,384	29,384	29,384	29,384	29,384	29,384
Firms	3,598	3,598	3,598	3,598	3,598	3,598	3,598	3,598

Notes: This table reports CEO's trust effect on quality-weighted patents using equation (1.5). Baseline sample includes all observations of firm f × year t × its current CEO d. The dependent variable is the inverse hyperbolic sine of firm f's patent application counts in year $t+1$, weighted by: forward citations (column 1); backward citations to non-patent (i.e., scientific) literature (column 2); patent technological scope (column 3); generality index (i.e., technological diversity of forward citations) (column 4); originality index (i.e., technological diversity of backward citations) (column 5); granted patents (column 6); and granted USPTO patents (column 7). The dependent variable in column 8 is the inverse hyperbolic sine of firm f's average forward citations per patent in year $t+1$ (or zero if firm f has zero patent applications in year $t+1$). The explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2). Baseline controls include (i) firm's age, age squared, ln(total assets), ln(sale), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. Standard errors are clustered by CEO's main ethnicity.
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table 1.6: CEO's trust effect by pre-transition researcher pool quality

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**asinh(Future patent applications)**					**asinh(R&D efficiency)**
	All patents				High quality	
CEO's trust	0.066***	0.066***	0.067***		0.057***	0.039***
	(0.018)	(0.018)	(0.019)		(0.015)	(0.013)
Trust × Proxy for pre-transition researcher quality	0.040**	0.034*	0.021		0.039**	0.025**
	(0.018)	(0.018)	(0.017)		(0.019)	(0.011)
Trust × Quality quintile 1				0.028		
				(0.042)		
Trust × Quality quintile 2				0.038		
				(0.028)		
Trust × Quality quintile 3				0.065		
				(0.052)		
Trust × Quality quintile 4				0.076**		
				(0.029)		
Trust × Quality quintile 5				0.128***		
				(0.037)		
Pre-transition window	-1 to 0	-2 to 0	All yrs	-1 to 0	-1 to 0	-1 to 0
Firm & Year FEs	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X
Observations	19,506	19,508	19,547	19,506	19,506	19,506
Events	2,278	2,279	2,285	2,278	2,278	2,278

Notes: This table explores the heterogeneous effects of CEO's trust on firm's patents by pre-transition researcher pool quality using equation (1.5) and the sample constructed from CEO transition events. For each event, I include all firm f × year t × its current CEO d observations that correspond to the predecessor's and successor's terms. The explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2). Firm-level proxy for researcher pool quality is computed from averaging the residuals from regressing patents on observable firm and CEO characteristics, controlling for SIC2 industry and year fixed effects (subsection 1.6.3) over different pre-transition windows. The dependent variable is the inverse hyperbolic sine of firm f's (i) patent application counts (columns 1 to 4), (ii) high quality (i.e., above median) patent application counts (column 5), and (iii) R&D efficiency, calculated as patent application counts over lagged R&D expenditure (column 6), in year $t+1$. Baseline controls include (i) firm's age, age squared, ln(total assets), ln(sale), asinh(R&D expenditure), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. Column (4) interacts CEO's trust measure with researcher pool quality quintile dummies (computed based on firm-level proxy for pre-transition researcher pool quality). The remaining columns interact CEO's trust measure with firm-level proxy for pre-transition pool quality. Standard errors are clustered by CEO's main ethnicity.

*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table 1.7: CEO's trust effect on "corporate trust culture"

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Corporate trust culture**					
CEO's trust	0.179***	0.193***	0.452*	0.234***	0.427**	0.755**
	(0.062)	(0.072)	(0.258)	(0.073)	(0.174)	(0.327)
CEO approval (%)				1.584***	1.540**	0.893
				(0.588)	(0.591)	(0.920)
College grad (%)					-0.687	-1.397
					(1.100)	(1.653)
Work ethic					-0.254*	-0.256*
					(0.143)	(0.152)
Risk preference					-0.286	-0.693*
					(0.246)	(0.362)
Industry (SIC3) FEs	X	X		X	X	
Baseline controls		X	X	X	X	X
Firm FEs			X			X
Observations	393	393	393	393	393	393
Firms	277	277	277	277	277	277
Industries (SIC3)	90	90	90	90	90	90

Notes: This table presents the effect of CEO's trust on measures of corporate culture computed from online employee reviews (Sull, 2018). Sample includes all firm $f \times$ CEO d observations over the 2008-2017 period. The explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2). The dependent variable measures firm f's "corporate trust culture" (i.e., the extent to which employees trust one another) during CEO d's term. Baseline controls include (i) firm's average age, average age squared, ln(average total assets), ln(average sale), and (ii) CEO's average age, average age squared, gender, education dummies, average tenure in firm (controls are averaged over CEO d's term in firm f). CEO approval rate is computed from online employee reviews. Additional controls for CEO's other cultural traits are as explained in Table 1.2's notes. Standard errors are clustered by SIC3 industry.

*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table 1.8: Heterogeneous effects by CEO's background

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	**asinh(Future patent applications)**				
CEO's trust	0.085***	0.082***	0.072***	0.097***	0.063***
	(0.017)	(0.019)	(0.016)	(0.020)	(0.018)
Education level	0.159				
	(0.113)				
Trust × Education level	-0.028				
	(0.022)				
Postgraduate degree dummy		0.176			
		(0.149)			
Trust × P.G. degree		-0.030			
		(0.029)			
Doctorate dummy			0.308		
			(0.299)		
Trust × Doctorate			-0.056		
			(0.058)		
Non-MBA P.G. degree dummy				0.702***	
				(0.230)	
Trust × Non-MBA P.G. degree				-0.135***	
				(0.044)	
R&D experience dummy					0.536
					(0.704)
Trust ×R&D experience					-0.135
					(0.148)
Firm & Year FEs	X	X	X	X	X
Baseline controls	X	X	X	X	X
Observations	29,384	29,384	29,384	29,384	29,384
Firms	3,598	3,598	3,598	3,598	3,598

Notes: This table explores the heterogeneous effects of CEO's trust on firm's patents by CEO education and experience using equation (1.5). Baseline sample includes all observations of firm f × year t × its current CEO d. The dependent variable is the inverse hyperbolic sine of firm f's patent application counts in year $t+1$. The explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2). Baseline controls include (i) firm's age, age squared, ln(total assets), ln(sale), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. Column (1) interacts CEO's trust measure with her highest education level (-1 – no degree, 0 – Bachelor degree, 1 – Masters degree, 2 – Doctor degree). Columns (2)-(4) interact CEO's trust measure with a dummy indicating if (i) she has a masters or doctorate (column 2), (ii) she has a doctorate (column 3), or (iii) she has a masters or doctorate but not an MBA degree (column 4). Column (5) interacts CEO's trust measure with a dummy indicating if she has prior R&D experience. Standard errors are clustered by CEO's main ethnicity.
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table 1.9: Heterogeneous effects by firm's characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**asinh(Future patent applications)**					
Sample:	By asset size	By employment size	IT/ Electr.	Non-IT/ Electr.	Pharma/ Chem.	Non-Pharma/Chem.
CEO's trust			0.105**	0.047**	0.019	0.062***
			(0.041)	(0.019)	(0.063)	(0.020)
Trust $\times$ Size quintile 1	0.020	0.011				
	(0.033)	(0.036)				
Trust $\times$ Size quintile 2	0.014	0.049				
	(0.032)	(0.031)				
Trust $\times$ Size quintile 3	0.132***	0.141***				
	(0.043)	(0.028)				
Trust $\times$ Size quintile 4	0.084**	0.045				
	(0.039)	(0.035)				
Trust $\times$ Size quintile 5	0.077***	0.075**				
	(0.022)	(0.031)				
Firm & Year FEs	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X
Observations	29,384	28,713	6,958	22,426	3,310	26,074
Firms	3,598	3,577	884	2,715	438	3,161

Notes: This table explores the heterogeneous effects of CEO's trust on firm's patents by firm size and industry using equation (1.5). Baseline sample includes all observations of firm f $\times$ year t $\times$ its current CEO d. The dependent variable is the inverse hyperbolic sine of firm f's patent application counts in year $t+1$. The explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2). Baseline controls include (i) firm's age, age squared, ln(total assets), ln(sale), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. Columns (1) and (2) interact CEO's trust measure with firm's size quintile dummies, computed with respect to firm's 3-digit industry $\times$ year cohort using total assets or employment as firm size measure. Columns (3) corresponds to the subsample of firms in ICT and electronic industries (i.e., computer and data processing services (SIC 737), computer and office equipment (SIC 357), electronic and other equipment (SIC 36)) and column (4) – the remaining subsample. Column (5) corresponds to the subsample of firms in pharmaceutical and chemical industries (i.e., chemicals and allied products (SIC 28)) and column (6) – the remaining subsample. Standard errors are clustered by CEO's main ethnicity.
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Chapter 2

Do Tax Incentives Increase Firm Innovation? An RD Design for RDD

This chapter is jointly co-authored with Antoine Dechezleprêtre, Elias Einiö, Ralf Martin, and John Van Reenen.

This chapter presents evidence of causal impacts of research and development (R&D) tax incentives on own-firm innovation and technological spillovers. Exploiting a change in the asset-based size thresholds that determine eligibility for R&D tax subsidies we implement a Regression Discontinuity Design using administrative tax data. There are statistically and economically significant effects of tax on R&D and (quality-adjusted) patenting that persist up to 7 years after the change. R&D tax price elasticities are large, with a lower bound of 1.1, consistent with the fact that the treated group is drawn from smaller firms that we show are more likely subject to financial constraints. Our RD design has power to identify positive spillovers on technologically-close peers, when these neighbors are in a sufficiently small technology class with the treated firms.

2.1 Introduction

Innovation is recognized as the major source of growth in modern economies. But because of knowledge externalities, private returns on research and development (R&D) are generally estimated to be lower than their social returns, suggesting the need for some public subsidy.[1] Consequently, not only does every country treat R&D investments more generously than capital investment, but the majority of OECD countries also have additional fiscal incentives for R&D. Over the last two decades, these tax incentives have grown more popular compared to more direct R&D subsidies to firms.[2]

But do R&D tax incentives really increase innovation? In this paper, we identify the impact of R&D tax incentives by exploiting a policy reform in the UK which raised the size threshold under which firms can access the more generous tax regime for small- and medium-sized enterprises (SMEs). Importantly, the new SME size threshold introduced was unique to the R&D Tax Relief Scheme, and does not overlap with access to other programs or taxes. Given this change, we can implement a Regression Discontinuity (RD) Design (Lee and Lemieux, 2010) at differences in innovation activity around the new SME threshold, which was based on accounting data pre-dating the policy change. We show that there are no discontinuities in any outcome around the threshold in the years prior to the policy change.

To analyze the impact of the R&D tax incentive on innovation activity, we use a newly assembled dataset that links the universe of UK companies with their confidential Corporate Tax returns (including firms' R&D expenditures) from the HMRC (the UK equivalent of the US IRS), their patent filings in all major patent offices in the world, and their financial accounts. The data is available before and after the R&D tax change, allowing us to analyze the causal impact of the tax credit up to seven years after the policy change.

A key advantage of our firm-level patent dataset is that it enables us to assess the effect of tax incentives not only on R&D spending (an input) but also on innovation outputs.[3] Indeed, the tax incentive could increase observed R&D without having much effect on innovation if, for example, firms relabeled existing activities as R&D to take advantage of the tax credits (e.g., Chen et al. (2018)) or only expanded very low-quality

[1]Typical results find marginal social rates of return to R&D between 30% and 50% compared to private returns between from 7% to 15% (Hall et al., 2010). Endogenous growth theories Romer (1990); Aghion and Howitt (1992) provide several reasons why private innovative activities do not take into account externalities over producers and consumers, and produce sub-optimal levels of R&D. For evidence showing R&D externalities, see for example Bloom et al. (2013). There is also evidence that these spillovers are partially localized geographically, so the country where the R&D is performed obtains a disproportionate share of the productivity benefits, at least initially (e.g., Jaffe et al., 1993).

[2]Over the period 2001-2011, R&D tax incentives were expanded in 19 out of 27 OECD countries (OECD, 2014). One reason for this shift is that subsidizing R&D through the tax system rather than direct grants reduces administrative burden and mitigates the risk of "picking losers" (e.g., choosing firms with low private and social returns due to political connections, e.g., Lach et al., 2017).

[3]There is a large literature on the effects of public R&D grants on firm and industry outcomes such as González et al. (2005), Takalo et al. (2013), Einiö (2014), Goodridge et al. (2015), Jaffe and Le (2015), and Moretti et al. (2016). The earlier literature is surveyed in David et al. (2000).

R&D projects. We can also directly examine the quality of these additional innovations through various commonly used measures of patent value, such as citations received and family size.

Another major advantage of our data and study design is that we are able to assess the impact of R&D tax credits on SMEs, allowing us to compare the response of SMEs to that of large firms, which have been the focus of most of the existing literature because accounting regulations in most countries only require larger (usually public listed) firms to report R&D. But since at least Arrow (1962), it has been recognized that financial markets may under-supply credit for R&D and these problems are likely to be particularly acute for SMEs.[4]

We find large effects of the tax policy on R&D and patenting activity. Following the policy change, R&D more than doubled in firms below eligibility threshold, followed by about a 60% increase in patenting. There is no evidence that these innovations are of lower value. We can reject absolute elasticities of R&D with respect to its user cost of less than 1.1 with a 5 percent level of confidence.[5] Our higher elasticities are likely to be because the sub-population "randomized in" by the RD Design is composed of smaller firms than have usually been examined and so are more likely to be credit constrained and therefore are also more responsive to R&D tax credits. We confirm this intuition by showing the response is particularly strong for firms in industries that are more likely to be subject to financial constraints.

The main economic rationale given for more generous tax treatment of R&D is that there are technological externalities, so that the social return to R&D exceeds the private return. Our design also allows us to estimate the causal impact of tax policies on R&D spillovers, i.e., innovation activities of firms that are technologically connected to policy-affected firms, through employing a similar RD Design specification with connected firms' patents as the outcome variable of interest. We find evidence that the R&D induced by the tax policy generated positive spillovers on innovations by technologically related firms, especially in small technology classes. Focusing on these smaller peer groups is exactly where we expect our design to have power to detect spillovers (see Angrist, 2014 and Dahl et al., 2014). Simple partial equilibrium calculations suggest that over 2006-2011 the UK R&D policy induced about $2 of private R&D for every $1 of taxpayer money and that aggregate UK business R&D would have been about 13% lower in the absence of the policy.[6]

Related Literature. Most directly, our paper contributes to the literature which seeks to evaluate the causal impact of tax policies on firms' R&D. Earlier evaluations were

[4]Since R&D costs are mainly people, it is hard to post collateral to borrow against R&D projects. Furthermore, asking outsiders for finance may reveal the innovation and so undermine its value.

[5]See surveys by Becker (2015), OECD (2013) or Hall and Van Reenen (2000) on R&D to user costs elasticities. The mean elasticities are usually between 1 and 2 whereas our mean results are twice as large.

[6]See Akcigit et al. (2017) and Acemoglu et al. (2018) for rigorous discussion of optimal taxation and R&D policy in general equilibrium.

conducted at the state or macro-economic level face the problem that changes of policies are likely to be coincident with many unobserved factors that may influence R&D.[7] Recent studies use firm-level data and more compelling causal designs, but focus on the impact of R&D tax credits on R&D expenditures. Agrawal et al. (2014) exploit a change in eligibility rules for R&D tax credits under Canada's tax incentive program to look at the tax effect on R&D investment for small private Canadian firms. They show that firms eligible to benefit from the more generous tax credit program spent 15 percent more on R&D following the program change, compared to firms with the same taxable income before the change. Rao (2016) uses administrative tax data and looks at the impact of US tax credits on R&D (but not other firm outcomes). In her approach she uses the changes in the Federal tax rules interacted with lagged firm characteristics to generate instrumental variables for the firm-specific user cost of R&D. Bøler et al. (2015) employ a difference-in-differences strategy to investigate how the introduction of R&D tax credit in Norway affects profits, intermediate imports and R&D.[8] Guceri (2018) and Guceri and Liu (2017) also use a diff-in-diffs approach to uncover positive effects from recent UK R&D tax policies.[9] Chen et al. (2018) is perhaps the closest paper to ours. The authors examine the impact of tax changes on R&D and other outcomes in a sample of Chinese firms using an RD Design. They find positive impacts, although there are large amounts of relabeling (much more than we find empirically in our application). Our paper, like Chen et al. (2018), relies on an RD design, which relaxes the stronger identification conditions implicit in a diff-in-diffs design. We are distinct from these papers by focusing on causal tax policy effects on innovation outcomes in addition to R&D inputs.

Second, we relate to the literature that examines the impact of research grants using ratings given to grant applications as a way of generating exogenous variation around funding thresholds. Jacob and Lefgren (2010) and Azoulay et al. (2015) examine NIH grants; Ganguli (2017) looks at grants for Russian scientists and Bronzini and Iachini (2014); and Bronzini and Piselli (2016) study firm R&D subsidies in Italy.[10] Howell (2017) uses the ranking of US SBIR proposals for energy R&D grants and finds significant effects of R&D grants on future venture capital funding and patents. Like us, she

[7] For example, Bloom et al. (2002), Wilson (2009), Chang (2018).

[8] See also Czarnitzki et al. (2011), Cappelen et al. (2012), and Bérubé and Mohnen (2009) who look at the effects of R&D tax credits on patents and/or new products. Branstetter and Sakakibara (2002) examine Research Joint Ventures and patents.

[9] Although complementary to our paper they look only at R&D and not at innovation outcomes or spillovers. Further, they condition on R&D performing firms which creates selection issues and means that they cannot look at the extensive margin (i.e., they cannot examine whether any firms start or stop performing R&D as a result of the tax changes).

[10] The authors look at the impacts of R&D subsidies on investment and patents in Northern Italy in an RD Design (they do not have R&D data). In their setting the running variable is based on project application scoring by a committee of experts. They observe a discontinuity in the score distribution around the eligibility cut-off, which they interpret as a sign of program managers being able to assign higher scores for projects just below the cut-off to avoid appeals.

also finds bigger effects for small firms.[11] However, none of these papers examine tax incentives directly.

Third, our paper also contributes to the literature on the effects of R&D on innovation (e.g., Doraszelski and Jordi, 2013 and the Hall et al., 2010 survey). We find that R&D has a positive causal effect on innovation, with elasticities that are underestimated in conventional OLS approaches. Although there is also a large literature on R&D spillovers (e.g., Bloom et al., 2013) we are to our knowledge, the first to provide evidence for the existence of technology spillovers in an RD setting.

Finally, we connect to an emerging field which looks at the role of both individual and corporate tax on inventors. This literature also appears to be finding an important role for taxation on mobility, quantity and quality of innovation.[12]

The paper is organized as follows. Section 2.2 details the institutional setting; Section 2.3 explains the empirical design; Section 2.4 describes the data; Section 2.5 presents the main results; Section 2.6 – technology spillover analysis; Section 2.8 concludes.

2.2 Institutional setting

We give more details of the institutional setting and tax policies in Appendix B.1 (e.g., Table B.1 details the policy changes over time), but summarize the most important features in this section. From the early 1980s the UK business R&D to GDP ratio fell, whereas it rose in most other OECD countries. In 2000, an R&D Tax Relief Scheme was introduced for small and medium enterprises (SMEs) and it was extended to cover large companies in 2002 (but SMEs continued to enjoy more generous R&D tax relief). The policy cost the UK government £1.4bn in 2013 alone (Fowkes et al., 2015).

The tax policy is based on the total amount of R&D, i.e., it is volume-based rather than calculated as an increment over past spending like the US R&D tax credit. It works mostly through enhanced deduction of R&D from taxable income, thus reducing corporate tax liabilities.[13] At the time of its introduction, the scheme allowed SMEs to deduct an additional *enhancement rate* of 50% of qualifying R&D expenditure from taxable profits (on top of the 100% deduction that applies to any form of current expenditure). If an SME was not making profits, it could surrender enhanced losses in return for a payable *tax credit*[14] amounting to 16% of enhanced R&D.[15] This design feature was aimed at dealing with the problem that smaller companies may not be making enough profits to benefit from the enhancement rate. The refundable aspect of

[11]Larger program effects for smaller firms are also found in several paper such as Wallsten (2000), González et al. (2005), Görg and Strobl (2007), Bronzini and Iachini (2014) and Zwick and James (2017).

[12]Akcigit et al. (2016); Akcigit et al. (2016); Moretti and Wilson (2017); Bell et al. (2018); Akcigit et al. (2018).

[13]Only current R&D expenditures, such as labor and materials, qualify for the scheme, but since capital only accounts for about 10% of total R&D, this is less important.

[14]Throughout we will use "tax credit" to refer to this refundable element of the scheme as distinct from the "enhanced tax deduction" element.

[15]Or equivalently, 24% (=16% x 150%) of total R&D expenditure. See Finance Act 2000 (Chapter 17, Schedule 20).

the scheme is particularly beneficial to firms which are liquidity constrained and we will present evidence in line with the idea that the large responses we observe are related to the alleviation of such financial constraints. Large companies had a less generous deduction rate of 25% of their R&D and could not claim the refundable tax credits in the case of losses (Finance Act, 2002).

The policy used the definition of an SME recommended by the European Commission (EC) throughout most of the 2000s. This was based on assets, employment and sales from the last two accounting years. It also takes into consideration company ownership structure and requires that in order to change its SME status, a company must fall in the new category in at least two consecutive years.

We focus on the major change to the scheme that commenced from August 2008. The SME assets threshold was increased from €43m to €86m, the employment threshold from 249 to 499 and the sales threshold from €50m to €100m.[16] Because of these changes, a substantial proportion of companies that were eligible only for the large company rate according to the old definition became eligible for the SME rate. In addition to the change in SME definition, the UK government also increased the enhancement rate for both SMEs and large companies in the same year. The SME enhancement rate increased from 50% to 75%.[17] For large companies, the rate changed from 25% to 30%. The policy change implies a reduction in the tax-adjusted user cost of R&D from 0.19 to 0.15 for the newly-eligible SMEs whereas the user cost for large companies was basically unchanged (subsection 2.7.2 below and Table B.2).

We examine the impact of this jump from 2008 onwards in tax-adjusted user cost of R&D at the new SME thresholds. There are several advantages of employing this reform instead of the earlier changes. First, unlike the previous thresholds based on the EU definition, which were extensively used in many other support programs targeting SMEs, the thresholds introduced in 2008 were specific to the R&D Tax Relief Scheme. This allows us to recover the effects of the R&D Tax Relief Scheme without confounding them with the impact of other policies. Second, identifying the impacts around newly introduced thresholds mitigates biases arising from tax planning which may cause endogenous bunching of firms around the thresholds. We show that there was no bunching around these thresholds in 2007 (or earlier) and covariates were all balanced at the cutoffs. This is important as although the policy was not completely detailed until July 2008 (and implemented in August 2008), aspects of the policy were understood in 2007 so firms may in principle have responded in advance. Information frictions, adjustment costs, and policy uncertainty mean that this adjustment is likely

[16] The other criteria laid down in the EC 2003 recommendation (e.g., two-year rule) were maintained in the new provision in Finance Act 2007 (Chapter 11). This act did not appoint a date on which new ceilings became effective. The date was appointed in the Finance Act 2007, Section 50 (Appointed Day) Order 2008 of July 16th, 2008.

[17] In parallel, the SME payable tax credit rate was cut slightly to 14% (from 16%) of enhanced R&D expenditure (i.e., 24.5% of R&D expenditure) to ensure that R&D tax credit falls below the 25% limit for state aid.

to be sluggish, especially for the SMEs we study.[18] The 2007 values of firm sizes are therefore what we use as running variables, as they matter for the firm's SME status in 2009 by the two-year rule, but are unlikely to be affected by tax-planning incentives.

We focus on assets as the key running variable. This is one of the three determinants of SME status and, unlike employment and sales, does not suffer from missing values in the available datasets. We discuss this in detail in Section 2.4, and we also consider using employment and sales in subsection 2.7.6, which yields similar results.

2.3 Empirical strategy

Consider a simple RD equation of the form:

$$R_{i,t} = \alpha_{1,t} + \beta_t^R E_{i,2007} + f_{1,t}(z_{i,2007}) + \epsilon_{1i,t} \tag{2.1}$$

where $R_{i,t}$ is the R&D expenditure of firm i in year t and $\epsilon_{1i,t}$ is an error term. We use polynomials of the running variable, assets in 2007 $f_{1,t}(z_{i,2007})$, which are allowed to be different either side of the new SME threshold ($\tilde{z}$). $E_{i,2007}$ is a binary indicator equal to one if 2007 assets are less than or equal to the threshold value and zero otherwise. The coefficient of interest β_t^R estimates the effect of being eligible for the more generous SME scheme on a firm's R&D spending at this threshold. In an RD Design, the identification assumption requires that the distribution of all predetermined variables is smooth around the threshold, which is testable on observables. This identification condition is guaranteed when firms cannot precisely manipulate the running variable (Lee, 2008). Under this assumption, eligibility is as good as randomly assigned at the cutoff. We reproduce regressions based on equation (2.1) for year by year outcomes, as well as their average over three post-policy years. We also estimate analogous regressions in the pre-policy years to assess the validity of the RD Design. The "new SME", i.e., those which became SMEs only under the new definition, could only obtain the higher tax deduction rates on R&D performed after August 2008. Hence, to the extent that firms could predict the change in thresholds in early 2008 (or they could manipulate the reported timing of within year R&D), such companies would have an incentive to reduce 2008 R&D expenditures before August and increase them afterwards. To avoid these complexities with the transition year of 2008, we focus on 2009 and afterwards as full post-policy years.

As is standard in RD Designs (Lee and Lemieux, 2010), we control for separate polynomials of the running variable on both sides of the asset threshold of €86m.[19] As noted above, because of the two-year rule, a firm's SME status in 2009 is partly based

[18]Sluggish adjustment to policy announcements is consistent with many papers in the public finance literature (e.g., Kleven and Waseem, 2013).

[19]In the baseline results, being mindful of Gelman and Imbens (2014) warning against using higher order polynomials when higher order coefficients are not significant, we use a first order polynomial. We show in robustness checks that including higher order polynomials produce qualitatively similar results across all specifications.

on its financial information in 2007. Using total assets in 2007 as our primary running variable thus mitigates the concern that there may have been endogenous sorting of firms across the threshold. Figure 2.1 shows that firms' 2007 total asset distribution appears continuous around the new 2008 SME threshold of €86m. The McCrary test gives a discontinuity estimate (log difference in density height at the SME threshold) of -0.026 with a standard error of 0.088, insignificantly different from zero. In later years there is a small but insignificant increase in bunching (subsection 2.7.5).

In terms of innovation outputs, we consider the following RD equation:

$$PAT_{i,t} = \alpha_{2,t} + \beta_t^{PAT} E_{i,2007} + f_{2,t}(z_{i,2007}) + \epsilon_{2i,t} \tag{2.2}$$

where the dependent variable $PAT_{i,t}$ is number of patents filed by firm i in year t, over a longer post-policy period from 2009 to 2015, due to the potential lag between R&D inputs and outputs. Under the same identification assumptions discussed above, $\widehat{\beta^{PAT}}$ consistently estimates the causal effect of being eligible for the more generous SME scheme at the threshold.

Thirdly, we consider the structural patent equation:

$$PAT_{i,t} = \alpha_{3,t} + \gamma_t R_{i,t} + f_{3,t}(z_{i,2007}) + \epsilon_{3i,t} \tag{2.3}$$

which can be interpreted as a "knowledge production function" as in Griliches (1979). Equations (1) and (3) correspond to a fuzzy RD model that estimates the impact of additional R&D spending induced by the difference in tax relief schemes on firm's patents, using $E_{i,2007}$ as the instrument for R&D. With homogeneous treatment effects, the IV estimate delivers the causal effect of R&D on patents under the exclusion restriction that the discontinuity-induced exogenous fluctuations in $E_{i,2007}$ does not affect innovation outputs through any channel other than qualifying R&D.[20]

Under the identification assumptions discussed above, the RD Design guarantees that $E_{i,2007}$ (conditional on appropriate running variable controls) affects innovations only through a firm's eligibility for the SME scheme, which directly translates into qualifying R&D expenditure. It is possible that firms benefitting from the SME scheme (i) also increase complementary non-qualifying spending, such as investments in capital or managerial capabilities (even though they would want to classify as much of this spending as qualifying R&D expenditure as possible), or alternatively (ii) relabel existing non-R&D spending as qualifying R&D expenditure to claim R&D tax relief. The first channel would bias our estimate of γ upward, while the second channel would bias it downward. Empirically, we do not find evidence of discontinuities in firm's capital expenses, (non-R&D) administrative expenses, or any expense category other than qualifying R&D at the eligibility threshold in the post-policy period. This suggests

[20]With heterogeneous treatment effects, IV requires a monotonicity assumption that moving a firm's size slightly below the threshold always increases R&D. In this case, γ is the Average Causal Response (Angrist and Imbens, 1995), a generalization of the Local Average Treatment Effect that averages (with weights) over firms' causal responses of innovation outputs to small changes in R&D spending due to the IV.

that these other channels through which $E_{i,2007}$ could affect innovations and the biases they imply are unlikely to be of first order concern. Relabeling is potentially a harder problem to deal with, but this affects only R&D expenditures and not patenting activity, which is the main outcome variable we focus on.

In Appendix 3.1 we shows how equations (2.1) and (2.3) can be derived from optimizing behavior of a firm with an R&D augmented CES production function and Cobb-Douglas knowledge production function. We discuss how the elasticity of R&D with respect to its user cost can be derived in subsection 2.7.2.

2.4 Data

2.4.1 Data sources

Appendix B.2 details our data sources. Our data comes from three main sources: (1) HMRC Corporate Tax returns (CT600) and its extension, the Research and Development Tax Credits (RDTC) dataset, which provide data on the universe of UK firms and importantly includes firm's R&D expenditures as claimed under the R&D Tax Relief Scheme, (2) Bureau Van Dijk's FAME dataset which provides data on the accounts of the universe of UK incorporated firms, and (3) PATSTAT which has patent information on all patents filed by UK companies in the main 60 patent offices across the world.

CT600 is a confidential administrative panel dataset provided by HMRC Datalab which consists of tax assessments made from the returns for all UK companies liable for corporation tax. The dataset covers financial years 2000 to 2011[21], with close to 16 million firm by year observations, and contains all information provided by firms in their annual corporate tax returns. We are specifically interested in the RDTC dataset, which consists of all information related to the R&D Tax Relief Scheme including the amount of qualifying R&D expenditure each firm has in a year and the scheme under which it makes the claim (SME vs. Large Company Scheme). Firms made 53,000 claims between 2000 and 2011 for a total of £5.8 billion in R&D tax relief, about 80% of the claims are under the SME scheme.

We only observe total R&D when firms seek to claim R&D tax relief. All firms performing R&D are in principle eligible for tax breaks which as we have discussed are generous. Further, all firms must submit tax returns each year and claiming tax relief is a simple part of this process. Hence we believe we have reasonably comprehensive coverage of current R&D spending. Ideally, we would cross check at the firm level with R&D data from other sources, but UK accounting regulations (like the US regulation of privately listed firms) do not insist on SMEs reporting their R&D, so there are many missing values. Statistics provided by internal HMRC analysis indicate that qualifying

[21] The UK fiscal year runs from April 1st to March 31st so 2001-02 refers to data between April 1st, 2001 and March 31st, 2002. In the text we refer to the financial years by their first year, so 2011-12 is "2011".

R&D expenditure amounts to 70% of total business R&D (BERD).[22] Note that the other outcomes (like patents) are observed for all firms, regardless of their R&D status.

CT600 makes it possible to determine the SME status of firms which claim the R&D tax relief, but not the SME status of the vast majority of firms which are not claiming. Employment and total assets are not available because such information is not directly required on corporate tax forms. Furthermore, only tax-accounting sales is reported in CT600, while the SME definition is based on financial-accounting sales as reported in company accounts.[23] Consequently, we turn to a second dataset, FAME, which contains all UK company accounts since about the mid-1980s. We match CT600 to FAME by an HMRC-anonymized version of company registration number (CRN), which is a unique regulatory identifier in both datasets. We merged 95% of CT600 firms between 2006 and 2011 with FAME and these firms cover 100% of R&D performing firms and patenting firms. Unmatched firms are slightly smaller but not statistically different from matched ones across different variables reported in CT600, including sales, gross trading profits, and gross and net corporate tax chargeable (Appendix B.2.4).

All firms are required to report their total assets in company accounts, but reporting of sales and employment is mandatory only for larger firms. In our FAME data, between 2006 and 2011, only 5% of firms reported employment and only 15% reported sales. By comparison 97% reported assets. Even in our baseline sample of relatively larger firms around the SME asset threshold of €86m, employment and sales are still only reported by 55% and 67% of firms respectively. For this reason, we focus on exploiting the SME asset threshold with respect to total assets and use this as the key running variable in our baseline specification. Financial variables are reported in sterling while the SME thresholds are set in euros, so we convert assets and sales using the same tax rules used by HMRC for this purpose. In addition, FAME provides industry, location, capital investment, profits, remuneration and other financial information through to 2013, though coverage differs across variables.

We use assets as our key running variable, although we also experiment with using employment and sales to determine SME status, despite the greater number of missing values. In principle, using both running variables should increase efficiency, but in practice (as we explain in subsection 2.7.6) it does not lead to material gains in the precision of the estimates. Hence, in our main specifications, we use the asset-based criteria for determining eligibility, because it allows us to cover a larger company population and it is clearer for graphical RD presentation.[24]

[22]There are various reasons for this difference, including the fact that BERD includes R&D spending on capital investment whereas qualified R&D does not (only current expenses are liable). It is also the case that HMRC defines R&D more narrowly for tax purposes that BERD which is based on the Frascati definition.

[23]Tax-accounting sales turnover is calculated using the cash-based method, which focuses on actual cash receipts rather than their related sale transactions. Financial-accounting turnover is calculated using the accrual method, which records sale revenues when they are earned, regardless of whether cash from sales has been collected.

[24]It is worth noting that using only one threshold for identification in a multiple threshold policy design does not violate the assumptions for RD Design; it may just reduce the efficiency of the estimates.

Our third dataset is PATSTAT which is the largest available international patent database and covers close to the population of all worldwide patents since the 1900s. It brings together nearly 70 million patent documents from over 60 patent offices, including all of the major offices such as the European Patent Office, the United States Patent and Trademark office (USPTO) and the Japan Patent Office (JPO). Patents filed with the UK Intellectual Property Office are also included. To assign patents to UK-based companies we use the matching between PATSTAT and FAME implemented by Bureau Van Dijk and available from the ORBIS database. Over our sample period, 94% of patents filed in the UK and 96% of patents filed at the EPO have been successfully associated with their owning company. We select all patents filed by UK companies up to 2015. Our dataset contains comprehensive information from the patent record, including application date, citations, and technology class. Importantly, PATSTAT includes information on patent families, which are sets of patents protecting the same invention across several jurisdictions. This allows us to identify all patent applications filed worldwide by UK-based companies and to avoid double-counting inventions that are protected in several countries.[25]

In our baseline results, we use the number of patent families – irrespective of where the patents are filed – as a measure of the number of inventions for which patent protection has been sought. That is, we count the number of patents filed anywhere in the world by firms in our sample, whether at the UK, European or US patent office, but we use information on patent families to ensure that an invention patented in multiple jurisdictions is counted once. Patents are sorted by application year.

Numerous studies have demonstrated a strong link between patenting and firm performance.[26] Nevertheless, patents have their limitations (e.g., Hall et al., 2013). To tackle the problem that the value of individual patents is highly heterogeneous, we using various controls for patent quality such as (i) the number of countries where IP protection is sought (e.g., US and Japan), and (ii) weighting patents by future citations.[27]

2.4.2 Baseline sample descriptive statistics

We construct our baseline sample from the above three datasets. Our baseline sample contains 5,888 firms with total assets in 2007 between €61m and €111m which survive based on a €25m bandwidth around the threshold, with 3,651 firms under the €86m SME asset threshold and 2,327 firms above the threshold. Our choice of bandwidth is guided by results from the Calonico et al. (2014) robust optimal bandwidth approach, yet we still have to decide on one single bandwidth for both R&D and patent outcomes

[25]This means that our dataset includes patents filed by foreign affiliates of UK companies overseas that relate to an invention filed by the UK-based mother company. However, patents filed independently by foreign affiliates of UK companies overseas are not included.

[26]For example, see Hall et al. (2005) on US firms or Blundell et al. (1999) on UK firms.

[27]Variations of these quality measures have been used by *inter alia* Lanjouw et al. (1998), Harhoff et al. (2003), and Hall et al. (2005).

to have a consistent baseline sample.[28] Therefore, we also show robustness to a range of alternative bandwidths. Firms which exit after 2008 are kept in the sample to avoid selection bias, but are given zero R&D and patents. Our key outcome variables include amount of qualifying R&D expenditure, and number of patents filed. All nominal variables are converted to 2007 prices using the UK Consumer Price Index, and all outcome variables are winsorized at 2.5% of non-zero values to mitigate outliers.[29] In 2006-2008, 259 firms in this baseline sample had positive R&D and this number rose to 329 over 2009-2011 (roughly 5% of aggregate R&D expenditure). 172 firms filed 1,127 patents over 2006-2008, and 189 firms filed 1,628 patents over 2009-2013.

Table 2.1 gives some descriptive statistics on the baseline sample. In the 2006-2008 period firms below the threshold spent on average £61,030 per annum on R&D and firms above the threshold spent an average of £93,788 (with an overall average of £73,977). After the policy between 2009 and 2011 these numbers changed to £80,269 and £101,917. In other words, the gap in R&D spending between the two groups of firms around the threshold reduced by more than 30% from £32,758 pre-policy to £21,649 post-policy. In terms of innovation outputs, the average number of patents per annum was similar between the two groups of firms before the policy change (0.061 vs. 0.067), while after the policy change, firms below the SME asset threshold filed around 40% more patents than those above the threshold over 2009-2013 (0.063 vs. 0.044).

These "difference-in-differences" estimates are consistent with our hypothesis that the 2008 policy change induces firms newly eligible for the SME scheme to increase their R&D and patents. The naive difference-in-difference estimates imply unadjusted increases of 15% in R&D and 38% in patents from being below the new SME asset threshold. However, differential time effects across firms of different size would confound these simple comparisons. In particular, recessions are likely to have larger negative effects on smaller firms (which are less likely to survive and are harder hit by credit crunch) than larger firms, which would lead to an underestimate of the positive causal impact of the policy. This is a particular concern in our context as the global financial crisis of 2008-2009 is coincident with the policy change. Even the addition of trends will not resolve the issue because the Great Recession was an unexpected break in trend. However, the RD Design is robust to this problem as it enables us to assume that the impact of the recession is similar around the threshold, whereas the difference-in-difference estimator is not. Consequently, we now turn to implementing the RD Design of equations (1)-(3) to investigate the causal effects directly.

[28] The Calonico et al. (2014) robust optimal bandwidth for using R&D as the outcome variable is 20, and for using patents as the outcome variable is 30. Our baseline bandwidth choice of 25 is in between these two. We also implemented the Imbens and Kalyanaraman (2011) optimal bandwidth approach, which gave similar results.

[29] This is equivalent to winsorizing the R&D of the top 5 to 6 R&D spenders and the number of patents of the top 2 to 4 patenters in the baseline sample each year. We also show robustness to excluding outliers instead of winsorizing outcome variables.

2.5 Results

2.5.1 R&D results

Table 2.2 examines the impact of the policy change on R&D (equation (2.1)). The key explanatory variable is the binary indicator for whether the firm's total assets in 2007 did not exceed the new SME asset threshold of €86m, and the running variable is the firms' total assets in 2007. The baseline sample includes all firms with total assets in 2007 between €61m and €111m, including non-R&D-performers. Looking at each of the two pre-policy years 2006 and 2007 and the transition year 2008 in columns (1)-(3), we find no significant discontinuity in R&D at the threshold. In the next three columns, we observe that from 2009 onward, firms just below the SME threshold have significantly more R&D than firms just above the threshold. Columns (7) and (8) average the three pre-policy/transition and three post-policy years respectively, and column (9) uses the difference between these averages as outcome variable. Although formally, our analysis indicates no pre-policy trends, we consider column (9) a conservative estimate (£60,400). A similar approach is to directly control for pre-policy R&D (the 2006-2008 average) in column (10) which yields a near identical estimate of £63,400 which is significant at the 5% level. These coefficients are not far below the pre-policy average annual R&D of £74,000, suggesting that the policy had a substantial impact from an economic as well as statistical perspective.

Figure 2.2 shows visually the discontinuous jumps in R&D at the SME asset threshold. Unsurprisingly, larger firms with more assets do more R&D as shown by the upward sloping regression lines, but right across the threshold there is a sudden jump in R&D consistent with a policy effect. The magnitude of the jump corresponds to the estimate in column (8) of Table 2.2.

Validity checks. Table 2.3 further examines the validity of our RD Design by looking at the balance of pre-determined covariates. Firms right below and above the threshold are similar to one another in their observable characteristics prior to the policy change. The differences in sales, employment, and capital between the two groups of firms in 2006 and 2007 are both small and statistically insignificant in columns (1) through (6). In column (7), we use a pseudo threshold of €71m with as an upper bound the true threshold of €86m and as a lower bound €46m (€25m below the pseudo-threshold as in the baseline). In column (8), we use a pseudo threshold of €101m with as a lower bound the true threshold of €86m and as an upper bound €116m (€25m above the higher pseudo-threshold). These placebo tests are based on the idea that, after the policy change, we should not observe any discontinuity in R&D around any asset threshold other than the true threshold. Neither placebo tests yield statistically significant effects. We also run similar placebo tests using all possible integer pseudo thresholds between €71m and €101m with a band ranging from €25m below to €25m above the pseudo threshold (we do not truncate the band at the true threshold for these specifications).

Figure B.3, which plots the resulting coefficients and their 95% confidence interval against the corresponding thresholds, shows that the estimated discontinuities in R&D peaks at the true threshold of €86m, while they are almost not statistically different from zero anywhere else.

Our results are robust to a wide range of robustness tests (Table B.3). First, if we add a second order polynomial to the baseline specification of column (8) in Table 2.2, the discontinuity (standard error) is larger at 189.9 (84.7).[30] Second, the results are robust to alternative choices of kernel weights and sample bandwidths.[31] Third, the discontinuity remains significant when we add industry and/or location fixed effects or use different winsorization or trimming rules. Fourth, we obtain statistically significant effects of comparable magnitude when using count data models instead of OLS.[32] Finally, we estimate the same specification as in Table 2.2 but use survival as the dependent variable, finding an insignificant coefficient.

2.5.2 Patent results

We now turn to our results on patents, which is the key outcome of interest. Table 2.4 reports the patent RD regressions (equation (2.2)) using the same specification and sample as Table 2.2. As with R&D, the first three columns show no significant discontinuity around the threshold for patenting activity prior to the policy change. By contrast, there is a significant increase in patenting in the post-policy period from 2009 onward, which persists through to the end of our patent data in 2015, 7 years after the policy change (columns (4)-(10) of Panel A).[33] Although we will focus on the 5 years from 2009 to 2013 (columns (5)-(7) in Panel B) as our baseline "post-policy period" for subsequent patent analyses, the results are qualitatively similar if we use the 2009-2011 average (columns (2)-(4)) or 2009-2015 average (columns (8)-(10)). According to column (5) of Panel B there is an average discontinuity estimate of 0.069 extra patents per year for firms below the policy threshold. The corresponding coefficient for the pre-policy period is less than half the size and statistically insignificant (column (1)). If we use the more-conservative before-after or lagged-dependent variable-specifications, the discontinuity estimates are 0.042 and 0.049 (columns (6) and (7)). Again, these coefficients are sizeable in comparison with the pre-policy mean patents of 0.064. Figure 2.3 illustrates the

[30] Adding a third order polynomial also yields a similar estimate and we cannot reject that the higher order terms are jointly zero.

[31] This include using Epanechnikov or triangular kernel weights, narrower bandwidths of €15m or €20m, or larger bandwidths of €30m or €35m. For larger bandwidths, we (i) add a second order polynomial to improve the fit (the coefficients on the second order assets terms are significant for both bandwidths), or (ii) use triangular kernel weights. All specifications yield statistically significant discontinuity estimates of comparable magnitude to our baseline result in column (8) of Table 2.2.

[32] We do this to allow for a proportionate effect on R&D (as in a semi-log specification). Using a Poisson specification yields coefficient (standard error) of 1.31 (0.49) and using a Negative Binomial specification yields 1.22 (0.49).

[33] These statistically-significant discontinuity estimates decrease in magnitude gradually over time, as 2007 assets becomes a progressively weaker predictor of firm's SME status. Part of this is because firms below the asset threshold in 2007 grow and eventually are no longer SMEs. In Table B.10, we report evidence of substantial policy-induced increase in employment that is consistent with this explanation.

discontinuity in the total number of patents filed over 2009-2013, which corresponds to the estimate in column (5) of Table 2.4 Panel B. As with R&D there is clear evidence of the discontinuity in innovation outcomes at directly the point of the SME threshold for R&D tax relief purpose.[34]

This is a key result: nothing in the R&D tax policy required a firm to show any patenting activity either in applying for R&D grants or in any auditing by the tax authority of how the money is spent. Therefore there is no administrative pressure to increase patenting. It may seem surprising that we observe a response in patenting as soon as 2009, but patent applications are often timed quite closely to research expenditures.[35] It is also possible that firms filed their off-the-shelf inventions when the policy change effectively reduced their patent filing costs. This would translate into a larger estimate in 2009, but cannot explain the persistent effects through 2015. Finally, we ran all the robustness and validity tests discussed for the R&D equation on the patent regressions. These include adding higher order polynomial controls or industry and/or location fixed effects, using alternative choices of kernel weights and sample bandwidths, using different winsorization or trimming rules, employing count data models instead of OLS (Table B.4), and employing pseudo SME thresholds (Figure B.4). The increase in patenting among firms below the SME threshold remains robust across these alternative specifications and peaks only at the true threshold, further confirming the validity of the RD Design and the policy effect on innovation.

As patents vary widely in quality, one important concern is that the additional patents induced by the policy could be of lower value. Table 2.5 investigates this possibility by considering different ways to account for quality. Column (1) reproduces our baseline result of patent counts. Column (2) counts only patents filed in the UK patent office, column (3) has those filed at the European Patent Office (EPO) and column (4) at the USPTO. Since filing at the EPO and USPTO is more expensive than just at the local UK office,[36] these patents are likely to be of higher value. It is clear that there is also a significant and positive effect on these high value patents. Although the coefficient is larger for UK patents, so is the pre-policy mean. Focusing on the relative effect in the final row (the RD coefficient divided by the pre-policy mean of the dependent variable), these are no smaller for EPO and USPTO patents than they are for UK patents (1.03 for UK, 1.20 for EPO and 1.58 for USPTO). We generalize this approach in column (5) by weighting by patent family size, i.e., the total number of jurisdictions in which each invention is patented. This also generates a significant relative effect of around 0.9.

Column (6) of Table 2.5 weights by future citations, which yields a positive and significant estimate. However, we need to keep in mind that our data is very recent

[34]Pseudo threshold tests similar to those in Figure B.3 show that the estimated discontinuities in patents peak at the real SME threshold of €86m and are not statistically different from zero elsewhere (Figure B.4).

[35]See the literature starting with Hall et al. (1986). There are exceptions, of course, where lags between R&D and patenting are longer such as pharmaceuticals.

[36]For example, filing at the EPO costs around €30,000 whereas filing just in the UK costs between €4,000 and €6,000 Roland Berger Market Research (n.d.).

for forward citation count purpose,[37] so the elasticity is less meaningful. To address this issue, we use the number of patents that are in the top citation quartile (in their technology class by filing year cohorts) in column (7). Here we obtain a relative effect of 1.06, very similar to the baseline. Finally, we examine heterogeneity with respect to technology segment looking specifically at chemicals (including biotechnology and pharmaceuticals) in column (8) and Information and Communication Technologies in column (10). These sectors do produce somewhat larger relative effects (both around 1.7 compared to 1.0 in the others), but columns (9) and (10) show that our results are not all driven by these technologically dynamic sectors.

In summary, there is no evidence from Table 2.5 of any major fall in innovation quality due to the policy.[38] The policy appears to robustly raise quality-adjusted patent counts across many measures of patent quality.

2.5.3 IV results for the knowledge production function

Table 2.6 estimates knowledge production functions (IV patents regressions) where the key right-hand-side variable, R&D, is instrumented by the discontinuity at the SME threshold (equation (2.3)).[39] As discussed in Section 2.3, the exclusion restriction, which requires that the instrument affects innovations only through qualifying R&D, is likely to hold in our setting given the lack of evidence of policy effect on other non-qualifying expense categories (Table B.12). Column (1) presents the OLS specification, which shows a positive association between patents and R&D. Column (2) reports a larger IV coefficient implies that one additional patent costs on average \$2.4 million (= 1/0.563 using a \$/£ exchange rate of 1.33) in additional R&D. At the pre-policy means of R&D and patents (£0.074m and 0.064 respectively), this implies an elasticity of patents with respect to R&D of 0.65 for our IV estimates (compared to 0.24 for OLS). If we also control for average pre-policy patents over 2006-2008 as in column (7) of Table 2.4 Panel B, the IV estimate decreases from 0.56 to 0.43 (Table B.5 Panel B) implying an elasticity of 0.50.

The next columns of Table 2.6 compare UK, EPO and US filings. All indicate significant effects of addition R&D on patents, which are again larger for IV than OLS. The corresponding costs for one additional UK, EPO, or USPTO patent are \$2.1, \$4.5, and \$4.0 million respectively (columns (4), (6), and (8)), which reflects the fact that only

[37]Patents are typically published 18 months after the application filing date, and it takes an average of 5 years after the publication date for a patent to receive 50% of its lifetime citations.

[38]We have also looked at many other indicators of quality such as weighting by (i) patent scope (i.e., the number of patent classes a patent is classified into), (ii) the originality index (a measure of how diverse a patent's backward citations are), and (iii) generality index (a measure of how diverse a patent's forward citations are). We also count the number of patents that are in their respective cohorts' top quality quartile as measured by these indices. All of these quality-weighted and top-quality-quartile patent counts yield positive and significant estimates with implied proportionate effects comparable to our baseline patent result (Table B.6 Panel A). Separately, we look at the number of patents subsequently granted (rather than all applications); this similarly yields a positive and signifi-cant estimate.

[39]In the corresponding IV model, the first-stage regression of R&D on the below-asset-threshold instrument is report-ed in column (8) of Table 2.2, and the reduced form regression of patents on the same instrument is reported in column (5) of Table 2.4 Panel B.

inventions of higher value typically get patented outside of the UK.[40] These figures are broadly in line with the existing estimates for R&D costs per patent of $1 to $5 million.[41] We again subject these IV regressions to the robustness tests discussed for R&D and patent regressions to show that the magnitudes are robust (Table B.5).

The fact that the IV effects are larger than OLS is consistent with the LATE interpretation that the IV specification estimates the impact of R&D on patents among complier firms which increase their R&D because of the policy. If these firms are more likely to be financially con-strained they are more likely to have higher-return R&D projects, which they could not have taken without the policy. Some direct evidence for this hypothesis is presented in Table 2.7. We calculate the average cash holdings to capital ratio in each three-digit industry in the pre-policy period using the population of UK firms.[42] All else equal we expect industries with higher cash-to-capital ratios as being less financially constrained. In columns (1) and (4) of Table 2.7, we fully interact all right-hand-side variables in our baseline specification with the industry cash-to-capital measure. The interaction terms indicate that the treatment effects on both R&D and patents are significantly larger for firms in financially constrained sectors. The other columns split sample into industries below and above the mean of the financial constraints measure (instead of using it as a continuous measure) which again show our results are positive and significant only for the firms who are more likely to be financially constrained. [43] In addition, we also calculate the Rajan and Zingales (1998) index of industry external-finance dependence and find qualitatively similar results (Table B.18).

2.6 R&D technology spillovers

The main economic rationale given for more generous tax treatment of R&D is that there are technological externalities, so the social return to R&D exceeds the private return. Our design also allows us to estimate the causal impact of tax policies on R&D spillovers, i.e., innovation activities of firms that are technologically connected to policy-affected firms, through employing a similar RD Design specification with

[40] Despite the weak adjusted first-stage F-statistic of 5.6, the Anderson-Rubin weak-instrument-robust inference tests indicate that all of the IV estimates are statistically different from zero even in the possible case of weak IV.

[41] See Hall and Ziedonis (2001), Arora et al. (2008), Gurmu and Pérez-Sebastián (2008), and Dernis et al. (2015).

[42] This ratio is computed using FAME data for the universe of UK firms between 2000 and 2005. Cash holding is the amount of cash and cash equivalents on the balance sheet; capital is proxied by fixed assets. We first (i) average cash holding and capital within firm over 2000-2005, then (ii) calculate the cash holding to capital ratio at the firm level, and finally (iii) average this ratio across firms by industry. Constructing the measure at the 2-digit and 4-digit industry levels, or using cash flow instead of cash holding, yield qualitatively similar results.

[43] The IV estimate for the effect of R&D on patents (similar to Table 2.6 column (2)) in the subsample of more financial-ly constrained firms is 0.602, significant at 5% level, and larger than the baseline estimate of 0.563. This is con-sistent with our hypothesis that the returns to R&D are higher among more financially constrained firms.

connected firms' patents as the outcome variable of interest (Dahl et al. (2014) for a similar methodological approach in a different context).

For this exercise, we consider two firms to be *technologically connected* if they (i) patent primarily in the same 3-digit technology class and (ii) have a high technological proximity – we use the conventional Jaffe (1986) measure of this.[44] We then construct a sample of all firm i and j dyads ($i \neq j$) in which (a) firm i is within our baseline sample of firms with total assets in 2007 between €61m and €111, and (b) firm j is technologically connected to firm i. Firms i and j are drawn from the universe of UK firms over 2000-2008 where we can construct these measures. There are 203,832 possible such dyads in our data, covering 547 unique firm i's and 17,632 unique firm j's in 91 different technology classes. For ease of exposition, we from now on call firm i the "baseline firm" and firm j the "connected firm."

Our baseline reduced-form spillover specification estimates the impact of base-line firm i's eligibility for the SME scheme based on the asset rule (i.e., being below or at the SME asset threshold) on connected firm j's average patents over 2009-2013:

$$PAT_{j,09-13} = \alpha_4 + \theta E_{i,2007} + f_4(z_{i,2007}) + g_4(z_{i,2007}) + \epsilon_{4ij} \tag{2.4}$$

Each observation is a pair of a baseline firm and a connected firm; $PAT_{j,09-13}$ is the connected firm's average patents over 2009-2013; $E_{i,2007}$ is the baseline firm's threshold indicator in 2007; and $f_4(z_{i,2007})$ and $g_4(z_{i,2007})$ are polynomials of baseline and connected firms' total assets in 2007.[45] As discussed in section 3, $E_{i,2007}$ is as good as random in the RD Design, which allows us to interpret $\hat{\theta}$ as a consistent estimate of the causal impact of baseline firm i's likely-eligibility on connected firm j's innovations.

In addition, we also estimate the following IV specification:

$$PAT_{j,09-13} = \alpha_5 + \xi R_{i,09-11} + f_5(z_{j,2007}) + g_5(z_{j,2007}) + \epsilon_{5ij} \tag{2.5}$$

using $E_{i,2007}$ as the instrument for R&D by baseline firm $R_{i,09-11}$ as in equation (2.3). The exclusion restriction requires that the discontinuity-induced random fluctuations in

[44]Let $F_i = (F_{i1}, \ldots, F_{iY})$ be a 1 x Y vector where $F_{i\tau} = \frac{n_{i\tau}}{n_i}$ is firm i's number of patents in technology class τ as a share of firm i's total number of patents. (i) Firm i's primary technology class is defined as $\tau^* = \arg\max_\tau F_{i\tau}$. (ii) Firms i and j's Jaffe technological proximity is defined as $\omega_{ij} = \frac{F_i F_j'}{[(F_i F_i')^{\frac{1}{2}}(F_j F_j')^{\frac{1}{2}}]}$, the uncentered angular correlation between F_i and F_j. ω_{ij} is equal to 1 if firms i and j have identical patent technology class distribution. It is zero if the firms patent in entirely different technology classes. To avoid picking up policy-endogenous connections Manski (1993), we compute F_i and F_j using the firms' pre-policy patent applications over 1900-2008 (most of these are filed after 1980). Our baseline firms patent primarily in 91 different technology classes, out of the 123 available 3-digit IPC classes. 0.75 is the median Jaffe technological proximity among firm pairs patenting primarily in the same 3-digit technology class, which we use as the cut off for high technological proximity. Using alternative definitions does not affect our qualitative findings.

[45]Conditional on $f_4(z_{i,2007})$, $E_{i,2007}$ is as good as random in the RD Design. It is therefore conditionally uncorrelated with connected firm j's characteristics. As a result, controlling for $g_4(z_{i,2007})$ is not needed for identification, although it helps improve precision as connected firm j's are drawn from a wide support in terms of firm size (as captured by $z_{i,2007}$). Our results are robust to dropping this additional $g_4(z_{i,2007})$ control.

the baseline firm's asset-based eligibility would only affect the connected firm's patents through spillovers from the baseline firm's innovation activities. Under this additional exclusion restriction assumption equation (2.5) consistently estimates the magnitude of the spillovers. Standard errors are clustered by baseline and connected firms' shared primary technology class to address the fact that the residuals may be correlated.

Column (1) of Table 2.8 reports the reduced-form spillover regression using the full sample of baseline firm-connected firm dyads, which yields a small and statistically insignificant coefficient. However, we expect spillovers to have measurable impact only in small-enough technology classes, where a single firm has a good chance of affecting the technological frontier in the field and thus other firms' innovations. This is why Angrist (2014) recommends and Dahl et al. (2014) implement a focus on looking at groups with small numbers of peers when examining spillover effects. Column (2) tests this by fully interacting the terms in equation (2.4) with the size of the dyad's primary technology class. The resulting interaction term is negative and statistically significant at the 5% level, confirming our hypothesis that spillovers are larger in smaller technology classes. Figure 2.5 presents this result visually by plotting the spillover coefficients by the size percentile of the dyad's primary technology class,[46] which yields a downward sloping curve.

Guided by Figure 2.5, we split the full sample of baseline firm-connected firm dyads by the size of the dyad's primary technology class (at 200 which is the 40$^{\text{th}}$ percentile). The subsample of small primary technology classes includes 2,093 dyads of 67 baseline firms and 1,190 connected firms in 36 technology classes. The reduced-form spillover coefficient in this subsample (column (4)) is positive and weakly significant despite the small sample size, and an order of magnitude larger than in the large technology classes in column (3). The presence of positive R&D spillovers on innovations only in small technology classes is robust to a range of robustness tests, including (i) excluding connected firms that are also in the baseline sample to avoid contamination from direct policy effect (column (5)),[47] (ii) extending the definition of technological connectedness to all firms patenting primarily in the same 3-digit technology class (column (6)),[48] and (iii) examining the evolution of spillovers effect over alternative post-policy periods (columns (7)-(8)).[49]

---

[46]This graph is estimated semi-parametrically: the spillover coefficient at each technology class size percentile (the X-axis variable) is obtained from the regression specified in equation (2.4), weighted by a kernel function at that percentile point (Appendix B.3.1).

[47]As $E_{i,2007}$ is as good as random in the RD Design, it is conditionally uncorrelated with whether connected firm j also benefits from the policy. Therefore, technically we do not have to control for possible direct policy effect. Empirically, the spillover point estimate in column (5) is almost the same as the baseline point estimate in column (4).

[48]Relaxing the definition of technological connectedness expectedly results in smaller spillover estimates, even in proportionate terms. More importantly, we observe the same pattern that spillovers are large and significant only in small technology classes (Figure B.7).

[49]The point estimates in columns (7) and (8) indicate that spillovers are positive and significant throughout the post-policy years. They also appear larger when we extend the post-policy period to later years. On the contrary, we find no evidence of spillovers on connected firm j's average pre-policy patents over 2006-2008.

Finally, in the last two columns, we estimate the IV specification using the subsample of small technology classes, with column (9) reporting the first stage and column (10) – the 2SLS. The first-stage effect of the baseline firm's below-asset-threshold indicator on its own post-policy R&D is not statistically significant due to the very small sample of only 67 baseline firms. However, the magnitude of the estimate in relation to the sample pre-policy mean is comparable to our baseline R&D results in Table 2.2, if not larger. Due to a weak first stage, the IV estimate of R&D spillovers in column (9) is also not statistically significant at conventional levels, yet the Anderson-Rubin test, which is robust to weak instruments, rejects the null hypothesis that this estimate is zero at 10% level. In term of magnitude, the spillover estimate is about 40% (= 0.22/0.56) of the direct effect of policy-induced R&D on own patents (Table 2.6).

These findings together provide evidence that policy-induced R&D has a sizable positive impacts on innovation outputs of not only the firms who receive R&D tax subsidies but also other firms in similar technology areas. To our knowledge this paper is the first to provide RD estimates of technology spillovers.

2.7 Extensions and robustness

2.7.1 Intensive versus extensive margins

The additional amount of R&D could come from firms which would not have done any R&D without the policy change (i.e., the extensive margin) or from firms which would have done R&D, although in smaller amounts (i.e., the intensive margin). In Table B.7, we estimate the baseline RD regression with dummies for whether the firm performs R&D or files patent as outcome variables, and find evidence of extensive margin effects only for patent outcomes. Alternatively, we split the baseline sample by firms' pre-policy R&D and patents in Table B.8, and by industry pre-policy patenting intensity in Table B.9. Both exercises show that firms and sectors already engaged in innovation activities have the strongest responses to the policy change. These results provide strong evidence that the policy does not meaningfully affect a firm's selection into R&D performance but works mostly through the intensive margin. In other words, the policy appears to mostly benefit firms that are already performing R&D and filing patents in the pre-policy period, which then helps increase these firms' chances of continuing to have patented innovations in the post-policy period.

We also split the baseline sample into firms which made some capital investments in the pre-policy period, and firms that did not (Table B.11). The policy effects on R&D and patents are larger among firms which had invested, suggesting that current R&D and past capital investments are more likely to be complements than substitutes. This is consistent with the idea that firms having previously made R&D capital investments have lower adjustment costs and therefore respond more to R&D tax incentives (Agrawal et al., 2014).

2.7.2 Magnitudes and tax-price elasticities

What is the implied elasticity of R&D with respect to its tax-adjusted user cost? We define the elasticity as the percentage difference in R&D capital with respect to the percentage difference in the tax-adjusted user cost of R&D.[50] Given the large policy-induced R&D increase in our setting, we calculate the percentage difference relative to the midpoint instead of either end points, following the definition of the arc elasticity measure.[51] Specifically, the tax-price elasticity of R&D η is given by:

$$\eta = \frac{\%\text{ difference in } R}{\%\text{ difference in } \rho} = \frac{\frac{R_{SME}-R_{LCO}}{(R_{SME}+R_{LCO})/2}}{\frac{\rho_{SME}-\rho_{LCO}}{(\rho_{SME}+\rho_{LCO})/2}}$$

where ρ_{SME} and ρ_{LCO} are the firm's tax-adjusted user cost of R&D under the SME and the large companies ("LCO") schemes, and R_{SME} and R_{LCO} are the firm's R&D.[52]

Deriving % difference in R: To obtain estimates of the treatment effects of the difference in tax relief schemes on R&D (i.e, $R_{SME} - R_{LCO}$) and patents, we need to scale $\widehat{\beta^R}$ and $\widehat{\beta^{PAT}}$ by how sharp $E_{i,2007}$ is as an instrument for a firm's eligibility as $E_{i,2007}$ does not perfectly predict firm i's post-policy SME status, $SME_{i,t}$. We estimate this "sharpness" (λ) using the following equation:

$$SME_{i,t} = \alpha_{6,t} + \lambda_t E_{i,2007} + f_{6,t}(z_{i,2007}) + \epsilon_{6i,t} \tag{2.6}$$

Equations (6) and (1) correspond to the first stage and reduced form equations in a fuzzy RD Design that identifies the effect of the change in the tax relief scheme on a firm's R&D at the SME asset threshold, using $E_{i,2007}$ as an instrument for $SME_{i,t}$.

Our setting differs from standard fuzzy RD Designs in that $SME_{i,t}$ is missing for the firms with no R&D (we do not have enough information in our data on sales and employment to determine their eligibility with reasonable precision). Therefore, we can only estimate equation (2.6) on the subsample of R&D performing firms.[53] Selection into this subsample by R&D performance raises the concern whether the resulting $\hat{\lambda}$ is a consistent estimator of the true λ in the full baseline sample which includes non-R&D performers. In Appendix B.1.4 we prove that a sufficient condition for $E(\hat{\lambda}) = \lambda$ is that the SME-scheme eligibility does not increase firm's likelihood of performing R&D compared to being ineligible, which is the case in our setting as shown in subsection

[50] For example, Hall and Jorgenson (1967) or Bloom et al. (2002).

[51] Calculating the percentage difference relative to one end point vs. the other end point yields very different results when the difference between the two points is large. Alternatively, we define the elasticity as the log difference in R&D capital with respect to the log difference in the tax-adjusted user cost of R&D: $\eta = \frac{\ln(R_{SME}/R_{LCO})}{\ln(\rho_{SME}/\rho_{LCO})}$, which yields quantitatively similar elasticity estimates (Table B.15).

[52] Formally, the numerator of the tax price elasticity should be the R&D capital stock rather than flow expenditure. However, in steady state the R&D flow will be equal to R&D stock multiplied by the depreciation rate. Since the depreciation rate is the same for large and small firms around the discontinuity, it cancels out (Appendix B.1).

[53] For the same reason, we cannot directly estimate the corresponding structural equation for the full baseline sample.

2.7.1. In this case the composition of eligible and non-eligible firms below and above the threshold in the estimation sample would be the same as in the full baseline sample. As a result, we are able to derive $\widehat{\beta^R}/\hat{\lambda}$ and $\widehat{\beta^{PAT}}/\hat{\lambda}$, in which $\widehat{\beta^R}$ and $\widehat{\beta^{PAT}}$ are estimated from the full baseline sample and $\hat{\lambda}$ – the R&D performing sample, as consistent estimators of the causal effect of tax policy change on R&D and patents at the threshold. Finally, we retrieve these estimators' empirical distributions and confidence intervals using a bootstrap procedure.

Table 2.9 reports the "first-stage" SME-status RD regressions of equation (2.6) using the same specification as Table 2.2 and the subsample of R&D performing firms in each respective year.[54] Columns (1)-(3) show that being under the new SME asset threshold in 2007 significantly increases the firm's chance of being eligible for the SME scheme in the post-policy years, even though the instrument's predictive power decreases over time as we would expect. Columns (4)-(6) aggregate a firm's SME status over different post-policy periods, which yield coefficients in the range of 0.25 to 0.46 that are all significant at the 1% level. In what follows we will use the mid-range coefficient on SME status of 0.353 (column (5)) as the baseline estimate of λ in equation (2.6). Table 2.2 column (9)'s R&D discontinuity estimate of £60,400 then implies a causal annual treatment effect of £60,400/0.353 = £171,200.[55] Together, these estimates yield a percentage difference in R&D of 1.07.[56]

Deriving % difference in ρ: We calculate the tax-adjusted user cost, ρ_f, based on the actual design of the R&D Tax Relief Scheme (Appendix B.1.5 for more details):

$$\rho_f = \frac{(1 - A_f)}{(1 - \tau_f)}(r + \delta)$$

where sub-script $f \in \{SME, LCO\}$ denotes whether the firm is a smaller (SME) or larger company (LCO), A is the value of R&D tax relief, τ is the effective corporate tax rate, r is the real interest rate, and δ is the depreciation rate. We calculate A separately for the deduction case and the payable credit case under each scheme using the policy parameters and derive the average value of A under each scheme using the probability that a baseline sample firm falls into each case.[57] The resulting average tax-adjusted user

[54]A firm's SME status over a period is the maximum of its SME status in each of the year within the period. We also report elasticity estimates derived from alternative estimates of λ (using different post-policy periods) in Table B.15.

[55]As the tax-adjusted user cost of R&D for large companies remains unchanged over 2006-2011 (Table B.2), it seems reasonable to use the average R&D over 2006-2008 as a proxy for how much an average firm would spend on R&D if it remained a large company over 2009-2011.

[56]That is, $\frac{R_{SME} - R_{LCO}}{(R_{SME} + R_{LCO})/2} = \frac{171.2}{(171.2+74.0+74.0)/2} = 1.07$.

[57]The value of the tax relief (i) in the deduction case is $A_{d,f} = \tau_f(1 + e_f)$ where e_f is the enhancement rate, (ii) in the payable credit case is $A_c = c(1 + e)$ where c is the payable tax credit rate. We use the share of baseline firms with corporate tax liabilities over 2006-2007 as a proxy for the probability that a baseline firm falls into the deduction case.

cost of R&D is 0.15 under the SME scheme and 0.19 under the large company scheme over 2009-2011, which translates into a percentage difference in user cost of 0.27.[58]

Deriving η: Putting the elements together we obtain a tax-price elasticity of R&D of about 4 (= 1.07/0.27), which is substantially higher than the typical values of between one and two found in other studies.[59] Note that Acemoglu and Linn (2004) also find R&D elasticity estimates in the range of 4 with respect to market size and suggest that this should be the same as R&D elasticity with respected to its user cost. In our view, a better way to think of these estimates is to consider the empirical distribution. We perform a bootstrap procedure with 1,000 replications where in each replication, we draw observations with replacement from the baseline sample and calculate the elasticities based on the resulting regression estimates and sample means.[60] Table B.16 Panel A summarizes the results which imply that any R&D tax-price elasticity lower than 1.1 can be rejected with a 5% level of confidence in our setting.

It is worth highlighting that our setting is different from those in previous studies. Most previous studies on R&D tax credits have effectively focused on larger firms such as publicly listed firms or using state/macro data that are dominated by larger firms' expenditures. Our sample, by contrast, is predominantly of smaller firms around the €86m threshold. As we have argued, these firms are more likely to be financially constrained and thus likely to be more responsive to R&D tax incentives. Many recent empirical studies find greater responses of smaller firms to such policy shifts.[61] Another consideration is that the new SME thresholds were introduced in the Global Financial Crisis where *all* firms were more likely to be credit constrained. Although our RD Design is robust to this, this may limit external validity. However, it is worth pointing out that the tax effect on R&D are strong as late as 2011 (and patents as late at 2015), well after the end of the credit crunch.[62]

[58]We set the real interest rate r to 5% and depreciation rate δ to 15%. As $(r+\delta)$ cancels out in the percentage difference between ρ_{SME} and ρ_{LCO}, the values of these parameters do not affect the final tax-price elasticity estimate.

[59]The same calculations yield an elasticity of patents with respect to R&D user cost of 3.6. The patent treatment effect derived from Table 2.4's baseline patent discontinuity estimate of 0.042 (Panel B column (5)) is 0.119 (= 0.042/0.353). This treatment effect and the pre-policy mean patents of 0.064 imply a patent percentage difference of $\frac{PAT_{SME}-PAT_{LCO}}{(PAT_{SME}+PAT_{LCO})/2} = \frac{0.119}{(0.119+0.064+0.064)/2} = 0.96$. This then yields a patent elasticity with respect to R&D tax-adjusted user cost of 3.6 (= 0.96/0.27).

[60]As the first-stage estimate of the effect of firm's below-asset-threshold indicator on its post-policy SME status is based on a smaller sample of 361 R&D performing firms, we separately draw 361 observations from this subsample and 5,527 (=5,888-361) observations from the remaining subsample.

[61]For example, Howell (2017), Zwick and James (2017) and Wallsten (2000) for the US, González et al. (2005) for Spain, Lach (2002) for Israel, and Bronzini and Iachini (2014) for Italy and Görg and Strobl (2007) for Ireland.

[62]Finally, it is also worth noting that we derive the elasticity estimate as $\frac{E(\Delta R_i)}{E(\Delta \rho_i)}$ (instead of $E(\frac{\Delta R_i}{\Delta \rho_i})$ as is standard in the literature), as we do not observe SME_i and implied ρ_i for non-R&D-performing firms. In the sample, it is expected that financially constraint firms have larger elasticity, and are also more likely to experience larger reduction in tax-adjusted user costs of R&D. This positive correlation implies that $|\frac{E(\Delta R_i)}{E(\Delta \rho_i)}| > |E(\frac{\Delta R_i}{\Delta \rho_i})|$.

2.7.3 Cost effectiveness of the R&D Tax Relief Scheme

A full welfare analysis of the R&D policy is complex as one needs to take into account general equilibrium effects through spillovers (Section 2.6) and possibly aggregate effects on scientists' wages (Goolsbee, 1998). We take one step in this direction by implementing a simple "value for money" calculation based on how much additional R&D is generated per pound sterling of taxpayer money ("Exchequer Costs"). The details of the calculations are in Appendix B.1.6, and we only summarize the key results of the analysis here.

Our estimates imply that over 2006-2011, the "value for money ratio" (i.e., the ratio of policy-induced R&D to policy tax payer costs) of the SME deductible scheme is 3.9, SME payable scheme – 2.9, and large company scheme – 1.5.[63] During this period, annually, £302m of Exchequer costs generate £991m additional R&D in the SME scheme, and £660m of Exchequer costs generate £992m additional R&D in the large company scheme. This translates into an aggregate value for money ratio of 2.1.

Figure 2.4 shows estimates of the counterfactual business R&D (BERD) to GDP ratio in the absence of the tax relief scheme. It is striking that since the early 1980s UK BERD became an increasingly small share of GDP, whereas it generally rose in other major economies. Our analysis suggests that this decline would have continued were it not for the introduction and extension of a more generous fiscal regime in the 2000s.[64] Business R&D would have been 13% lower over the 2006-2011 period (total BERD is larger than tax qualifying R&D).

A full welfare analysis could produce even larger benefit to cost ratios. First, since the tax-payer costs are transfers, only the deadweight cost of tax should be considered (e.g., Gruber (2011) uses 40%). Second, the additional R&D is likely to have technological spillovers to other firms, raising their innovation rates (e.g., Bloom et al., 2013) as examined in Section 2.6. On the other hand, there may be general equilibrium effects raising the wages of R&D scientists which would dampen the overall effect.

2.7.4 R&D tax effects on other aspects of firm performance

We examine if the tax policy generated changes in other aspects of firm performance through to 2013 (Table B.13). We again use the baseline specification but use (i) sales, (ii) employment, (iii) capital, and (iv) Total Factor Productivity (TFP) as the outcome variables.[65] Panel B reports sizable, robust, and growing estimates of the policy impact

[63]To be consistent with how policy tax payer costs are reported in HMRC data, we calculate these value for money ratios without accounting for pre-enhancement lost tax revenue from policy-induced R&D. If we also include this amount into tax payer costs, the respective value for money ratios of the three schemes are 2.2, 2.9, and 1.1, and the aggregate value for money ratio of the whole R&D Tax Relief Scheme over 2006-2011 is 1.5. Note that for the SMEs, we use the median elasticity estimates (4.0) in our calculations, and for the large companies, we use the lower-bound elasticity estimates (1.1).

[64]The trend annual decline in business R&D intensity was 0.0190% between 1981 and 1999. We estimate that in the absence of the policy change the decline would have continued at 0.0195% a year 1999 to 2012.

[65]We also include 2-digit industry fixed effects to absorb across-industry heterogeneity in production function. Our results are qualitatively similar without these fixed effects.

on employment over 2009-2013, consistent with a dynamic in which firms increase R&D, then innovate, and then grow larger. In Panel A, estimates are less precise but exhibit similar pattern, suggesting that the policy also have some positive impact on sales. On the other hand, we find little evidence of policy-induced increase in capital (Panel C). This may reflect contemporaneous substitution towards intangible capital (R&D) and away from tangible capital. Finally, we examine if more innovations translate into higher productivity by computing and estimating the policy impact on TFP (Panel D).[66] Similar to Panel A, the resulting estimates, although noisy, are substantially larger in the post-policy years, especially in comparison to the pre-policy estimates of close to zero.

These results should be interpreted with caution. As discussed above, there are many miss-ing values on accounting values of employment and sales as UK accounting regulations do not insist on these being reported for smaller and medium sized enterprises (as in the US). Nevertheless, the results suggest that the policy positively impacts other measures of size and productivity as well as innovation.

2.7.5 Bunching at the threshold in later years

As discussed in Section 2.3, we chose total assets in 2007 as our primary running variable to avoid potential endogenous sorting of firms across the threshold once the policy effective date was announced in 2008. We test the validity of our primary running variable choice and our concern by performing the McCrary test for each year from 2006 to 2011,[67] which estimates the discontinuity in firms' total asset distribution at the SME threshold of €86m. The respective McCrary tests for 2006 and 2007 confirm that firms did not manipulate their total assets to bene-fit from the SME scheme before 2008.[68] On the other hand, there is some graphical evidence of firms' bunching right below the €86m from 2009 onward although it is small and insignificant. Finally, Figure B.5 pools together the two years before the policy change (2006-2007) and Figure B.6 the three years after the change (2009-2011). Endogenous sorting does seem to happen, but only after the policy became effective.[69]

[66]We compute TFP by estimating a production function using the Olley and Pakes (1996) method, based on value added (calculated as sales minus imputed materials), capital, and wages, and at 2-digit industry level. We also compute alternative TFP measures, including Olley-Pakes TFPs based on alternative measures of production inputs and outputs, and Solow-residual TFPs. All measures give qualitatively similar results.

[67]We exclude 2008 as the increase in deduction rate for large companies became effective before the effective date for the changes in the SME scheme (including increase in deduction rate for SMEs and SME definition change) was announced much later in the year. As such, it is hard to predict which way the bunching would happen in this year, or if it would happen at all.

[68]The log differences in density height at the SME threshold in 2006 and 2007 are not statistically different from zero, with coefficients (standard errors) of 0.029 (0.065) and -0.026 (0.088) respectively.

[69]If knowledge production benefits from economy of scale, then firm's attempt to "stay small" to benefit from the SME scheme could lead to an underestimation of the true returns of R&D on patents (and vice versa). However, the small difference in firm size between those right below and above the threshold is unlikely to generate bias large enough to be of first order concern.

2.7.6 Exploiting other elements of the SME definition

We also ran RD regressions using other elements of the SME definition (sales and employment) to estimate the impacts of the policy (Table B.14). We must interpret these results with caution because, as noted above, there are many missing values on sales and especially employment. Furthermore, we also find evidence that the asset criterion is more binding than the sales one. A firm is considered an SME if it meets either one of the criteria, thus the asset criterion is binding only when the firm already fails the sales one and vice versa.[70]

As expected, while we still find positive effects on R&D and innovation outputs using the sales or employment criterion, these effects are not always statistically significant. They are also of smaller magnitude compared to our baseline effects estimated using the asset criterion when taking into consideration the baseline pre-policy R&D and patent means of the respective sample. The discontinuity estimate to pre-policy mean ratios for R&D using asset, sales, and employment criteria are 1.67, 1.16, and 0.41 respectively, and the same set of ratios for patents are 1.09, 0.31, and 0.41.[71] We also examined whether combining the different SME criteria could increase the efficiency of our estimates, but found no significant improvement.[72]

2.8 Conclusion

Fiscal incentives for R&D have become an increasingly popular policy of supporting innovation across the world. But little is known about whether these costly tax breaks causally raise innovation. We address this issue by exploiting a change in the UK R&D tax regime in 2008 which raised the size threshold determining whether a firm was eligible for the more generous SME tax relief scheme. This enables us to implement a RD Design and assess impact of the policy on R&D and innovation (as measured by patenting). Using total assets in the pre-policy period of 2007 we show that there is no evidence of discontinuities around the threshold prior to the policy, which is unsurprising as the new threshold is only relevant for the R&D Tax Relief Scheme and not for other programs targeting SMEs.

[70] The binding/non-binding ratio (i.e., the number firms for which the criterion binds divided by the number of firms for which the criterion does not bind) for the *asset* criterion is 0.36, while the same ratio for the *sales* criterion is only 0.20 (Appendix B.2.6 for further details).

[71] Even when we restrict the sample to firms for which the sales criterion binds when using the sales running variable, the proportionate effects are still lower than our baseline results, and the estimates are not statistically significant.

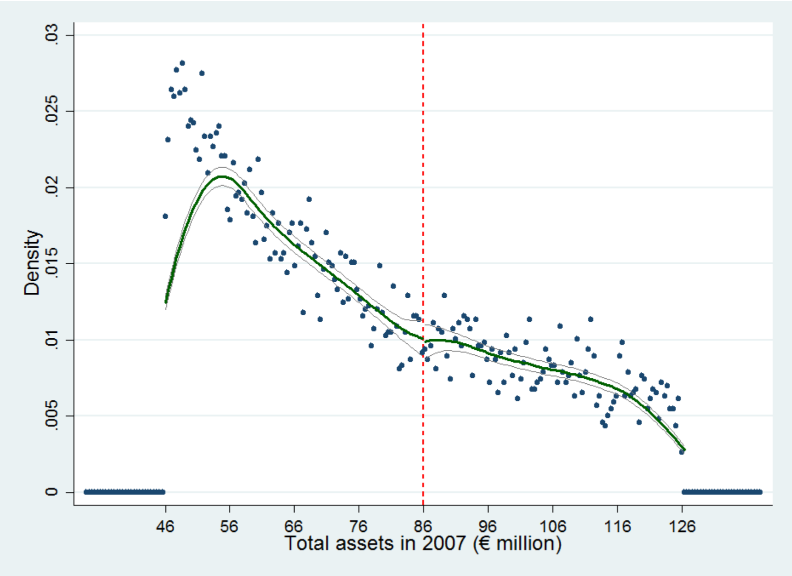
[72] The *below-asset-threshold indicator* almost always generates large and statistically significant effects on both R&D and patents, while the *below-sales-threshold indicator* does not (Table B.10 Panel B). This is consistent with the observation that the asset criterion is more binding and therefore the *below-asset-threshold indicator* is a more precise instrument for firm's SME status. Joint F-statistics for *below-asset-threshold* and *below-sales-threshold indicators* indicate that their effects on R&D and patents are always jointly significant. Finally, the IV estimates for R&D effect on patents using both criteria as instrumental variables for R&D are similar to our baseline. However, they are less precise due to the inclusion of an additional weak *below-sales-threshold indicator* instrument.

The policy caused an economically and statistically significant increase in R&D and patent-ing (even after quality-adjusting). Furthermore, the tax policy appears to stimulate positive technology spillovers. This suggests that R&D tax policies do seem effective in increasing innovation, and not simply devices for relabeling existing spending or shifting innovative activities between firms. The implied elasticities of patenting and R&D with respect to changes in its (tax adjusted) user cost are large. We argue that the R&D elasticity is large compared to existing estimates because we focus on firms that are smaller, and therefore more likely to be subject to financial constraints, than those conventionally used in the extant literature. Over the 2006-2011 period we calculate that the tax relief scheme meant aggregate business R&D was 13% higher than it would otherwise have been, halting the secular decline of the UK's share of business R&D in GDP.

There are many caveats when moving from these results to policy. Although the results are optimistic about the efficacy of tax incentives, the large effects come from smaller firms and should not be generalized across the entire size distribution – this does imply that targeting R&D policy on financially constrained SMEs is worthwhile (although a first best policy would be to deal directly with credit market imperfections). Furthermore, our estimates are based on the period after the global financial crisis when credit frictions may have been particularly acute. However, the fact that the impact is also large seven years after the crisis period suggest that this should not be overstated.

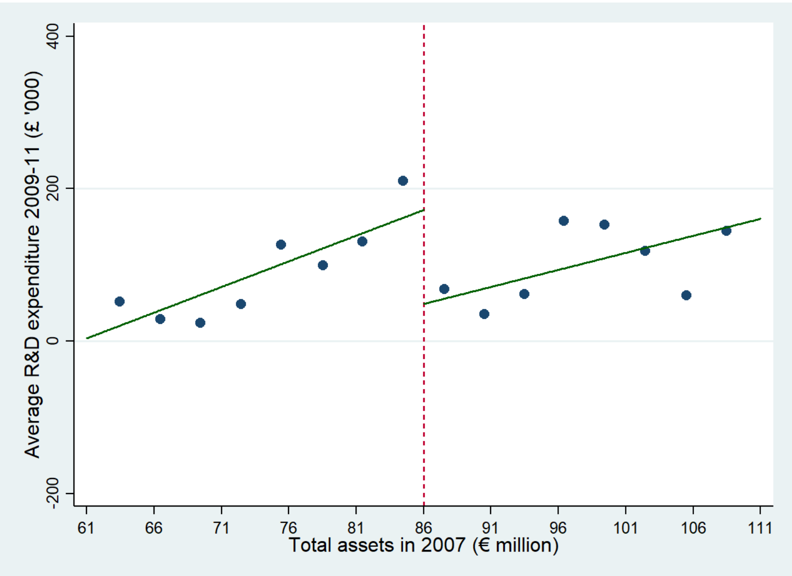
We have partially examined general equilibrium effects by demonstrating that the R&D tax policy stimulates patenting activity not only for the firms directly benefited, but also creates spillovers for other firms which were indirectly affected. However, there may be other equilibrium effects that reduce innovation. For example, subsidies are captured in the form of higher wages rather than a higher volume of R&D, especially in the short-run. We believe that this is less likely to be a first order problem when there is large international mobility of inventors, as is the case in the UK (e.g., Akcigit et al., 2016; and within the US see Moretti and Wilson, 2017). Furthermore, the policy's effect on patenting implies that the increase in R&D is driven by volume and not just wages.

Figure 2.1: McCrary test for no manipulation at the SME asset threshold in 2007



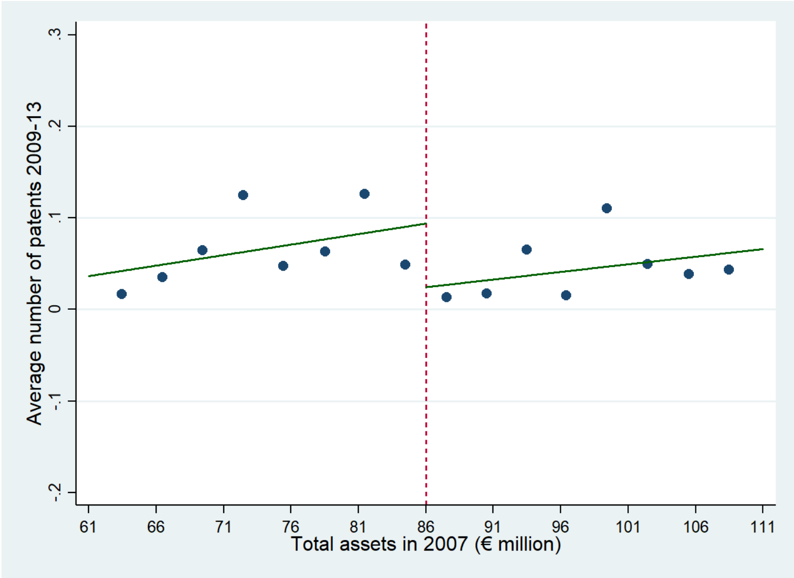
Notes: McCrary test for discontinuity in distribution density of total assets in 2007 at the SME asset threshold of €86m. Sample includes firms with total assets in 2007 between €46m and €126m. The discontinuity estimate (log difference in density height at the SME threshold) is -0.026, with standard error of 0.088.

Figure 2.2: Discontinuity in average R&D expenditure over 2009-2011



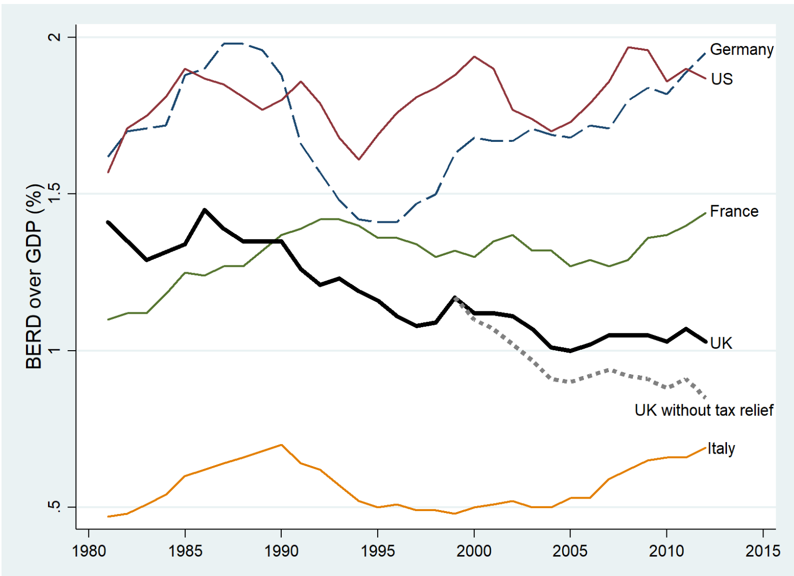
Notes: The figure corresponds to the baseline R&D regression based on equation (2.1), using an OLS Regression Discontinuity (RD) Design. The dependent variable is average R&D expenditure over 2009-2011. The running variable is total assets in 2007 with a threshold of €86m. The baseline sample includes firms with total assets in 2007 €25m above and €25m below the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of the running variable separately for each side of the threshold are included. The OLS discontinuity estimate at the €86m threshold is 123.2 with a standard error of 52.0. Each point represents a bin of 368 firms on average, over a range of €3m. (Bin size is large due to data confidentiality requirement.)

Figure 2.3: Discontinuity in average number of patents over 2009-2013



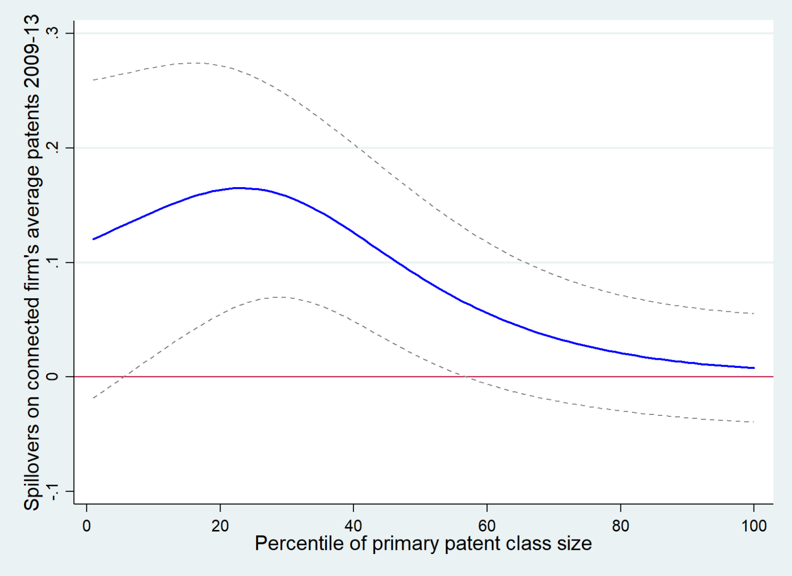
Notes: The figure corresponds to the baseline patent regression based on equation (2.3), using an OLS Regression Discontinuity (RD) Design. The dependent variable is average number of patents over 2009-2013. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 €25m above and below the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of the running variable separately for each side of the threshold are included. The OLS discontinuity estimate at the €86m threshold is 0.069 with a standard error of 0.026. Each point represents a bin of 368 firms on average, over a range of €3m. (Bin size is large due to data confidentiality requirement.)

Figure 2.4: Business Enterprise R&D over GDP, selected countries



Notes: The data is from OECD MSTI downloaded February 9th, 2016. The dotted line ("UK without tax relief") is the counterfactual R&D intensity in the UK that we estimate in the absence of the R&D Tax Relief Scheme (sub-section 2.7.3 and Appendix B.1.6 for details).

Figure 2.5: Spillovers on connected firm's patents by primary patent class size



Notes: This figure presents semi-parametric estimates of the spillover coefficient on technologically-connected firm's patents as a function of the technology class size percentile (the X-axis variable). The semiparametric estimation is based on equation (2.4), using a Gaussian kernel function of the X-axis variable and a bandwidth of 20% of the range (Appendix C.1 for details). The 40th percentile of technology class size is 200. The dashed lines indicate the 90% confidence interval for the spillover coefficients.

Table 2.1: Baseline sample descriptive statistics

Subsample	**Firms with 2007 total assets between €61m and €86m**			**Firms with 2007 total assets between €86m and €111m**			**Difference between two subsamples**		
Year	2006-08 average	2009-11 average	2009-13 average	2006-08 average	2009-11 average	2009-13 average	2006-08 average	2009-11 average	2009-13 average
Total no. of firms in subsample		3,561			2,327			1,234	
No. of R&D performing firms	160	210		99	119		61	91	
No. of patenting firms	105	104	120	67	57	69	38	47	51
Mean R&D expenditure *(£)*	61,030	80,269		93,788	101,917		-32,758	-21,649	
Mean patent applications (family)	0.061	0.064	0.063	0.067	0.047	0.044	-0.006	0.017	0.018
Mean EPO patent applications	0.078	0.070	0.069	0.074	0.053	0.051	0.004	0.017	0.018
Mean UK patent applications	0.031	0.030	0.030	0.028	0.024	0.024	0.003	0.006	0.006
Mean US patent applications	0.026	0.028	0.028	0.024	0.025	0.025	0.002	0.003	0.003

Note: The baseline sample includes 5,888 firms with total assets in 2007 between €61m and €111m. Total assets are from FAME and are converted to € from £ using HMRC rules. Qualifying R&D expenditure comes from CT600 panel dataset and are converted to 2007 prices. Patent counts come from PATSTAT.

Table 2.2: R&D regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**R&D expenditure (£ *'000*)**									
	Before (pre-policy)			After (post-policy)			Before	3yr After	3yr Diff.	LDV
Year	2006	2007	2008	2009	2010	2011	2006-08 average	2009-11 average	3yr After - Before	2009-11 average
Below-asset-threshold indicator (in 2007)	43.4	81.9	63.1	97.3*	133. 6**	138.9**	62.8	123.3**	60.4*	63.4**
	(50.6)	(59.2)	(44.9)	(51.4)	(53.5)	(55.1)	(48.9)	(52.1)	(31.5)	(32.1)
Past R&D exp. (*£'000*), 2006-08 average										0.95***
										(0.08)
Firms	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of the running variable separately for each side of the threshold are included. Robust standard errors are in brackets. Mean R&D expenditure between 2006 and 2008 was £73,977 and between 2009 and 2011 was £88,824. 2007 real prices.

Table 2.3: Pre-treatment covariate balance tests and placebo tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	**Ln(Sales)**		**Ln(Employment)**		**Ln(Capital)**		**R&D exp. *(£ ’000)***	
Year	2006	2007	2006	2007	2006	2007	2009-11 average	
Below-asset-threshold indicator (in 2007)	-0.124	0.086	0.117	0.157	0.023	-0.006	-8.0	53.1
	(0.162)	(0.161)	(0.135)	(0.131)	(0.112)	(0.103)	(38.0)	(75.1)
SME threshold (€)	86m	86m	86m	86m	86m	86m	71m	101m
Sample bandwidth	61-111m	61-111m	61-111m	61-111m	61-111m	61-111m	46-86m	86-126m
Firms	4,155	4,348	2,973	3,089	4,766	5,078	7,095	3,354

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of the running variable separately for each side of the threshold are included. Robust standard errors are in brackets. Columns (1)-(6) report pre-treatment covariate tests for sales, employment, and capital. Columns (7) and (8) report placebo tests using placebo asset threshold of €71m and €101m.

Table 2.4: Reduced-form patent regressions

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**All patent family count**									
	Before (pre-policy)			After (post-policy)						
Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Below-asset-threshold indicator (in 2007)	0.002	0.036	0.044	0.095***	0.070**	0.073**	0.050**	0.059*	0.059**	0.047*
	(0.035)	(0.034)	(0.033)	(0.034)	(0.031)	(0.034)	(0.024)	(0.030)	(0.023)	(0.023)
Firms	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**All patent family count**									
	Before	3 years After			5 years After			7 years After		
Year	2006-08 average	2009-11 average	3yr After - Before	2009-11 average	2009-13 average	5yr After - Before	2009-13 average	2009-15 average	7yr After - Before	2009-15 average
Below-asset-threshold indicator (in 2007)	0.028	0.079***	0.052**	0.057**	0.069***	0.042*	0.049**	0.065***	0.037*	0.046**
	(0.030)	(0.030)	(0.023)	(0.022)	(0.026)	(0.022)	(0.020)	(0.024)	(0.022)	(0.019)
Past patent family count, 2006-08 average				0.818***			0.729***			0.670***
				(0.107)			(0.106)			(0.106)
Firms	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firm's total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of the running variable separately for each side of the threshold are included. Robust standard errors are in brackets. Mean all patent family count between 2006 and 2008 was 0.064, and between 2009 and 2011 was 0.057, between 2009 and 2013 was 0.055, and between 2009 and 2015 was 0.52.

Table 2.5: Effects of R&D tax relief on quality-adjusted patents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Dependent variable (2009-13 average)	Baseline	UK patents	EPO patents	US patents	Family size (countries)	Patent citations	Patents in top citation quartile	Chem/ pharma patents	Non-chem/ pharma patents	ICT patents	Non-ICT patents
Mean of dependent variable (2006-08)	*0.064*	*0.076*	*0.030*	*0.026*	*0.254*	*0.292*	*0.031*	*0.009*	*0.050*	*0.003*	*0.059*
Below-asset-threshold indicator (in 2007)	0.069***	0.078**	0.036**	0.041**	0.218**	0.133**	0.033**	0.0149*	0.049**	0.005	0.058**
	(0.026)	(0.031)	(0.016)	(0.016)	(0.108)	(0.067)	(0.013)	(0.008)	(0.021)	(0.003)	(0.024)
elasticity (estimate divided by mean of dependent variable)	*1.08*	*1.03*	*1.20*	*1.58*	*0.86*	*0.46*	*1.06*	*1.66*	*0.98*	*1.67*	*0.98*
Firms	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of the running variable separately for each side of the threshold are included. Robust standard errors are in brackets. Quality measures are baseline patent family count (column 1), EPO patent count (column 2), UK patent count (column 3), US patent count (column 4), patent by family size count (i.e., patent by country count) (column 5), patent by citation count (column 6), patent count in the top 25% in citation count of their patent class x year cohort (column 7), chemistry/pharmaceutical patent count (column 8), non-chemistry/pharmaceutical patent count (column 9), ICT patent count (column 10), and non-ICT patent count (column 11). Chemistry/pharmaceutical patents include all patents classified into patent sector (3) Chemistry. Information and communication technology (ICT) patents include all patents classified into either patent field (4) Digital communication, (6) Computer technology, or (7) IT methods for management.

Table 2.6: Effects of R&D on patents (IV regressions)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable (2009-13 average)	**All patent family count**		**UK patent count**		**EPO patent count**		**US patent count**	
Specification	OLS	IV	OLS	IV	OLS	IV	OLS	IV
R&D expenditure *(£ million)*, 2009-11 average	0.206***	0.563**	0.231***	0.629*	0.122***	0.293*	0.121***	0.330**
	(0.070)	(0.282)	(0.084)	(0.328)	(0.046)	(0.153)	(0.043)	(0.166)
Anderson-Rubin test p-value		0.008		0.012		0.025		0.012
Firms	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888

Note: *** significant at 1% level, ** 5% level, * 10% level. Instrumental variable is the indicator whether total assets in 2007 is below €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of the running variable (total assets in 2007) separately for each side of the threshold are included. Robust standard errors are in brackets. Adjusted first-stage F-statistic is 5.6. P-values of Anderson-Rubin weak-instrument-robust inference tests indicate that the IV estimates are statistically different from zero even in the possible case of weak IV.

Table 2.7: Heterogeneous effects of R&D tax relief by financial constraints

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	**R&D expenditure *(£ '000)*, 2009-11 average**			**All patent family count, 2009-13 average**		
Sample	Full	Low Cash/K	High Cash/K	Full	Low Cash/K	High Cash/K
Below-asset-threshold indicator (in 2007)	157.8**	286.6**	-17.8	0.104***	0.171***	-0.003
	(70.6)	(112.0)	(31.4)	(0.040)	(0.064)	(0.011)
Below-asset-threshold indicator # Cash/K	-13.6*			-0.011***		
	(7.7)			(0.004)		
Difference		304.4*** (116.3)			0.174** (0.065)	
Firms	4,504	2,237	2,267	4,504	2,237	2,267

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of the running variable separately for each side of the threshold are included. Robust standard errors are in brackets. Cash/K is calculated as the 3-digit industry average of firms' cash and cash equivalents holding as the share of capital over 2000-2005. Firms in industries with low Cash/K measure are more likely to be financially constrained. Low (high) Cash/K subsample includes firms with below (above) median industry Cash/K measure. All right-hand-side variables are fully interacted with industry Cash/K measure in columns (1) and (4).

Table 2.8: R&D technology spillovers on patents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Specification	Reduced form								First stage	IV
Dependent variable	**Connected firm's all patent family count**								**Baseline firm's R&D**	**Connected firm's all patents**
Year	2009-13 average				09-13 avg.	09-13 avg.	09-11 avg.	09-15 avg.	09-11 avg.	09-13 avg.
Sample	Full	Full	Large tech. class	Small tech. class	Small tech. class; excl. baseline	Small tech. class; broad def.	Small tech. class	Small tech. class	Small tech. class	Small tech. class
Baseline firm's below-asset-threshold indicator (in 2007)	0.019 (0.029)	0.067 (0.042)	0.018 (0.029)	0.196* (0.097)	0.192* (0.103)	0.109** (0.049)	0.170* (0.090)	0.198* (0.099)	0.884 (0.633)	
Baseline firm's below-threshold indicator # tech. class size *('000)*		-0.029** (0.012)								
Baseline firm's R&D expenditure *(£ million)*, 2009-11 average										0.222 (0.166)
Difference			0.178* (0.094)							
Anderson-Rubin test p-value										0.051
Dependent variable mean over 2006-08	*0.396*	*0.396*	*0.397*	*0.291*	*0.294*	*0.203*	*0.291*	*0.291*	*0.248*	*0.291*
No. of connected firms	17,632	17,632	16,477	1,190	1,152	2,340	1,190	1,190	1,190	1,190
No. of baseline firms	547	547	487	67	67	78	67	67	67	67
No. of 3-digit IPC classes	91	91	55	36	36	36	36	36	36	36
Observations	203,832	203,832	201,739	2,093	2,015	5,708	2,093	2,093	2,093	2,093

Note: *** Significant at 1% level, ** 5% level, * 10% level. Sample of technologically-connected pairs of a baseline firm and a connected firm. Two firms are technologically connected if they (i) patent primarily in the same 3-digit technology class, and (ii) have Jaffe technologically proximity of at least 0.75 (except for column (6), in which only (i) applies), computed using all patent applications over 1900-2008. Baseline firms include patenting firms (before 2008) with total assets in 2007 between €61m and €111m. Connected firms include the universe of patenting firms (before 2008). OLS estimates based on the RD Design. The running variable is baseline firm's total assets in 2007 with a threshold of €86m. Controls for (i) first order polynomials of the running variable separately for each side of the threshold and (ii) second order polynomial of connected firm's total assets in 2007 are included. Standard errors in brackets are clustered by baseline and connected firms' shared primary technology class. Technology class size is the number of firms whose primary technology class is the said class. Small (large) technology class subsample includes firms whose primary technology classes are below (above) 200 in size (technology class size's 40th percentile). In column (2), all right-hand-side variables are fully interacted with technology class size (in thousands).

Table 2.9: SME status regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	**Indicator: Has R&D claims under SME Scheme**					
Year	2009	2010	2011	2008-2009	2008-2011	2009-2011
Below-asset-threshold indicator (in 2007)	0.326***	0.301***	0.184*	0.464***	0.353***	0.248***
	(0.085)	(0.089)	(0.100)	(0.087)	(0.090)	(0.093)
Firms	215	218	248	265	361	333

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of the running variable separately for each side of the threshold are included. Robust standard errors are in brackets. The sample for a certain year (period) effectively includes firms in the baseline sample with R&D tax relief claims in that year (period). A firm's SME status over a period is the maximum of its SME status in each of the year within the period.

Chapter 3

Evidence from Vietnam

This chapter shows that from 2001 to 2008, the gradual privatization program in Vietnam has strong positive spillovers on firms' total factor productivity through backward linkages. A firm gains on average 4 percentage points in TFP if private firms in all of its downstream industries increase their market share by 10%. This effect comes mainly from privatized domestic firms within local markets. It is stronger for upstream industries under more competition from imports, and weaker for those that export more. The effect on allocative efficiency is imprecisely negative and mitigates the overall effect on industry TFP index, as new entrants but not incumbents capture very large gains in TFP. I present evidence that the effect works through elevated pressure from privatized client firms. Furthermore, the spillovers are stronger when provincial institutions favor state-owned firms and create higher entry costs, suggesting that good governance and privatization spillovers are substitutes.

3.1 Introduction

Ever since Plato and Aristotle, ownership by the state or by private entities has always been a central topic in debates regarding economic systems, public policies and their consequences. In recent decades, thousands of state-owned enterprises (SOEs) have been privatized in both developed and developing countries, most often guided by economic theories that predict higher productivity and efficiency under private ownership. However, there are large gaps between theory and empirical evidence on the effects of privatization. Recent empirical findings have found a wide range of estimates of privatization benefits (Estrin et al., 2009), as not all privatization campaigns have lived up to the projected progress in productivity, thus fueling debates among economists and policy makers on the consequences of privatization. This paper contributes to those debates with the novel suggestion that privatization's effects can be strongly advantageous when one looks for benefits beyond the privatized firms, into their upstream and downstream industries. With empirical evidence that privatization yields large spillovers on upstream firms' productivity, the paper supplies a strong reason in favor of privatization because of its macroeconomic advantages.

The foremost arguments for or against privatization have always been its direct impacts on privatized firms and related factors such as excessive labor or performance. Consequently, most economists and policy makers study and debate on the potential direct impacts. Little is asked and answered, however, on the spillovers of privatization on other firms in the same and in other industries. Yet, this may be key to understanding the overall success or failure of large privatization campaigns seen in transition economies since the 1990s.

While SOEs can be found almost everywhere in the world, the importance of privatization has been mostly stressed among transition economies, where SOEs once were the sole key drivers of the economy, and where many countries are still struggling to find consistent growth. Understanding privatization effects, especially at the macro level, is particularly important in those countries. In the 1990s, transition policies in Eastern Europe and former Soviet republics were mostly radical, resulting in sharp increases of the share of private firms from almost zero to at least half the economy over a couple of years (Table 3.1, Estrin et al., 2009). The rich political and economic environment surrounding those policies also carry along many events at the same time, such as the introduction of free elections and labor market liberalization, making it hard to disentangle the effects of privatization, especially when one wants to consider cross-industry spillover. Vietnam is among the very few countries that at some point followed tightly Soviet-style planning before market liberalization, and yet chose to undergo a gradual pattern of privatization, where private and state-owned firms coexist for an extended period of time in an economy where most aspects (such as salary and employment) had been liberalized before. Over the period from 2001 to 2008, the share of private firms in manufacturing has slowly grown from to 66.1% to 86.8%, or a modest but gradual gain of less than 3% a year. It has happened throughout most manufacturing industries, but at slow, varying speeds, so that one could use such variation to estimate privatization effects on firms and industries' productivity.

Because privatized firms are usually not chosen randomly, existing studies of the direct effect of privatization on firm performance have spent much effort on controlling for selection issues, usually with fixed effects or instrumental variables (as noted by Estrin et al., 2009). By studying how upstream and downstream privatization leads to firm performances, I shift the focus to indirect effects of privatization through linkage spillover. I thus avoid the most thorny issue of endogeneity, by which individual firms' specific shocks may partially cause both privatization and performance changes. Instead, my identification assumption only restricts that there be no correlated shocks that have influences on downstream (or upstream) privatization and also same-industry performance. I will further argue that if such shocks exist, they will have biased the results against my findings.

To understand privatization's indirect impacts on productivity, I apply Olley and Pakes's (1996) (OP) structural estimation to estimate individual firms' total factor productivity (TFP) for all manufacturing firms in the Vietnam Enterprise Census of all firms with 10 employees or more. I then decompose each industry's TFP index into an unweighted average TFP and a covariance term measuring allocative efficiency. This decomposition is used to study privatization's impacts on all components of the overall TFP index.

Firm-level regressions of TFP on privatization measures show a strong, robust effect of downstream privatization. TFP increases by 4 percentage points when downstream privatization increases by 10%. The effect comes from increases in domestic private, rather than foreign ownership, and is only in place for private firms' TFP.

As a whole, the manufacturing sector in Vietnam, represented by 43,545 firms in my sample, has enjoyed substantial growth of TFP during the period 2001 to 2008. The annual TFP growth of 2.6% over this period is half explained by unweighted average TFP growth, and half by increased allocative efficiency. Downstream privatization has a very strong positive impact on unweighted average TFP. On the other hand, its effect on allocative efficiency is negative, although not significant, resulting in an overall effect on industrial TFP index that is negligible. Further decomposition shows that the effect comes mostly from largely improved TFPs among entrants, while incumbent firms are not significantly affected.

I investigate the potential channels of influence of downstream privatization on TFP. I control for proxies of demand and industrial competition, and find that the effect is dominated by privatization of downstream firms in the same local market. Because the effect works through entrants rather than incumbent firms, it is unlikely to be technology spillovers via learning by doing. Instead, the evidence points to an explanation that privatized downstream industries raise the pressure on suppliers, thus pushing suppliers, especially entrants, to improve their productivity if they want to enter and survive in the market.

I ask whether this privatization effect is complement or substitute with institution quality regarding SOEs, private firms and especially entrants. I find that they are substitutes. In good-governance provinces, where SOEs are treated with less favor, and/or barrier to entry is lower, the privatization effect is smaller. Understandably, when SOEs are not treated differently, they tend to resemble private firms more.

This paper resonates from a line of research in the trade economics literature that studies spillover effects of FDIs through linkages in the host country. Javorcik (2004) finds sizeable linkage effects of FDI on TFP in Lithuania, and Arnold et al. (2011) finds linkage effects of service liberalization on downstream manufacturing TFP in the Czech Republic, which work mostly through entry in downstream industries. Gorodnichenko et al. (2010) looks at a more comprehensive sample of surveyed firms across 27 emerging markets and finds a positive relationship between same-industry foreign competition and firm innovations, probably because of higher pressure from competition. However, this literature has looked only at the effects on individual firms, and overlooked the effects on different components of the TFP index. I do find important effects on individual firms, but also point out the much mitigated effect on overall industry TFP, because of the allocative efficiency component.

Several economic theories have examined the channels through which firm TFP may improve in those situations. Notably, Aghion et al. (2005) and Aghion et al. (2009) model how competitive pressure, either from among existing firms in the same industry, or from the threat of entry, shapes firms' behaviors in improving productivity. Rodriguez-Clare's (1996) international trade model shows that a multinational's entry, under sufficient complementarity with domestic suppliers, could spur growth in the upstream industry, thanks to larger demands coming from the international market. Lin and Saggi (2007) proposes a different channel of linkage effects via the supply side, where foreign firms could contract to transfer new technology to domestic suppliers. Sutton (2007) also models vertical transfers of capabilities from foreign firms to domestic firms, because of quality requirement on the international market.

Looking beyond vertical linkages, the paper contributes to a larger literature on the determinants of productivity within and across firms Syverson (2011). The paper also connects to the recent literature that emphasizes the large role of allocative efficiency, notably since Hsieh and Klenow (2009) and Bartelsman et al. (2013), with for example an application to China Brandt et al. (2012, 2013). Once again, I find that allocative efficiency is important in understanding overall TFP, in that the positive effect of privatization on individual firms could be offset by that on allocative efficiency.

There has been a large literature that addresses the causes and consequences of privatization in transition economies, surveyed in Megginson and Netter (2001), Djankov and Murrell (2002), Estrin et al. (2009). In tandem with the rising interests in the Chinese economy (e.g. Song et al., 2011), this paper belongs to a few that have studied Vietnam's recent grow phenomenon using the Vietnam Enterprise Census Bai et al. (2013); McCaig and Pavcnik (2012); Tran (2013). It shows similarities between Vietnam and China when it comes to firm and industry characteristics Brandt et al. (2012), suggesting that results from Vietnam may be generalizable to China and other transition contexts too.

The paper is organized as follows. Section 3.2 presents a background discussion on privatization in Vietnam. Section 3.3 details all methodologies used in the paper, for which an uninterested reader could skip. Section 3.5 exposes the results, and section 3.6 discusses the key implications and caveats, suggests future work, and concludes.

3.2 Background of privatization in Vietnam

Before the mid-80s, Vietnam followed a central-planning economic model based on a large public sector with thousands of state-owned enterprises (SOEs). In 1986, the Vietnamese government launched a market liberalization Reform, known as *Doi Moi*, to eliminate the role of central-planning and open the country's borders to international capital and trade flows. Over the past 20 years, Vietnam has achieved a remarkable average annual growth rate of 7.3%, making it one of the fastest growing economies in the world over this period. Based on my calculation, Vietnam's manufacturing total factor productivity (TFP) also grew at 2.6% per year between 2001 and 2008, which is on par with TFP growth in China and high in comparison with other Asian countries Brandt et al. (2012).

The reform of SOEs plays a key role in this comprehensive economic renovation. A key problem SOEs face is their multiple, ambiguous, and conflicting goals, which constrain the SOEs' managerial autonomy. Under this system, managers and workers are often poorly motivated, leading to weak business performance, low efficiency, and operating losses. The first SOE reform in Vietnam provided these firms with more managerial autonomy to make business plans and set product prices. However, it ended up causing many SOEs to fall in the red as government subsidies were also cut.

In 1992, the Vietnamese government decided to deal with SOEs more comprehensively through a privatization program, under the name *equitization*, with the goal to raise firms' productivity through supposedly more efficient private ownership. More specifically, the program objectives include improving firms' performance and competitiveness via ownership diversification; mobilizing capital from employees and outside investors, including domestic and foreign investors, for technology renewal and business development; and balancing state, employee, and shareholder interests within the privatized firm. Vietnamese SOEs were allowed to choose one of the following forms of privatization: (i) maintain existing state capital and issue additional shares to mobilize more capital; (ii) sell a part of existing state capital; (iii) sell entire existing state capital; or (iv) sell existing state capital and concurrently issue additional shares (which is by far the most popular form of privatization). In short, the method of privatization in Vietnam takes form of privatization through sale of state property and privatization from below.

The actual privatization process in Vietnam can be divided into pilot stage and expansion stage. The pilot privatization program was initiated in 1992, focusing on small and medium SOEs that were non-strategic and profitable. From 1996 onward, all non-strategic small and medium SOEs were encouraged to be privatized. However, only 123 firms got privatized during the period 1992-1998, according to the National SOE Reform Board, because of opposition to privatization among interest groups and administrative difficulties.

The government then promulgated Decrees 44/1998/ND-CP (1998) and 64/2002/ND-CP (2002) to accelerate privatization, with an official target to privatize all three thousand non-strategic SOEs. The progress of the privatization program during this stage was more impressive. By the end of 2004, 2,242 SOEs had been completely privatized. Most of them were small and medium SOEs (81.5% of these firms had less than VND 10 billion, equivalent to USD 500,000, in capital) and their new shareholders were predominantly insiders (managing board and employees). Finally, the milestone of the privatization program came in 2005 when the government introduced the new Enterprise Law, which unified SOE Law, Foreign Direct Investment Law, and the old Enterprise Law. This legal unification was part of Vietnam's preparation for joining the World Trade Organization, for which Vietnam had promised to create a level playing field for all enterprises. The new Enterprise Law helped boost privatization efforts and as a result, 929 SOEs were privatized in 2005 alone. By the end of 2005, Vietnam's privatization program has achieved its official target.

3.3 Measurements and estimation

3.3.1 Main estimation strategy and identification

I ask the following empirical questions. First, what are the impacts of a higher degree of private firms in upstream and downstream industries on a firm's and an industry's TFP? Second, how can the industry-level impact be decomposed into different parts, notably between incumbents and entrants, and between unweighted average logTFP and allocative efficiency across firms? Third, what are the channels of influence? Fourth, is this privatization impact stronger or weaker under low versus high quality governance? In other words, is privatization and governance quality substitutable or complementary?

I proceed as follows. First, I use Olley and Pakes's (1996) (OP) structural estimation procedure to obtain firm TFPs for each industry. Second, I construct aggregate indices of revenue-weighted shares of upstream/downstream/same-industry private firms. Third, I regress firm and industry logTFPs on privatization measures, taking into account the dynamic autoregressive nature of TFP measures as modeled by OP. Fourth, I interact privatization meaures with industry characteristics to understand the channels. Fifth, I decompose industry logTFP indices into an unweighted average logTFP across firms and a covariance term that measures allocative efficiency Olley and Pakes (1996). I also decompose industry TFPs into incumbent firms' TFPs and entrants' TFPs. I then regress those components on privatization measures. Sixth, I interact privatization measures with measures of provincial governance to understand the substitutability between privatization and governance.

The estimation strategy depends on the identification assumption that privatization in upstream and downstream industries happens independently from the determinants of firm TFPs. After controlling for firm fixed effects and a year trend, the identification comes from off-trend variations in upstream and downstream privatization across industries. I argue that those variations are reasonably exogenous to firm TFPs. As discussed previously, the government's general plan was determined to privatize most industries, either partly or fully, but left the specific privatization schedule open. It implies different paces of privatization across different industries, which were jointly influenced by insider management teams, interested outsider private firms and businessmen, and also regulators across different ministries and provinces. As a result, after controlling for firm fixed effects and a year trend, the remaining variation in privatization over the period in question could be considered as unpredictable, so reverse causation is unlikely.

The identification assumption would be at odd with the existence of a macroeconomic shock that affects all industries differently, such that its impacts on an industry is correlated with how downstream the industry is in the economy's structure. Section 3.6 will explore if this alternative explanation could explain my results.

As will be discussed next, OP's model allows for a reasonably flexible evolution in TFP, that is, it follows a first-order Markov process. It is taken into account by a dynamic panel specification, for firm i in industry s and year t, as follows:

$$logTFP_{ist} = \rho logTFP_{is,t-1} + \beta_S SameInd_{st} + \beta_U Up_{st} + \beta_D Down_{st} + \alpha_{is} + \delta t + \epsilon_{ist}, \quad (3.1)$$

where $SameInd_{st}$, Up_{st} and $Down_{st}$ denote the privatized shares in the same industry and the linkage-weighted privatized shares in upstream and downstream industries, respectively.

A simple fixed-effect regression of this dynamic specification will suffer from the Nickell (1981) bias, because the demeaned variable of lagged outcome $logTFP_{is,t-1}$ will be correlated with the demeaned error term ϵ_{ist}. The bias in this context could be particularly large, because it is of order $O(T^{-1})$ where T is the time dimension of the panel. To avoid biases in dynamic models of panel data, I follow Holtz-Eakin et al. (1988) and Arellano and Bond (1991) in using lags of the outcome variable as instruments in the difference equation.[1] Because the variable $SameInd_{ist}$ (same industry privatization) is potentially determined endogenously in period t, I further use its lags as instruments

[1] I do not apply Arellano and Bover's (1995) method of using lagged differences as instruments in the level equation, because it requires additional assumptions, in particular one of stationarity, that is not justifiable in this context.

in the difference equation. I limit the number of lags to at most 3, so as to keep reasonably strong instruments (Arellano and Bover's (1995) critique). All results are robust to specifications with different numbers of lags, different levels of standard error clustering, and further treating $SameInd_{ist}$ as exogenous.

I also apply the same method to other measures of TFP, including the revenue-weighted average logTFP Ω_{st} in each industry and its components according to different decompositions:

$$\Omega_{st} = \rho\Omega_{s,t-1} + \beta_S SameInd_{st} + \beta_U Up_{st} + \beta_D Down_{st} + \alpha_s + \delta t + \eta_{st}, \tag{3.2}$$

The following subsections will detail methodologies to construct the key measures used in the paper, including TFP estimation using OP's structural estimation method, privatization measures through linkages, OP's decomposition of TFP index, and identification of influence channels.

3.3.2 Total factor productivity estimation

Similar to the large literature on total factor productivity estimation, I start with a simplifying assumption of Cobb-Douglas production function of capital K and labor L for each firm i in industry s and year t: $\log Y_{ist} = \beta^k \log K_{ist} + \beta^l \log L_{ist} + \omega_{ist} + \eta_{ist}$. Here ω_{ist} denotes the part of logTFP term observed by the firm before making decisions on inputs, and η the additional exogenous shock that is thus uncorrelated with inputs. As summarized in Griliches and Mairesse (1998) and Ackerberg et al. (2007), a number of key issues are present in empirical studies of firm-level productivity. First, firms make choices over inputs in order to maximize their objectives, so the input variables K and L are naturally endogenous, and likely correlated with the TFP term ω. Second, in sufficiently long panels, many firms with very low TFP will choose to exit the market, which creates a selection problem for the econometrician. Those two endogeneity problems could be addressed by either finding appropriate instrumental variables for inputs and exits (from natural experiments or economic theory), or imposing a restriction on TFP as a firm fixed effect ($\omega_{ist} = \omega_{is}$).

In practice, the specification of constant TFP as fixed effects can be excessively restrictive Ackerberg et al. (2007), especially in manufacturing industries. It takes root from models where firms learn passively about their fixed TFP over time Jovanovic (1982); Hopenhayn (1992). The fixed TFP assumption implies that the distributions of TFP and firm size remains dependent on the initial draws at entry (i.e. long term memory) Pakes and Ericson (1998). In a more realistic setting, Ericson and Pakes's (1995) model allows TFP to evolve over time, possibly endogenously, according to a Markov process. In this case, the TFP and size distributions are much more dependent on the recent past.[2] Fixed-effect specifications of production functions will be theoretically ill-

[2]Indeed, Pakes and Ericson (1998) finds evidence among manufacturing industries in support of Ericson and Pakes (1995) against Jovanovic (1982).

fitted in a context where firms' production and productivity evolve quickly, such as in a transition economy.[3]

To address concerns of endogeneity and heterogeneity, I thus adopt Olley and Pakes's (1996) (OP) approach to estimate firms' production function and TFP in separate estimations for each industry, as it builds on theoretical results from Ericson and Pakes's (1995) model. The model considers a general, flexible form of industrial structure among firms with heterogenous productivity and firm-specific uncertainties following a Markov process (which encompasses the special case of time-invariant TFP). Each firm's expected profit depends flexibly on the current and future structure of all existing firms, and so do their free decisions to enter or exit. Ericson and Pakes (1995) proves existence of a unique Markov perfect Nash equilibrium, shows that in equilibrium the industrial structure evolves continually as an ergodic Markov process, and determines conditions where (i) investment is a bijective function of productivity, and (ii) exit decision is a step function of productivity. These two properties are crucial in OP's estimation method. We will also use the model's key insights on entry and exit decisions.

The first property implies that investment $i_{it} = i_t(k_{it}, \omega_{it})$ can be inverted to obtain: $\omega_{it} = h_t(k_{it}, i_{it})$. (Lower-case variable names denote the log of corresponding variables in upper cases.) The subscript t indicates potential influences of time-varying factors that affect all firms alike, such as demand shocks. In my analysis, it also represents changes in upstream and downstream privatization. One can then rewrite log output as: $y_{ist} = \beta^k k_{ist} + \beta^l l_{ist} + h_t(k_{it}, i_{it}) + \eta_{ist}$. Now denote $\phi_t(k_{it}, i_{it}) = \beta^k k_{ist} + h_t(k_{it}, i_{it})$:

$$y_{ist} = \beta^l l_{ist} + \phi_t(k_{it}, i_{it}) + \eta_{ist}. \tag{3.3}$$

In the **first step**, this equation is estimated semiparametrically by representing the function $\phi(k, t)$ as a polynomial series of (k_{it}, i_{it}). Total fixed assets are used as capital stocks, and investment is computed from $K_{i,t+1} = (1 - \delta)K_{it} + I_{it}$ for a discount rate $\delta = 0.05$. From the first step, we obtain consistent estimates of $\hat{\beta}^l$ and $\hat{\phi}_t(k_{it}, i_{it})$.

OP's method models ω_{it} to evolve following an exogenous Markov process, so we can write: $\omega_{it} = \mathbb{E}[\omega_{it}|\mathcal{I}_{i,t-1}] + \xi_{it} = \mathbb{E}[\omega_{it}|\omega_{i,t-1}] + \xi_{it} = g(\omega_{i,t-1}) + \xi_{it}$ for some function g where ξ_{it} is an innovation uncorrelated with $k_{i,t-1}$. Thus, we could write:

$$\begin{aligned} y_{it} - \beta^l l_{it} &= \beta^k k_{it} + g(\omega_{i,t-1}) + \xi_{it} + \eta_{it} \\ &= \beta^k k_{it} + g(\phi_t(k_{i,t-1}, i_{i,t-1}) - \beta^k k_{i,t-1}) + \xi_{it} + \eta_{it}. \end{aligned} \tag{3.4}$$

Given the exogeneity of the innovation $(\xi_{it} + \eta_{it})$ in each period t, we can estimate the last equation, using the previous semiparametric estimates $\hat{\phi}_t(k_{i,t-1}, i_{i,t-1})$ and a polynomial series to represent the function g. We need to use nonlinear least squares, because β^k is present both within g and outside.

[3]Furthermore, when inputs are measured with errors (from surveys or calculation assumptions of capital value and labor bill), fixed-effect regressions tend to attenuate the coefficients of inputs, as observed in Ackerberg et al. (2007).

So far, the procedure will produce consistent estimates of production function parameters if the exit decision is not correlated with productivity. From Ericson and Pakes (1995), the exit decision (denoted by $\chi_{it} = 0$) is a step function of productivity: $\chi_{it} = 1 \ iff \ \omega_{it} \geq \overline{\omega}_t(k_{it})$. To correct for endogenous selection, we rewrite the previous output formula in expectation, conditional on past state variables and survival:

$$\mathbb{E}[y_{it} - \beta^l l_{it} | \mathcal{I}_{i,t-1}, \chi_{it} = 1] = \beta^k k_{it} + \mathbb{E}[\omega_{i,t} | \mathcal{I}_{i,t-1}, \chi_{it} = 1] = \beta^k k_{it} + g(\omega_{i,t-1}, \overline{\omega}_t(k_{it})). \tag{3.5}$$

We could apply the nonlinear estimation above to identify β^k only if we have a separate proxy for $\overline{\omega}_t(k_{it})$.

This proxy comes from actual observations of exits.[4] In the **second step**, we estimate the probability of exit $Pr(\chi_{it} = 1 | \mathcal{I}_{i,t-1}) = Pr(\omega_{it} \geq \overline{\omega}_t(k_{it}) | \mathcal{I}_{i,t-1})$ in a probit model and predict it as $\varphi_t(i_{i,t-1}, k_{i,t-1}) = P_{it}$. Let us denote $\overline{\omega}_t(k_{it}) = f(\omega_{i,t-1}, P_{it})$, and put it back into function g in equation (3.5). We can rewrite it to include exogenous innovation terms ζ_{it} and η_{it}, where $\tilde{g}$ is defined by $\tilde{g}(\omega_{i,t-1}, P_{it})) = g(\omega_{i,t-1}, \overline{\omega}_t(k_{it}))$:

$$y_{it} - \beta^l l_{it} = \beta^k k_{it} + \tilde{g}(\phi_{i,t-1} - \beta^k k_{i,t-1}, P_{it}) + \zeta_{it} + \eta_{it}. \tag{3.6}$$

In the **third step**, using estimates $\hat{\beta}^l$, $\hat{\phi}_{i,t-1}$, and $\hat{P}_{it}$ from previous steps, we can use nonlinear least squares to estimate the remaining coefficient β^k. Lastly, TFP is backed out as the residual $y_{ist} - \hat{\beta}^k k_{ist} - \hat{\beta}^l l_{ist}$.

This 3-step procedure produces consistent estimates of the parameters of the production function within each industry.[5] Variables that do not vary across firms in the same the industry, such as my measures of upstream and downstream privatization, enter the estimation as time-variant components. I will therefore retrieve TFP measures and regress them on those variables to estimate their impacts on firm TFP.

The OP procedure assumes that TFP follows a first-order Markov process, of which the constant TFP in fixed effect specifications is a special case. When we study the impacts of different measures of privatization on TFP, it is important to take into account this feature by modeling an autoregressive component of TFP.

I want to highlight three implementation choices of my OP procedure. First, because of the lack of reliable data on firm-level quantities (or prices), I follow the literature in using revenues, not quantities, as the measure of firm output. Labor L is measured as wage bill, and capital K as total fixed assets on the balance sheet. I thus measure the revenue-based TFP, or TFPR (see the distinction in for example Hsieh and Klenow, 2009). The overall effects on TFPR could be coming from either physical TFP (TFPQ) or output prices.

[4] I follow the literature in assuming firms exit if they are not observed before the end date of the panel. Missing data remain a thorny issue for this methodology.

[5] Standard errors from this procedure could be calculated from analytical formulae, by bootstrapping, or by implementing an equivalent one-step GMM procedure suggested by Wooldridge (2009). For my purpose, it is unnecessary to compute standard errors from OP, since I will only use the derived TFP as dependent variable in the next stage.

Second, I do not model a firm's endogenous decision to invest in improving TFP (e.g. Aw et al., 2008) for two reasons: (i) I do not have reliable measures of research and development, or any form of investment towards this end, to be able to identify meaningful variations in firm's decisions to improve TFP, and (ii) I believe most Vietnamese manufacturing firms improve their TFPs via passive learning of new management and technological practices, rather than actively search for new technologies. Good measures of research and development are subject of my future research.

Third, it is possible in OP's procedure to identify in addition an evolution of TFP that depends on macroeconomic changes over time (i.e. identify a function $\mathbb{E}[\omega_{it}|\mathcal{I}_{i,t-1}] = g(\omega_{i,t-1}, \Delta_t) = g_t(.)$, where Δ_t groups all industry-wide shocks at time t). However, I have not done so because that will automatically remove a lot of meaningful variations on the time dimension, including variations pertaining to upstream and downstream privatization.[6] A resulting caveat is that I need to assume that firms would set their capital in expectation of an exogenous process of TFP evolution, and they cannot anticipate the variation of changes in upstream/downstream privatization. In future work, I will include privatization measures directly in OP's estimation to obtain their impacts within each industry.

3.3.3 Decomposition of industry TFP

To understand the channels by which linkage privatization may affect an industry's TFP, I decompose the revenue-weighted average of firm logTFPs ω_{ist} in an industry s, defined as the industry TFP index Ω_{st}, into the sum of the unweighted average of firm TFPs and a covariance term Olley and Pakes (1996):

$$\begin{aligned}\Omega_{st} &= \sum_{i=1}^{N} S_{ist}\omega_{ist} = \frac{1}{N}\sum_{i=1}^{N}\omega_{ist} + \sum_{i=1}^{N}(S_{ist} - \frac{1}{N})(\omega_{ist} - \frac{1}{N}\sum_{i=1}^{N}\omega_{ist}) \\ &= \overline{\omega}_{st} + Ncov(S_{it}, \omega_{ist}),\end{aligned} \tag{3.7}$$

where $S_{ist} = \frac{Y_{ist}}{\sum_{i=1}^{N} Y_{ist}}$ is the revenue share of firm i in industry s. $\overline{\omega}_{st}$ is the unweighted average of logTFP across all firms in industry s. The OP covariance term $Ncov(S_{it}, \omega_{ist})$, measures the degree of efficient allocation of resources across firms in a single industry.[7] It is higher when resources are allocated more efficiently, in that firms with higher TFPs are also producing more output in the industry. It reaches its hypothetical maximum if all revenues are concentrated in the firm with the highest TFP. I will study upstream/downstream privatization's effects on both components of the decomposition.

A different way to measure allocative efficiency, according to Hsieh and Klenow's (2009) model, is to consider the variance term of all firms' revenue-based TFPs (TFPR)

[6] For the same reason, I only include an annual trend in my dynamic panel regressions of TFPs.

[7] See Bartelsman et al. (2013) for a model with labor overhead and semi-fixed capital where firms have heterogenous revenue-based TFPs in equilibrium even without distortion in allocation, and OP's covariance term is the key measure of allocative inefficiency.

in an industry. In their framework, this variance term measures directly the overall effect of distortion on an industry's aggregate TFP. The best allocation of resources happens in an economy without distortion, in such a way to equalize all TFPRs, so the variance will be equal to zero. The two measures of allocative efficiency point to different directions: Better allocative efficiency is increasing in OP's covariance term, while in Hsieh and Klenow's (2009) model it is decreasing in the variance of TFPRs.

Application to two groups: Incumbents and entrants One can apply the decomposition in equation (3.7) to only two groups, such as incumbent and entrant firms. The TFP index of the industry Ω_{st} is equal to half the sum of the TFP index of incumbents Ω^I_{st} and entrants Ω^E_{st}, plus OP's covariance term between those two that indicates allocative efficiency between incumbents and entrants:

$$\Omega_{st} = \frac{1}{2}(\Omega^I_{st} + \Omega^E_{st}) + \frac{1}{2}(S^I_{st} - S^E_{st})(\Omega^I_{st} - \Omega^E_{st}). \tag{3.8}$$

Since the market share of incumbents is much larger than that of entrants, the covariance term is positive when $\Omega^I_{st} > \Omega^E_{st}$. In this case, if we keep the first term constant, allocative efficiency will increase when either $S^I_{st} - S^E_{st}$ or $\Omega^I_{st} - \Omega^E_{st}$ increases.

From equation (3.8), I can apply the decomposition of equation (3.7) on each weighted average logTFP: $\Omega^I_{st} = \overline{\omega}^I_s t + Cov^I_{st}$ and $\Omega^E_{st} = \overline{\omega}^E_s t + Cov^E_{st}$. I end up with a decomposition of Ω_{st} into five different terms. Four of them measure mean and allocative efficiency within each group (incumbents or entrants), and the other measures the allocative efficiency between the two groups.

3.3.4 Measures of privatization and linkages

My measure of same-industry privatization is a revenue-weighted average of an indicator of privatization for each firm. In order to capture key differences between firms that are managed mostly towards "state-owned" objectives and those that are managed more towards private gains (for which profit maximization should be the principal purpose), I choose the indicator whether private ownership exceeds 50% as the benchmark measure of firm-level privatization, or whether there is nonzero foreign ownership. In Vietnam, partially private firms with less than 50% of private ownership are usually heavily subject to state controls, and behave similarly to fully state-owned firms (their private shares are small, and mostly distributed among their own employees without giving them real control rights). I also count all firms with foreign ownership, because they are usually heavily influenced by foreign shareholders, and behave remarkably different from state-owned firms.[8] For robustness checks, I also create other indicators of firms with non-zero private shares and firms with non-state

[8] Because there are multiple barriers to the repatriation of capital and profits, foreign shareholders, mostly direct investors, usually request strong influence on a firm's management as a necessary condition for them to invest.

control.[9] The measure is then: $SameInd_{st} = \frac{Y_{ist}P_{ist}}{\sum_{i=1}^{N} Y_{ist}}$, given the privatized indicator P_{ist} and revenue Y_{ist}.

I use Vietnam's input-output matrix (2 digit industry level) to measure backward linkages from privatization in downstream industries, and forward linkages from privatization in upstream industries. I compute the matrix $(\alpha_{sk})_{s,k=\overline{1,J}}$ where α_{sk} denotes the proportion of industry s's output supplied to industry k: $\alpha_{sk} = \frac{FY_{sk}}{\sum_j FY_{sj}}$ where FY_{sj} is the output flow from industry s to j (output from s used in final consumption and export is thus excluded). Similarly, I compute $(\zeta_{ks})_{k,s=1}^{J}$ where ζ_{ks} denotes the proportion of industry s's intermediary inputs supplied from industry k: $\zeta_{ks} = \frac{FY_{ks}}{\sum_j FY_{js}}$. By definition, $\sum_k \alpha_{sk} = 1$ and $\sum_k \zeta_{ks}$=1. An industry s's measure of downstream privatization is a weighted sum of privatization in all industries except itself, with weights α_{sk}. Its measure of upstream privatization is a weighted sum of privatization in all industries except itself, with weights ζ_{ks}:

$$Down_{st} = \sum_{k \neq s} \alpha_{sk} SameInd_{kt} \ ; \ \ Up_{st} = \sum_{k \neq s} \zeta_{ks} SameInd_{kt}. \tag{3.9}$$

Because the flow of intermediary goods within each industry s is excluded in the calculation of its measures of upstream and downstream privatization, the weights in $Down_{st}$ add up to only $1 - \alpha_{ss}$, and those in Up_{st} to $1 - \zeta_{ss}$.

3.3.5 Investigation of influence channels

I am further interested in the channels of spillovers from downstream privatization via backward linkages. Privatization in downstream industries may affect their suppliers' productivity through (i) increasing demand for intermediate products, which allows suppliers to benefit from economies of scales, such as suggested by Rodriguez-Clare (1996), (ii) higher requirements for product and service quality, which pressure suppliers to improve their management and technology, and/or (iii) direct technological spillovers. I control for the demand channel with a measure of demand addressed to each industry s: $Demand_{st} = \sum_k \alpha_{sk} Y_{kt}$. I also control for the Herfindahl-Hirschman index of industry concentration to account for the type of pressure from competition similar to that in Aghion et al. (2005) and Aghion et al. (2009).

The pressure channel should work best among firms that are geographically close, and depend on each industry's exposure towards its downstream industries. In a model such as Ericson and Pakes's (1995), the pressure channel could be understood as more stringent conditions of survival. Consequently, only entrants with sufficiently high draws of TFP choose to enter the market, so the effect of downstream privatization on entrants' TFP should be positive.

The direct technological spillover channel is related to technological spillovers from FDIs' vertical linkages Javorcik (2004). In the context of privatized firms, they probably

[9] See Appendix Table C.1 for descriptions of firm ownership classification and private firm definitions.

do not have better (explicit) knowledge of production technology than state-owned firms. If it is about the transfers of tacit knowledge, including for example incentive systems and management practices, it should likely work through learning by doing, and should affect mostly incumbent firms, not entrants. Therefore, by studying the effect of downstream privatization on incumbents versus entrants, we could deduce whether it is about pressure or transfer of tacit knowledge.

3.4 Data description

3.4.1 Data sources

My main data source is the Vietnam Enterprise Census (VEC), which is conducted annually by the General Statistics Office of Vietnam (GSO) since 2000. All firms with more than 9 workers are legally obliged to fill out the census' questionnaire. Although this questionnaire has been slightly revised every year, the basic indicators remain unchanged. These indicators cover information on firms' name, address, ownership, industry, revenue, assets, investment, employment, wage bill, imputed material costs as well as tax records and industry-specific variables. I match firms in VEC data across different years using their names and corporate tax codes to construct a panel dataset covering the period 2001-2008.[10] This raw panel dataset contains 886,997 firm $\times$ year observations and covers a total of 294,323 firms. Tables 1 and 2 report the key summary statistics of firms included in this benchmark panel dataset.

I construct my benchmark dataset of manufacturing firms from this raw dataset. A challenge I face is that GSO changed its industry classification system in 2006. Before 2006, GSO used the Vietnam Standard Industrial Classification (VSIC) 1993 system, a version of the International Standard Industrial Classification (ISIC) Revision 2. From 2006 onward, GSO moved to VSIC 2007, which is closer to the ISIC Revision 3. The Vietnam's 2000 input-output matrix also employs another classification system that slightly differs from the VSIC 1993. As a result, it takes me a considerable amount of time to match the industry codes across these systems of classification. The next challenge is that some firms in my panel dataset have different industry codes (at 2-digit VSIC level) in different years. I identify these firms' main industry as the one recorded for at least 50% of the firms' occurrences in the dataset and drop the firms for which I could not identify their main industry (these firms account for about 10% of total manufacturing firms). I also drop all observations with zero or negative revenue from sales of goods and services (proxy for output), fixed assets (proxy for capital), or wage bill (proxy for labor). Finally, I have to exclude three industries – manufacturing of tobacco (VSIC 16), refined petroleum products (VSIC 23), and office machinery and computers (VSIC 30), as the small sample size makes it impossible to apply the Olley-

[10]VEC data in year 2000 do not contain information on revenue. I also exclude years 2009 and 2010 to avoid the large noises coming from the global crisis that started to hit Vietnam in this period.

Pakes procedure to these industries.[11] My final benchmark panel dataset contains 145,380 firm $\times$ year observations, covering 43,545 firms across 18 two-digit level VSIC manufacturing industries. Table 3.1 reports the distribution by industries of firms included in this benchmark panel dataset.

I classify firms' ownership using the firm-type question included in every year of the census. This question asks firms to identify themselves as one of the 14 different firm types listed, ranging from fully SOE to fully-private enterprise to fully-foreign-owned enterprise (more details on this in Appendix Table C.1). I also cross-check a firm's answer to this question with information on its capital structure to ensure that I correctly identify its ownership type. Throughout most of this paper, I define a private firm as one having at least 50% domestic private ownership or one having non-zero foreign ownership. In a few cases, I use slightly different definitions of private firms for robustness checks. These definitions are detailed in Appendix Table C.1.

For total factor productivity (TFP) estimation using Olley-Pakes procedure, I calculate real output as reported revenue from sales of goods and services, deflated by the Producer Price Index (PPI) of the corresponding industry.[12] Capital is defined as the value of fixed assets at the beginning of the year, deflated by the simple average of the deflators for five industries: machinery and equipment, office machinery and computers, electrical equipment and apparatus, motor vehicles, and other transport equipment. Labor is proxied by total wage bill, deflated by the Consumer Price Index (CPI). Investment is equal to the value of total reported investment, deflated by the same deflators used for capital. As robustness checks, I use total employment count for labor instead of total wage bill or include imputed material costs, deflated by CPI, in my TFP estimations.[13] The summary statistics of these key variables and resulting estimated TFP are reported in Table 3.2.

Finally, my other data sources include Vietnam's 2000 input-output matrix, obtained from GSO, and Vietnam Provincial Competitiveness Indices, a set of survey-based indices of industries' governance perceptions constructed with the help of UNDP, for the period 2005-2008.

3.4.2 The process of privatization in Vietnam

Table 3.1 describes the manufacturing industries in my sample, and the overall picture of privatization from 2001 to 2008. Privatization has taken place gradually across all

[11]There are 37 firms in manufacturing of tobacco (VSIC 16), 62 firms in manufacturing of refined petroleum products (VSIC 23), and 44 firms in manufacturing of office machinery and computers (VSIC 30) in 2001-2008.

[12]GSO only publishes industry-level PPI data for 2005 and 2010 onward. I construct PPI data for the period 2001-2008 by regressing available PPIs on relevant CPI components of the same period and use the resulting coefficients to predict past PPIs from past CPI components (which are available for this period). I am confident about this approach, as most of these regressions have R-squared above 0.9.

[13]VEC does not directly collect information on material costs in its questionnaire. The VEC dataset only contains imputed material costs calculated for a subsample of firms based on information collected in a much smaller-scaled survey on manufacturing activities. Therefore, I include imputed material costs only in my alternative TFP estimations used for robustness checks.

industries, albeit at different paces. By the end of the period, most firms have become private, but the remaining SOEs are quite sizeable. Privatization has been fastest in publishing and printing, wearing apparel, and textiles, with changes of more than 30 percentage points in $SameInd_{st}$, but also quite rapid in non-metallic mineral products and machinery and equipment. At the other end, private firms' shares increased very little for beverages, motor vehicles and other transport equipment. Private firms' shares have increased from 66.1% in 2001 to 86.8% in 2008, or an increase of 20.7%.

Table 3.1 shows the evolution over time of each of the three measures of privatization across all industries. The overall upwardness in all three measures is universal, although their trends could differ a lot in each industry and year. Upstream and downstream measures are substantially smaller, because the flow of intermediary goods within each industry s is excluded in their calculation. 3.1

Table 3.2 shows the summary statistics of the major variables used in my estimations across firms, and Table 3.3 describes privatization summary statistics per year across industry $\times$ years. The standard deviation of the three privatization measures is around 16-18%. Across industries, same-industry, downstream, and upstream privatization measures have grown by 17.5, 18.1 and 12.1 percentage points respectively over the course of 7 years from 2001 to 2008. Privatization has thus been mostly gradual.

3.5 Estimation results

3.5.1 The distribution and evolution of firm TFPs in Vietnam

Table 3.4 exhibits the different decompositions of firm TFPs as discussed in section 3.3.3 for every year. The annual growth rate in TFP index amounts to 2.6%, although there was a slight decline around 2001-2003 before recovery from 2004 to 2008. Over the period 2001-2008, roughly half of the gain in TFP comes from the unweighted average among firms, and the other half from better allocation of resources across firms. Those numbers are illustrated in Table 3.2.

Among all firms, private firms have much higher TFPs than state-owned firms, as illustrated in Table 3.3. Furthermore, this gap seems to widen over time, as state-owned firms struggle to improve their TFPs after the early slump from 2001 to 2003. Table 3.4 shows that within the remaining state-owned firms, unweighted average logTFP actually increased. However, resource allocation among them deteriorates, which more than mitigates the apparent improvement in unweighted average logTFP.

A closer look at the evolution of TFPs within firms that privatized sheds further light on state-owned firms' unweighted average logTFP. Table 3.4 examines the average TFP of firms that are eventually privatized, in reference to the year of privatization. Pre-privatization TFP is low and decreasing until privatization, but bounces back after privatization. Those facts put together imply that state-owned firms with low TFP are likely to get privatized and switch to the sample of private firms. They also suggest that privatization improves TFP within firms. The validity of this conclusion relies on the differences-in-differences assumption that firms that do not privatize would follow the pre-privatization trend.

If this assumption does not hold in reality, the chance of privatization is endogenous to firms' recent TFP, so one cannot use the double differences to obtain the direct effect of privatization on privatized firms. This argument echoes the reason why identification of privatization's direct effects is difficult. Even when I only look at the industry-wide measure of same-industry private shares, there still is a possibility of immediate correlation between the privatization measure and the recent innovation term in TFP. Therefore, in my dynamic models of panel data of firms and industries, I include same-industry privatization as an endogenous variable, so that its coefficient is identified using its lagged level as instrument for the difference equation. In contrast, upstream and downstream privatization measures are much less likely to suffer from this direct endogeneity problem. Hence they are included directly in the regressions.

The allocative efficiency of resources between private and state-owned firms is measured by $\frac{1}{2}(S_{st}^{P} - S_{st}^{SO})(\Omega_{st}^{P} - \Omega_{st}^{SO})$ as in equation (3.8) where superscripts P and SO indicate private and state-owned, respectively. This term has substantially improved from 2001 to 2008, since S_{st}^{P} increases from 65.6% to 86.3%, and the TFP gap widens from 0.42 log points to 0.59 log points (roughly 17 percentage points increase). Overall, this has substantially contributed to the efficiency of all manufacturing industries.

In the decomposition of TFP index of private firms into those of incumbents and entrants, incumbents' TFP index is larger than entrants' TFP index, but the gap shrinks considerably from 0.18 log points in 2001 down to 0.03 log points in 2008. While incumbents' TFP has increased at rate of about 1.3% a year, entrants' TFP grows at 3.4% annually. Among incumbents, the gain comes entirely from better allocative efficiency. In contrast, among entrants, the gain comes only from better TFP among new batches of entrants, while allocative efficiency among entrants fluctuates but does not improve. Lastly, the allocation of resources between entrants and incumbents is measured by $\frac{1}{2}(S_{st}^{I} - S_{st}^{E})(\Omega_{st}^{I} - \Omega_{st}^{E})$ from equation (3.8). This term drops in 2008 when the TFP gap becomes very small.

The evolution of TFPs in Vietnam is similar to that of China in Brandt et al.'s (2012) description, especially in that entrants made the strongest progress in TFP growth during the same period. The annual contribution of entrants to overall TFP growth may not be large though, because of their small shares each year.

3.5.2 Spillover on firm-level productivity

Strong spillovers on firm TFP from downstream privatization

Table 3.5 explores the association between firm TFP and privatization in downstream and upstream industries, using Section 3.3.1's dynamic specification that allows for firm fix effects and autoregressive TFP.[14] The privatization measures for same-, upstream, and downstream industries are detailed in section 3.3.4. Columns (1) reports that the coefficient of spillovers from downstream privatization through backward linkages to the sample of private firms is positive and statistically significant. The point estimate of 0.336 implies that one standard deviation increase in downstream privatization corresponds to a sizable increase of 0.06 log points in firm TFP. The coefficient of same-industry privatization is similarly positive and significant, while I do not detect any spillovers from upstream privatization. Columns (2) and (3) show results for the samples of state-owned firms and firms privatized during the period, respectively. None of the coefficients of same-industry, upstream, and downstream privatization is positive and significant in these regressions. Table 3.5 thus offers complementary evidence that not only do state-owned and newly privatized firms in transition economies struggle to reap the direct benefits of privatization (e.g. Estrin et al., 2009), they do not even gain from the indirect spillovers from privatization in upstream and downstream industries.

Columns (4) considers the baseline sample of private firms, and controls for (i) the demand channel using demand for intermediate products from downstream industries, and (ii) industry concentration using same-industry Herfindahl-Hirschman Index as mentioned in section 3.3.5. Column (5) further controls for downstream and upstream HHI, where $DownHHI_{st} = \sum_{k \neq s} \alpha_{sk} HHI_{kt}$ and $UpHHI_{st} = \sum_{k \neq s} \zeta_{ks} HHI_{kt}$. If downstream privatization increases downstream industries' demand for intermediate products, which in turn increases firm productivity through economies of scale, this situation can still be regarded as a spillover effect. The same argument also applies for spillover effect through decreasing downstream or upstream concentration. However, even after controlling for these channels, the coefficients of downstream privatization remain positive and significant, suggesting a more direct spillover channel than the ones controlled for. The results from Table 3.5 together indicate that private firm productivity improves with not only privatization in the same industry but also privatization in downstream industries. The productivity autoregressive coefficient is consistently around 0.31-0.32, which confirms and strengthens the validity of my estimation strategy.

Table 3.6 presents robustness checks for the spillover effect of downstream privatization on private firm productivity. Columns (1) and (2) use the samples of firms that were always private and firms that were never private throughout the period, respectively. The results are analogous to those reported in columns (1) and (2) of Table 3.5. Columns (3) and (4) use alternative definitions of private firms to calculate the privatization measures and define the sample of private firms. The effect disappears when I use a looser definition (i.e. firms having non-zero domestic private or non-zero foreign ownership) and strengthens when I use a stricter definition (i.e. firms controlled by domestic private or foreign entities.) The last two columns use alternative TFP

[14] OLS regressions with firm fixed effects yield similar point estimates for the spillover effects of downstream privatization on firm productivity in different samples of firms. These results are reported in Appendix Table C.2.

measures estimated using employment count instead of wage bill (column (5)) or using wage bill and imputed material costs (column (6)). The spillover effect of downstream privatization remains robust in these regressions with similar point estimates.[15]

Spillover driven by domestic privatization rather than foreign ownership

The existing literature has generally found that foreign ownership has strong positive effects on firms' performance in transition economies. These effects may apply directly to firms with foreign participation Estrin et al. (2009); Commander and Svejnar (2011), or indirectly to domestic firms through knowledge transfer or via competitive pressure Javorcik (2004); Gorodnichenko et al. (2010); Arnold et al. (2011). On the other hand, research on the direct effect of domestic privatization on firms' performance has yielded mixed results Estrin et al. (2009), and none thus far has attempted to access the indirect spillover effect of privatization through upstream and downstream linkages. Table 3.7 thus decomposes privatization measures into purely-domestic and foreign-affiliated components to examine which is the driver of the spillovers from downstream privatization. The first two columns split the baseline sample of private firms into purely-domestic private firms (column (1)), for which the effect remains strong and significant, and foreign-affiliated firms (column(2)), for which the effect disappears. I then split same-industry, upstream, and downstream privatization measures into shares of purely-domestic private firms and shares of foreign-affiliated firms in same, upstream, and downstream industries.[16] The result reported in column (3) suggests that in my case, domestic privatization and not foreign entry is the key driver of firm productivity improvement. Columns (4) to (6) are similar to the first three columns but use alternative definitions based on control rights, namely domestic-controlled private firms and foreign-controlled firms. The results from these regressions confirm my novel finding on the importance of domestic privatization.

Spillover from downstream privatization mainly due to pressure channel

As discussed in section 3.3.5, I will include the two following control variables. Increased demand for intermediate products is accounted for by the variable $Demand_{st}$ constructed from downstream demands, and concentration by the industry's Herfindahl-Hirschman Index HHI_{st}. In column (4) of Table (5), while the coefficient of demand is positive and statistically significant as expected, the coefficient of downstream privatization also remains sizable and significant, suggesting that the spillovers also work through

[15]I replicate the whole set of regressions in Tables 5 and 6 with TFP measures estimated using wage bill and imputed material costs. My key results remain qualitatively similar and are reported in Appendix Table C.3.

[16]The share of purely-domestic private firms in industry s is $SameInd^D_{st} = \frac{\sum_i Y_{ist} D_{ist}}{\sum_i Y_{ist}}$, where D_{ist} is the indicator of purely-domestic private firms. Similarly, the share of foreign firms in industry s is $SameInd^F_{st} = \frac{\sum_i Y_{ist} F_{ist}}{\sum_i Y_{ist}}$, where F_{ist} is the indicator of foreign firms. The upstream and downstream measures of privatization for purely-domestic private firms or foreign firms are calculated by applying equation (3.9) to $SameInd^D_{st}$ and $SameInd^F_{st}$.

increasing pressure to deliver from downstream industries. I now check the prediction regarding the pressure channel that spillovers from more closely-linked downstream firms are larger. This happens when (i) the downstream firms are in the same local market, or (ii) intermediate consumption from downstream industries constitutes a larger shares of gross outputs.

Table 3.8 compares the spillover effect of privatization within the same market and that of privatization in different markets. I conduct the analysis with two different geographical definitions of markets: (i) GSO's classification of 8 different regions, and (ii) a classification into three large regions, North, Center, South, that are historically and geographically separable.[17] Columns (1) to (3) of Table 3.8 consider a local market as one of the three large regions, while columns (4) to (6) replicate them on the 8 smaller regions. Columns (1) and (2) contrast results using privatization measures calculated only for the same local market (column (1)) and only for other markets (column (2)).[18] Consistent with my prediction, the coefficient of downstream privatization in column (1) more than doubles that in column (2), and both are statistically significant. Column (3) includes both sets of same-market and different-market privatization measures in one same regression. Only the coefficient of downstream privatization in the same market remains significant with similar point estimate to that in column (1). In the replications on 8 smaller regions, shown in columns (4) to (6), the results are qualitatively similar, even though the coefficient of downstream privatization in the same market is much bigger as markets are more narrowly defined. The point estimate of 1.178 implies that one standard deviation increase in same-market downstream privatization measure corresponds to a sizable increase of 0.22 log points in firm TFP. Together, results from Table 3.8 strongly suggest that spillovers from downstream privatization are much larger when the privatization takes place within geographical proximity.

Table 3.9 examines the prediction that downstream privatization has stronger spillover effects across industries with higher intermediate consumption shares of gross outputs. Intermediate consumption share of gross outputs IC/GO is calculated as $IC/(IC + FC) \times (1 - X/GO)$, and imports shares of intermediate consumption as $M/(IC + FC)$, where GO is the industry's gross outputs, IC – intermediate consumption, FC – final consumption, X – exports, and M – imports, as $IC + FC + X = GO + M$.[19] If most of a industry's gross outputs go on to be inputs for its downstream industries (high intermediate consumption shares of gross outputs), then

[17] The North includes the Red River Delta, North East, and North West regions. The Center includes the Northern Central Coast, Southern Central Coast, and Central Highlands regions. The South includes Mekong River Delta and South Eastern regions.

[18] For a local market m and industry s, the share of same-market private firms in industry s is $SameInd^{SM}_{mst} = \frac{\sum_i Y_{ist} P_{it} SM_{imt}}{\sum_i Y_{ist}}$, where SM_{imt} is the indicator whether firm i is in market m, and P_{it} indicates its private status. Similarly, the share of different-market private firms in industry s is $SameInd^{DM}_{mst} = \frac{\sum_i Y_{ist} P_{it}(1-SM_{imt})}{\sum_i Y_{ist}}$. The upstream and downstream measures of privatization for same versus other local markets are calculated by applying equation (3.9) to $SameInd^{SM}_{mst}$ and $SameInd^{DM}_{mst}$.

[19] IC/GO and $M/(IC + FC)$ ratios are calculated for each industry based on Vietnam's 2000 IO table.

higher product and service requirements from these downstream industries are likely to exert more pressure, and thus downstream privatization is likely to have strong spillover effect on the industry's TFP. The converse is expected if most of the industry's gross outputs are channeled into final consumption or exports instead (low intermediate consumption shares of gross outputs.) Columns (1) and (2) split the baseline sample into firms in intermediate-good-producing industries, defined as industries with above-median intermediate consumption shares of gross outputs, and firms in remaining non-intermediate-good-producing industries, respectively. Column (3) directly compares the different spillover effects of downstream privatization across these industries by interacting the downstream privatization measure with an indicator whether the firm belongs to an intermediate-good-producing industry. The results from these regressions confirm my prediction. The spillovers are pronounced among firms in industries that provide relatively large shares of their gross outputs to downstream industries and is non-detectable among the remaining firms. The difference between these spillover coefficients is also sizeable and statistically significant.

If higher requirements from downstream industries put pressure on supplier firms to improve, this effect should be reinforced for supplier firms facing more competition with imports. Columns (4) to (6) examine this question by comparing the spillover effects of downstream privatization on import-competing industries, defined as industries with above-median imports shares of intermediate consumption, and remaining non-import-competing industries. Again, the spillovers are large among firms in industries facing high competition with imports and is much smaller otherwise. The difference is also sizable and statistically significant, as reported in column (6). Finally, columns (7) and (8) report that the spillovers from downstream privatization are most pronounced among firms in industries with both strong interactions with downstream industries (measured as high intermediate consumption shares of gross outputs) and intense competitive pressure from imports (high imports shares of intermediate consumption.)

Overall, the results from this section establish that domestic privatization in downstream industries has a positive spillover effect on the productivity of their supplier firms. This effect is driven by the increasing pressure to deliver from downstream industries following their privatization, and is larger where there are stronger interactions with downstream industries. Therefore, I observe the strongest spillover effect from downstream privatization in the same market, in industries with high intermediate consumption shares of gross outputs, and in industries facing intense competition with imports.

3.5.3 Spillover on industry $\times$ market-level productivity

Improvements from unweighted average logTFP, not allocative efficiency

I now explore whether the within-firm productivity improvement that corresponds to downstream privatization found in the previous section is meaningful at an aggregate level. If within-firm productivity improvement is coupled with more efficient resource allocation across firms, the spillover effect on aggregate productivity could even be larger than that on firm-level productivity. In contrast, if allocative efficiency deteriorates, downstream privatization may have little or no effect on aggregate productivity. To address this question, I aggregate firm productivity at the industry $\times$ market level (using the three-region definition of markets), and calculate measures of privatization, concentration, and demand at the same level.[20]

In Table 3.10, I apply the baseline dynamic specification that includes industry $\times$ market fixed effects and an autoregressive coefficient (section 3.3.1) on the decomposed parts of industry TFP index, including the unweighted average logTFP and OP's covariance term as detailed in section 3.3.3. As the baseline results in Table 3.5 suggest, I look at aggregate productivity of private firms and state-owned firms separately. Columns (1) to (3) confirm that downstream privatization does not have any spillover effect on state-owned firms on aggregate. The coefficients of downstream privatization on state-owned firms' weighted average logTFP, unweighted average logTFP, and OP's covariance term are all negative, though none of them is statistically significant.

Among private firms, the coefficient of downstream privatization on unweighted average logTFP is positive and significant (column (5)), consistent with evidence of spillovers on firm-level productivity from the previous section. I do not detect any downstream privatization spillovers on weighted average logTFP (column (4)) or OP's covariance term (column (6)). In the latter regression, the coefficient of downstream privatization is negative and sizable, though not statistically significant. Allocative efficiency in manufacturing industries Vietnam, reported in Table 3.4, has improved during this period, similar to what happened in many other transition economies Bartelsman et al. (2013). However, results from Table 3.10 imply that spillovers from downstream privatization do not contribute to this development.[21]

Improvement on extensive rather than intensive margins

Table 3.11 studies spillover effects of downstream privatization on private firms' aggregate TFP index through intensive and extensive margins. I apply the two-group decomposition detailed in section 3.3.3 to incumbents versus entrants. Columns (1) to (3) examine downstream privatization's spillover effect on incumbents' aggregate productivity, including the TFP index, unweighted average productivity, and the OP covariance term among incumbent firms. Columns (4) to (6) replicate them for entrants' aggregate productivity. In the first three regressions, downstream privatization coefficients are statistically nonsignificant. The coefficient on unweighted average logTFP is positive and sizeable, while that on OP's covariance term is negative.

[20] All results remain qualitatively similar if I aggregate firm productivity at the industry level instead. These results are reported in Appendix Tables 4 and 5.

[21] I also replicate this exercise using log real labor productivity (LPR) as the measure of productivity, as used by Bartelsman et al. (2013). The results are qualitatively similar to those reported in this table.

On the contrary, among entrants, downstream privatization's coefficients on weighted average logTFP (column (4)) and unweighted average (column (5)) are both large and significant, and the coefficient on OP's covariance term (column (6)) is also positive although not significant. The point estimate of 1.861 in column (4) implies that one standard deviation increase in downstream privatization measure corresponds to an impressive increase of 0.33 log point in entrants' TFP index. Column (12) shows in a robustness check that the same coefficient in a firm-level regression among entrant firms[22] doubles that in the baseline firm-level regression (0.813 versus 0.405). These results indicate that there are positive spillovers from downstream privatization on entrants' productivity and they are meaningful at both aggregate and firm levels.

So far, I have observed that downstream privatization positively affects entrants' TFP index, but has an imprecise effect on incumbents' TFP index or overall private firms' TFP index. This leaves us with the remaining OP covariance term between incumbent and entrant groups of firms. Results in columns (7) (covariance of weighted average logTFP and revenue shares) and (8) (covariance of unweighted average logTFP and revenue shares) confirm that downstream privatization reduces the OP covariance term, which explains the nonsignificant result on overall private firms' TFP (column (4) in Table 3.7).

To understand those results better, recall that the OP covariance term in the two-group decomposition of TFP index is $\frac{1}{2}(S^I_{st} - S^E_{st})(\Omega^I_{st} - \Omega^E_{st})$. Columns (9) and (10) show that privatization in downstream industries helps reduce the productivity gap between incumbent and entrant private firms, respectively in terms of both weighted and unweighted average logTFP (though incumbents remain more productive than entrant firms on average), which is consistent with my previous results. Column (11) shows that downstream privatization also contributes to the increasing revenue shares of entrants, and thus the decreasing revenue share gap with incumbents, as reported in Table 3.4. Both effects explain the negative coefficient of downstream privatization on the OP covariance term found in columns (7) and (8). While it is indicative that downstream privatization reduces allocative efficiency, I need to be cautious that this exercise regards only the short-term year-on-year effects. If new, better entrants will continue to grow in TFP to a level above incumbents', then this apparent reduction of allocative efficiency in the short run could be the root of better allocative efficiency in the long run. This is a key caveat of section 3.3.3's decomposition.

The results in this section indicates clearly that spillovers from downstream privatization work mostly through entrants' productivity. As argued from section 3.3.5, the evidence is thus against the influence channel of direct knowledge transfers between existing firms, which probably pass through learning by doing, since privatized firms are unlikely better endowed in the type of hard knowledge transferrable via blueprints and contractual agreements, as usually discussed in the literature on FDI spillovers Javorcik (2004); Lin and Saggi (2007). On the contrary, it is much more likely that

[22] In this column, I only use a standard fixed effect specification without dynamics, because individual entrant firms cannot have lagged TFP in their first year in the sample.

firms are put under higher pressure to improve efficiency by privatized downstream industries, so that entrants will only enter if they are already very productive. This channel has been overlooked in extant research on both FDI spillovers and privatization.

3.5.4 Privatization effect and institution quality are substitutes

In this section, I examine the interaction between spillovers from downstream privatization and institutional environment. If private ownership and institution quality are complements, the spillover effect is likely stronger in better institutional settings. The converse is true if private ownership and institution quality are substitutes instead.

I use provincial governance indicators taken from the Vietnam Provincial Competitiveness Indices (PCI) to proxy for institution quality. Many studies of Vietnam have documented that the provincial government, more than the central government, is the relevant level of government when thinking about the institutional climate facing firms Bai et al. (2013); Malesky (2008). The PCI is a set of survey-based indices of industries' governance perceptions that has been systematically constructed with the help of UNDP since 2005 (see details in Malesky (2006) and subsequent reports.) Among the available indicators, I select two that are relevant to my context: (i) state-owned enterprise (SOE) bias index, which measures the extent of provincial policy bias towards state-owned firms (for example, in terms of access to cheaper capital), and (ii) entry cost index, which assesses the differences in entry costs for new firm across provinces. The first indicator speaks volume to the gap between private and state-owned firms considered in previous sections, and the second indicator helps clarify the channel of influence through entrants' improved TFPs. Other indicators available in the PCI are not directly related to my discussion.[23] Since the PCI is available only from 2005 onwards, I use the averages of these indices in each province between 2005 and 2008 as proxies for that province's institution quality throughout my sample period (2001-2008). This approach is valid since institution quality does not change drastically between 2005 and 2008.

Columns (1) to (4) of Table 3.12 explore the interaction between downstream privatization spillovers and provincial SOE bias. In columns (1) and (2), the spillover effect is large and significant in both samples of private firms in provinces with above-median and below-median SOE-bias scores (higher score indicates less bias), with similar point estimates. When I interact downstream privatization measure with either a dummy indicating if the firm's province has above-median SOE-bias score (column (3)) or directly with province's SOE-bias score (column (4)), the coefficients of the interaction terms in both regressions are negative and statistically significant. These results suggest that downstream privatization has larger spillover effect in provinces with larger market imperfection caused by SOE bias.

Columns (5) to (8) repeat the exercise and show similar patterns of substitutability between entry-cost score and privatization spillover effects. The only difference with the

[23]Further details on PCI's methodology and indices are available online at `http://eng.pcivietnam.org/index.php`.

previous 4 columns is in column (5), where the coefficient of downstream privatization among private firms in provinces with below-median entry costs is much smaller and not significant. Column (9) combines both measures of governance imperfection to show that downstream privatization spillovers matter more when both imperfections are strong. Finally, column (10) saturates the model with all interactions between SOE bias, entry costs and downstream privatization, and shows that downstream privatization spillovers matter most in provinces with largest market imperfection in both dimensions of strong SOE bias and high entry costs.

These results provide evidence that private ownership and institution quality are substitutes in my setting. When there is initially a strong gap between state-owned and private firms, privatization brings a larger change to the business environment, especially reinforcing the pressure on suppliers of the privileged state-owned firms. In areas where entry costs are high, it is likely that incumbents have enjoyed some market power without worry about competing potential entrants. Downstream privatization exerts additional pressure that likely improves both incumbents and entrants' productivity.

The substitutability between downstream privatization effect and institution quality implies that when a large-scale reform including governance improvements and privatization is undertaken, its effects on economic performances may be lower than the sum of potential effects of all components, at least in the short run. This observation clearly does not recommend against large-scale reforms. Rather, one should guard against premature conclusions that their effects are too low.

3.6 Discussion and concluding remarks

This paper has found robust evidence of a strong positive association between downstream privatization and private firm TFP improvement. The effect comes more from privatization to domestic than foreign shareholders. However, the influence on industry TFP indices is substantially mitigated by a reverse effect on within-industry allocative efficiency. The effect comes mostly from improved TFP among entrants, whose market shares are growing and whose TFPs are catching up with incumbents'.

This could be taken as evidence of causation under the identification assumption that no detrended shocks can have differential effects across industries that cause both downstream privatization and upstream TFP growth. Even if there are correlated shocks that affect privatization, I would argue that they will bias the results against my findings. For a single firm, as we have seen in Table 3.4, a negative shock to TFP is likely to lead to privatization. A negative shock is likely to also affect immediate upstream and downstream industries negatively, which creates a negative correlation between TFP and downstream privatization. So it works against my finding of a positive relationship. Therefore, I am confident of my causal interpretation of the results.

In the paper, I have controlled for several potential channels of influences from downstream privatization to upstream TFP. To account for privatization's effect on the size of the downstream industry, I have controlled for a measure of aggregate demand addressed to each industry. I include the Herfindahl-Hirschman Index of industrial concentration to control for competition. Regarding technological transfers, I argue that privatized firms do not possess better technological blueprints after privatization, so the technological edge they may diffuse must be tacit knowledge of management and organization. Such knowledge is hardly contractible for transfers, and would have been learned by doing instead. Since the strong effect of downstream privatization comes on entrants, instead of incumbents, I conclude that the major channel of influence must be direct effect of higher pressure on suppliers coming from newly privatized client firms.

I further discover that privatization and local governance quality are substitutes. Privatization is likely to bring the best TFP improvements out of upstream industries in provinces where private, entrant firms are most likely to suffer from bad governance. With the natural caveat of generalizability, this finding suggests that further privatization would bring negligible gains in good-governance environments (e.g. Singapore or Taiwan), while it could bring additional economy-wide gains in developing countries.

Figure 3.1: Privatization measures (2001-2008)

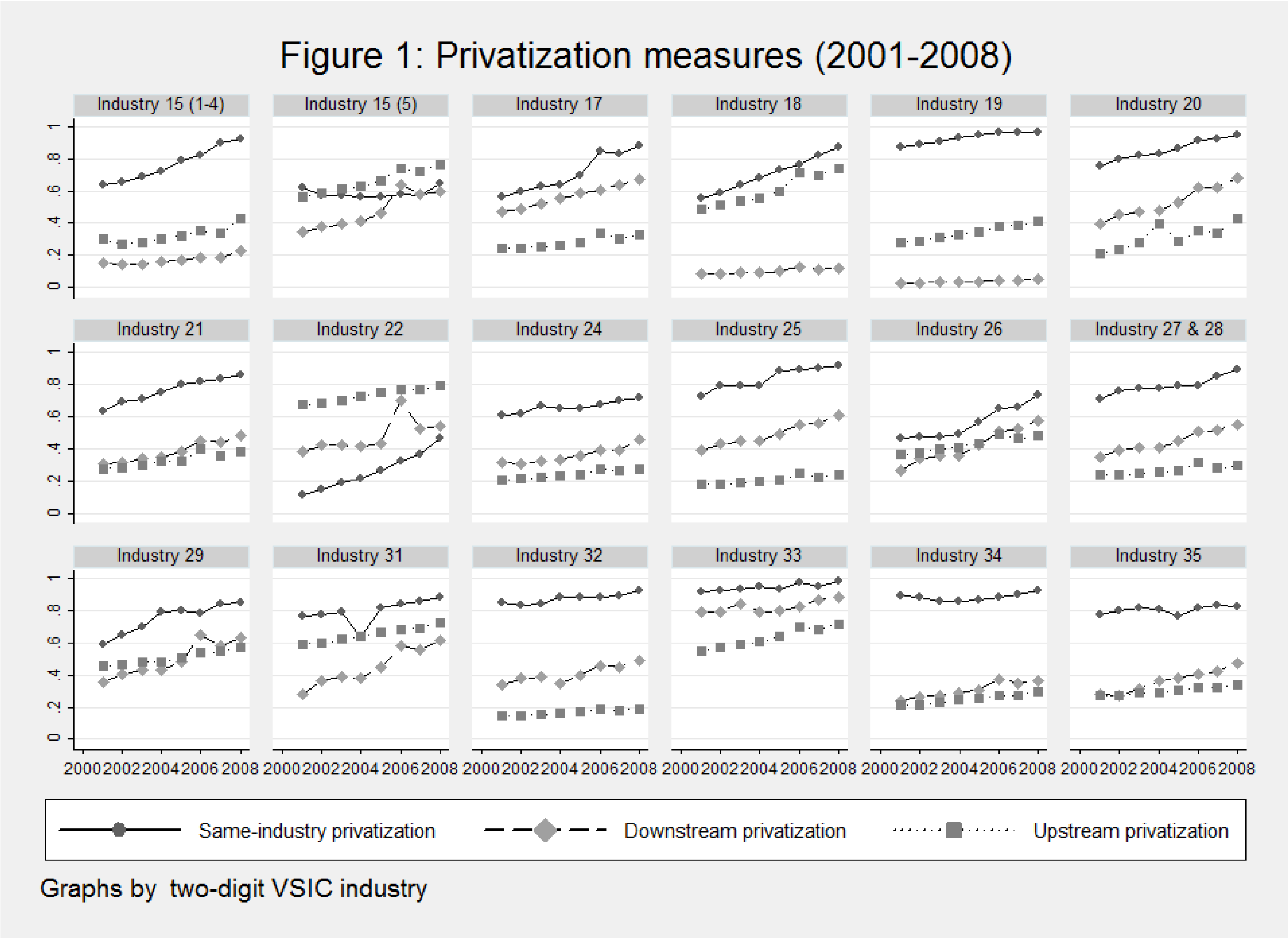


Figure 3.2: Manufacturing TFP decomposition (2001-2008)

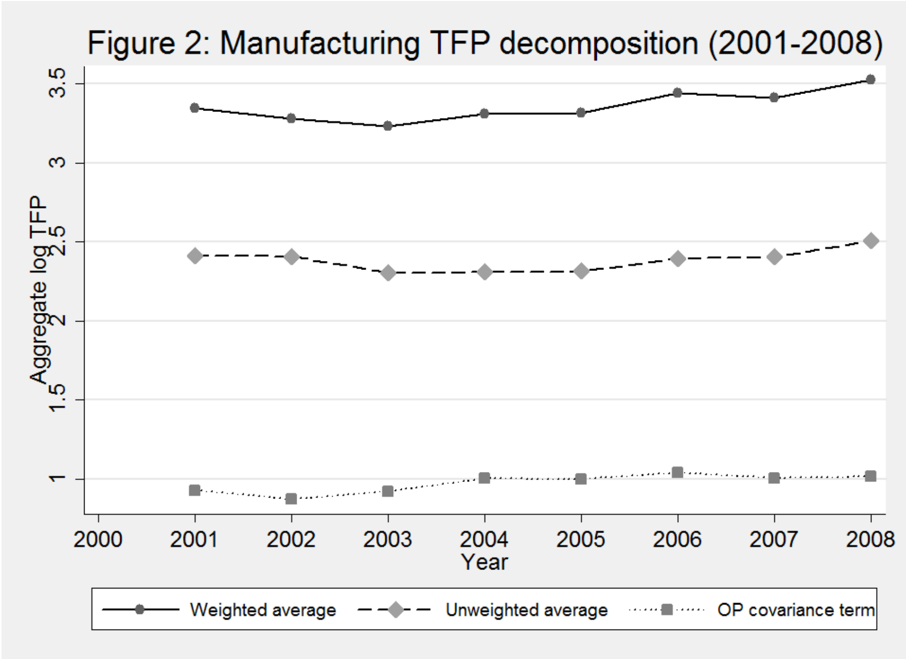


Figure 3.3: Manufacturing TFP (2001-2008)

Private vs. state-owned firms

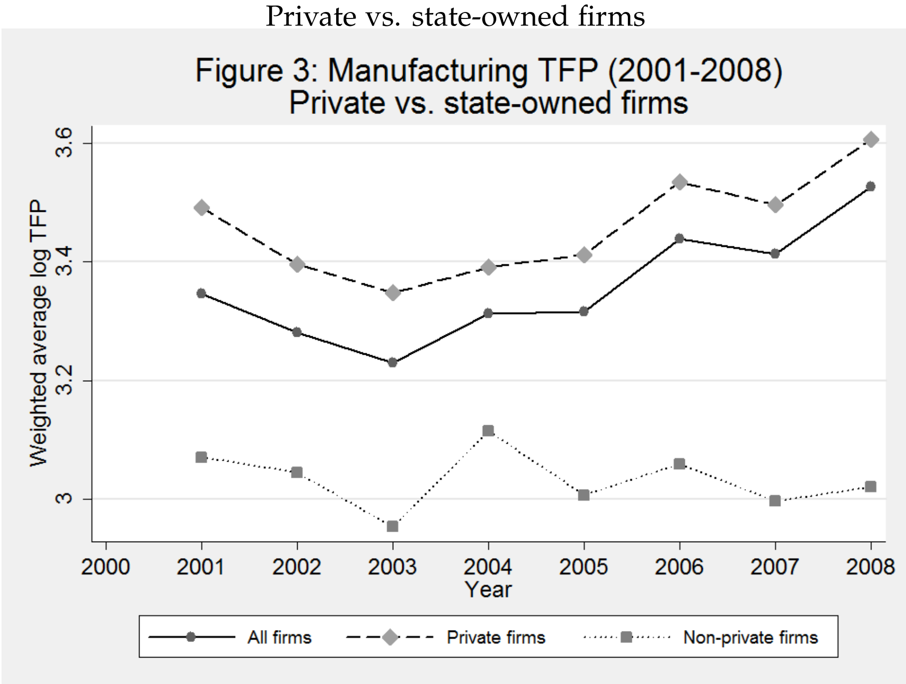


Figure 3.4: TFP evolution of privatized firms

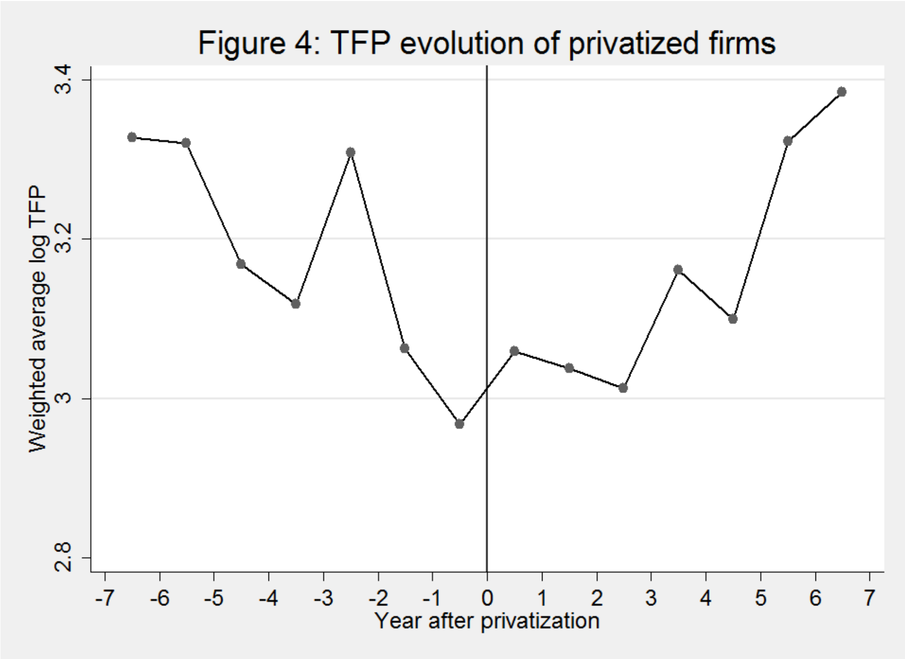


Table 3.1: Distribution of firms by industry (2001-2008)

VSIC code	Industry	No. of obser-vations	No. of firms	No. of firms privatized	Shares of private firms (by firm counts)		*Same-industry privatization*		*Downstream privatization*		*Upstream privatization*	
		2001-08	2001-08	2001-08	2001	2008	2001	2008	2001	2008	2001	2008
15 (1-4)	Food products	25,306	6,572	136	92.6%	97.8%	64.2%	92.1%	15.2%	22.5%	30.1%	42.9%
15 (5)	Beverages	7,019	2,247	40	87.7%	97.9%	62.3%	65.1%	34.2%	59.7%	56.5%	76.7%
17	Textiles	6,596	1,937	29	85.5%	97.5%	56.6%	87.9%	46.9%	67.4%	24.2%	32.6%
18	Wearing apparel	10,956	3,800	66	83.1%	98.5%	55.6%	87.6%	8.4%	11.8%	49.0%	74.0%
19	Leather and related products	3,466	989	19	85.5%	97.4%	87.3%	96.9%	2.4%	5.3%	28.2%	41.1%
20	Wood and wood products (including furniture)	18,643	6,086	43	94.2%	99.2%	75.7%	94.5%	39.9%	68.2%	20.9%	43.1%
21	Pulp, paper, and paper products	6,280	1,786	32	92.2%	98.7%	63.7%	86.4%	30.5%	48.5%	27.9%	38.5%
22	Publishing, printing, and recorded media	8,194	2,761	46	54.0%	93.9%	11.8%	47.2%	38.2%	54.7%	67.4%	79.4%
24	Chemicals and chemical products	6,412	1,813	47	81.8%	96.7%	61.0%	71.6%	31.5%	45.9%	20.7%	27.8%
25	Rubber and plastic products	9,655	2,837	25	93.5%	99.0%	73.2%	92.1%	39.3%	60.8%	18.1%	24.0%
26	Other non-metallic mineral products	11,790	3,031	140	82.9%	96.1%	47.1%	73.4%	27.0%	58.1%	36.7%	48.1%
27 & 28	Basic metals and fabricated metal products	18,199	6,035	48	89.6%	98.6%	71.5%	89.2%	34.9%	55.4%	24.2%	29.6%
29	Machinery and equipment	4,023	1,246	22	73.9%	96.1%	59.0%	85.3%	35.5%	63.1%	45.3%	57.2%
31	Electrical equipment and apparatus	2,113	551	13	85.2%	95.8%	76.3%	88.2%	28.4%	61.5%	59.1%	72.7%
32	Radio, television, and communication equipment	1,216	382	3	75.6%	94.6%	84.7%	92.2%	33.9%	49.0%	14.4%	19.1%
33	Medical, precision, and optical instruments	588	175	4	83.3%	97.7%	91.4%	98.2%	79.4%	88.5%	54.5%	71.6%
34	Motor vehicles	1,515	396	7	81.8%	95.1%	89.1%	92.8%	24.2%	36.7%	21.0%	29.6%
35	Other transport equipment	3,409	901	24	73.2%	90.7%	77.8%	82.7%	28.2%	47.6%	26.8%	34.1%
	Total	**145,380**	**43,545**	**744**	**86.9%**	**97.6%**	**66.1%**	**86.8%**	**25.6%**	**44.1%**	**31.1%**	**41.3%**

Notes: Private firms are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. All same-industry, downstream, and upstream privatization measures are calculated for each industry x year using this definition of private firms. Industries 16, 23, and 30 are not included in due to small sample size (i.e. insufficient to run Olley-Pakes estimation).

Table 3.2: Summary statistics for key variables (2001-2008)

Variable	Mean	Standard deviation	Variable	Mean	Standard deviation
Log Y	8.63	2.12	*Same-industry privatization (private firms)*	75.5%	16.1%
Log K	7.04	2.14	*Downstream privatization (private firms)*	40.0%	18.3%
Log W	6.41	1.83	*Upstream privatization (private firms)*	39.7%	16.6%
Log TFP	2.40	1.14	*Same-industry privatization (domestic private firms)*	38.7%	14.7%
HHI	0.02	0.02	*Downstream privatization (domestic private firms)*	25.3%	14.3%
Log demand	18.35	0.67	*Upstream privatization (domestic private firms)*	20.4%	8.1%

Notes: Private firms are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. Domestic private firms are defined as those with at least 50% domestic private ownership and no foreign ownership. Same-industry, downstream, and upstream privatization measures reported above are calculated using the corresponding definition of private firms or domestic private firms. Summary statistics are calculated for 145,380 firm x year observations.

Table 3.3: Additional summary statistics for privatization measures (2001-2008)

	Same-industry privatization		*Downstream privatization*		*Upstream privatization*	
Year	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
2001	67.1%	18.7%	32.1%	16.2%	34.7%	16.3%
2002	69.2%	18.3%	34.9%	16.7%	35.5%	16.9%
2003	71.2%	17.6%	36.6%	17.6%	37.2%	17.2%
2004	72.1%	17.7%	37.0%	16.9%	39.1%	17.4%
2005	75.8%	16.5%	40.2%	17.4%	40.3%	18.4%
2006	79.1%	15.5%	47.9%	20.2%	44.9%	19.2%
2007	81.2%	15.1%	46.6%	19.8%	43.6%	19.6%
2008	84.6%	12.8%	50.3%	20.5%	46.8%	20.0%

Notes: Private firms are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. All same-industry, downstream, and upstream privatization measures are calculated for each industry x year using this definition of private firms. Summary statistics are calculated for 18 industries each year.

Table 3.4: Manufacturing TFP decomposition by firm type (2001-2008)

	2001	2002	2003	2004	2005	2006	2007	2008	Annual growth
All firms									
Weighted average log TFP	3.35	3.28	3.23	3.31	3.32	3.44	3.41	3.53	2.6%
Unweighted average log TFP	2.42	2.41	2.31	2.31	2.31	2.40	2.41	2.51	1.3%
OP covariance term	0.93	0.87	0.92	1.00	1.00	1.04	1.01	1.02	1.2%
State-owned firms									
Weighted average log TFP	3.07	3.05	2.95	3.12	3.01	3.06	3.00	3.02	-0.7%
Unweighted average log TFP	2.24	2.27	2.25	2.31	2.33	2.37	2.31	2.32	1.1%
OP covariance term	0.84	0.78	0.70	0.80	0.68	0.69	0.69	0.71	-1.9%
Private firms									
Weighted average log TFP	3.49	3.40	3.35	3.39	3.41	3.53	3.50	3.61	1.6%
Unweighted average log TFP	2.44	2.43	2.31	2.31	2.31	2.40	2.41	2.51	1.0%
OP covariance term	1.05	0.97	1.04	1.08	1.10	1.14	1.09	1.09	0.6%
Incumbent private firms									
Weighted average log TFP	3.52	3.40	3.34	3.40	3.42	3.54	3.50	3.61	1.3%
Unweighted average log TFP	2.52	2.47	2.37	2.37	2.37	2.40	2.42	2.50	-0.2%
OP covariance term	1.00	0.93	0.98	1.03	1.05	1.14	1.08	1.10	1.5%
Entrant private firms									
Weighted average log TFP	3.34	3.38	3.44	3.22	3.28	3.40	3.34	3.58	3.4%
Unweighted average log TFP	2.27	2.15	1.98	1.97	1.97	2.36	2.35	2.54	3.9%
OP covariance term	1.07	1.23	1.45	1.25	1.30	1.04	0.99	1.04	-0.5%
Private firm shares of all firms (by revenue)	65.6%	67.3%	70.3%	71.9%	76.1%	79.9%	83.5%	86.3%	3.0%
Entrant shares of private firms (by revenue)	14.7%	3.3%	3.4%	4.1%	2.3%	3.9%	4.8%	6.1%	0.5%

Notes: Private firms are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. Annual growth for entrant shares of private firms is calculated for the period of 2002-2008.

Table 3.5: Dynamic panel regressions at firm level

	(1)	(2)	(3)	(4)	(5)
Depedent variable			*Log TFP*		
Sample	Private firms	State-owned firms	Firms that got privatized	Private firms	Private firms
Lagged dependent variable	0.320***	0.308***	0.306***	0.305***	0.305***
	(0.0224)	(0.0435)	(0.0385)	(0.0242)	(0.0242)
Same-industry privatization	0.739*	-0.559	-0.261	-0.0331	-0.0385
	(0.436)	(0.444)	(0.370)	(0.301)	(0.289)
Downstream privatization	0.336***	-0.156	-0.0972	0.405***	0.362*
	(0.122)	(0.108)	(0.163)	(0.121)	(0.195)
Upstream privatization	-0.103	-0.578*	-0.189	-0.0943	-0.0724
	(0.225)	(0.328)	(0.314)	(0.226)	(0.224)
Private			0.0104		
			(0.0323)		
Log demand				0.266**	0.285***
				(0.109)	(0.110)
HHI				-0.512	-0.785*
				(0.419)	(0.406)
Downstream HHI					-0.247
					(0.289)
Upstream HHI					-0.486
					(0.353)
Year trend	Yes	Yes	Yes	Yes	Yes
Number of observations	64,868	4,875	4,637	64,868	64,868
Number of firms	20,194	1,371	882	20,194	20,194

Notes: Private firms are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. All same-industry, downstream, and upstream privatization measures are calculated for each industry x year using this definition of private firms. Standard error in parentheses are at clustered at industry x province level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

Table 3.6: Robustness checks for spillover from downstream privatization

	(1)	(2)	(3)	(4)	(5)	(6)
Depedent variable		*Log TFP*			*Log TFP (L)*	*Log TFP (M)*
Sample	Always private firms	Never private firms	Non-zero private-share firms	Non-state-controlled firms	Private firms	Private firms
Lagged dependent variable	0.305***	0.286***	0.303***	0.306***	0.361***	0.269***
	(0.0246)	(0.0588)	(0.0247)	(0.0238)	(0.0270)	(0.0240)
Same-industry privatization	0.0205	-1.486***	0.0527	0.0225	0.183	0.154
	(0.316)	(0.325)	(0.214)	(0.226)	(0.288)	(0.218)
Downstream privatization	0.422***	-0.0735	-0.0733	0.342***	0.402***	0.497***
	(0.124)	(0.138)	(0.112)	(0.118)	(0.118)	(0.157)
Upstream privatization	-0.0999	-0.333	0.297	-0.292	-0.400*	0.360**
	(0.235)	(0.428)	(0.207)	(0.283)	(0.237)	(0.148)
Log demand	0.272**	0.146	0.270**	0.252**	0.225*	0.0723
	(0.114)	(0.106)	(0.105)	(0.105)	(0.115)	(0.0928)
HHI	-0.511	-0.0963	-0.926**	-0.292	-0.361	0.304
	(0.439)	(1.106)	(0.397)	(0.482)	(0.431)	(0.405)
Year trend	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	62,129	2,977	66,198	64,481	63,645	53,896
Number of firms	19,414	754	20,521	20,136	20,036	16,783

Notes: Private firms in colums (1), (2), (5) and (6) are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. Private firms in column (3) are defined as those with non-zero domestic private ownership or non-zero foreign ownership. Private firms in column (4) are defined as those controlled by domestic private or foreign entities. Same-industry, downstream, and upstream privatization measures in each column are calculated for each industry x year using the corresponding defintion of private firms. Log TFP in column (5) is estimated using employment count instead of wage bill. Log TFP in column (6) is estimated using wage bill and imputed material costs. Standard error in parentheses are clustered at industry x province level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

Table 3.7: Domestic vs. foreign ownership and spillover

	(1)	(2)	(3)	(4)	(5)	(6)
Depedent variable			*Log TFP*			
Sample	Purely-domestic private firms	Foreign-affiliated firms	All private firms	Domestic-controlled private firms	Foreign-controlled firms	Non-state-controlled firms
Lagged dependent variable	0.320***	0.193***	0.296***	0.321***	0.187***	0.297***
	(0.0257)	(0.0282)	(0.0251)	(0.0252)	(0.0294)	(0.0252)
Same-industry privatization	-0.0554	0.0119		-0.0279	0.0913	
	(0.346)	(0.303)		(0.252)	(0.318)	
Downstream privatization	0.459***	-0.0970		0.434***	-0.247	
	(0.132)	(0.274)		(0.129)	(0.188)	
Upstream privatization	-0.0527	-0.359		-0.270	-0.545**	
	(0.255)	(0.249)		(0.322)	(0.262)	
Same-industry privatization (domestic)			1.124***			0.897**
			(0.390)			(0.363)
Downstream privatization (domestic)			0.231*			0.255*
			(0.137)			(0.138)
Upstream privatization (domestic)			-0.0304			-0.174
			(0.467)			(0.484)
Same-industry privatization (foreign)			-0.726**			-0.510**
			(0.319)			(0.259)
Downstream privatization (foreign)			-0.361			-0.460
			(0.783)			(0.312)
Upstream privatization (foreign)			-0.139			-0.383
			(0.189)			(0.253)
Log demand	0.279**	0.181**	0.0259	0.269**	0.150*	0.0270
	(0.124)	(0.0757)	(0.107)	(0.119)	(0.0775)	(0.107)
HHI	-0.590	-0.518	1.756***	-0.298	-0.541	1.561**
	(0.497)	(0.519)	(0.626)	(0.572)	(0.534)	(0.606)
Year trend	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	55,253	9,615	64,868	55,253	9,228	64,481
Number of firms	17,601	2,616	20,194	17,641	2,533	20,136

Notes: Private firms in column (3) are defined as (i) those with at least 50% domestic private ownership and no foreign ownership (i.e. purely-domestic private firms, column (1)), or (ii) those with non-zero foreign ownership (i.e. foreign-affiliated firms, column (2)). Overall, domestic, and foreign privatization measures in columns (1) to (3) are calculated for each industry x year using the above definitions of private firms, pure-domestic private firms, and foreign-affiliated private firms respectively. Non-state-controlled firms in column (6) are defined as (i) those controlled by domestic private entities (i.e. domestic-controlled private firms , column (4)), or (ii) those controlled by foreign entities (i.e. foreign-controlled firms, column (5)). Overall, domestic, and foreign privatization in columns (4) to (6) are calculated for each industry x year using the above definitions of non-state-controlled firms, domestic-controlled private firms, and foreign-controlled firms respectively. Standard error in parentheses are clustered at industry x province level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

Table 3.8: Spillover form downstream privatization in same vs. different markets

	(1)	(2)	(3)	(4)	(5)	(6)
Depedent variable	*Log TFP*					
Sample	Private firms					
Market definition	3 large regions			8 smaller regions		
Lagged dependent variable	0.306***	0.304***	0.306***	0.306***	0.308***	0.307***
	(0.0224)	(0.0226)	(0.0220)	(0.0231)	(0.0230)	(0.0228)
Same-industry privatization (same market)	-0.215		-0.133	-0.127		-0.0667
	(0.199)		(0.170)	(0.157)		(0.131)
Downstream privatization (same market)	0.539***		0.574**	1.124***		1.178***
	(0.174)		(0.224)	(0.230)		(0.270)
Upstream privatization (same market)	-0.193		-0.133	-0.146		-0.159
	(0.132)		(0.123)	(0.151)		(0.155)
Same-industry privatization (different markets)		-0.110	0.173		-0.140	0.169
		(0.261)	(0.249)		(0.200)	(0.195)
Downstream privatization (different markets)		0.200*	-0.0450		0.164	-0.122
		(0.118)	(0.176)		(0.126)	(0.166)
Upstream privatization (different markets)		-0.0140	0.0531		-0.0294	0.115
		(0.181)	(0.181)		(0.223)	(0.231)
Log demand (same market)	0.240***	0.221***	0.207***	0.112	0.125	0.0856
	(0.0759)	(0.0769)	(0.0776)	(0.0784)	(0.0766)	(0.0804)
HHI (same market)	-0.240	-0.302	-0.337	-0.250	-0.207	-0.168
	(0.376)	(0.389)	(0.381)	(0.419)	(0.394)	(0.430)
Year trend	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	64,868	64,868	64,868	64,868	64,868	64,868
Number of firms	20,194	20,194	20,194	20,194	20,194	20,194

Notes: Private firms are defined as those with at least 50% domestic private ownership or those with non-zero foreign ownership. Markets in columns (1) to (3) are defined based on three large regions (i.e. North, Central, South). Markets in columns (4) to (6) are defined based on 8 smaller regions (i.e. Red River Delta, North East, North West, North Central Coast, South Central Coast, Central Highlands, Mekong River Delta, South Each). Same-market and different-market privatization measures in each column are calculated for each industry x market x year using the above definition of private firms and the corresponding market definition. Standard error in parentheses are clustered at industry x province level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

Table 3.9: Spillover on intermediate-good-producing and import-competing industries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable				*Log TFP*				
Sample	Intermediate-good producing	Non inter.-good producing	All private firms	Import competing	Non import competing	All private firms	Inter.-good producing & import competing	All private firms
Lagged dependent variable	0.255***	0.327***	0.308***	0.237***	0.331***	0.306***	0.235***	0.310***
	(0.0175)	(0.0334)	(0.0238)	(0.0184)	(0.0320)	(0.0237)	(0.0210)	(0.0234)
Same-industry privatization	-0.548*	0.704*	-0.140	0.125	0.447	0.108	0.117	0.0608
	(0.315)	(0.380)	(0.307)	(0.235)	(0.411)	(0.307)	(0.259)	(0.323)
Downstream privatization	0.598***	-0.000710	0.154	0.501**	0.213	0.240*	1.465***	0.290
	(0.138)	(0.233)	(0.179)	(0.245)	(0.138)	(0.123)	(0.339)	(0.195)
Downstream privatization # Intermediate-good producing industry			0.360*					
			(0.209)					
Downstream privatization # Import competing industry						0.748***		
						(0.252)		
Downstream privatization # Non inter.-good producing # Import competing								-0.00450
								(0.417)
Downstream privatization # Intermediate-good producing # Non import competing								-0.0425
								(0.242)
Downstream privatization # Intermediate-good producing # Import competing								1.061***
								(0.303)
Upstream privatization	-0.463	0.141	-0.0432	-0.0970	0.0316	-0.0520	-0.852	-0.0313
	(0.390)	(0.228)	(0.223)	(0.254)	(0.267)	(0.225)	(0.583)	(0.220)
Log demand	0.355***	0.113	0.268**	0.684***	-0.0429	0.251**	0.653***	0.244**
	(0.134)	(0.163)	(0.110)	(0.131)	(0.115)	(0.109)	(0.184)	(0.111)
HHI	-0.154	-10.48***	-0.379	-0.594	9.221	-0.237	0.106	-0.0583
	(0.378)	(3.479)	(0.422)	(0.373)	(7.936)	(0.421)	(0.514)	(0.432)
Year trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	31,986	32,882	64,868	27,936	36,932	64,868	20,446	64,868
Number of firms	10,151	10,043	20,194	9,003	11,191	20,194	6,550	20,194

Notes: Private firms are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. All privatization measures are calculated for each industry x year using this definition of private firms. Intermediate-good-producing industries are those with above-median intermediate consumption shares of gross outputs; non-intermediate-good-producing industries are the remaining ones. Import-competing industries are those with above-median import shares of intermediate consumption; non-importing-competing industries are the remaining ones. Standard error in parentheses are clustered at industry x province level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

Table 3.10: Dynamic panel regression at industry x market level

	(1)	(2)	(3)	(4)	(5)	(6)
Sample used for aggregation	State-owned firms			Private firms		
Dependent variable	*Weighted average log TFP*	*Unweighted average log TFP*	*Olley-Pakes covariance (TFP, share)*	*Weighted average log TFP*	*Unweighted average log TFP*	*Olley-Pakes covariance (TFP, share)*
Lagged dependent variable	0.0757	0.346***	0.0704	0.241	0.424***	0.121
	(0.157)	(0.113)	(0.123)	-0.159	-0.0942	-0.164
Same-industry privatization	-0.564	-0.167	-0.312	0.12	-0.00855	0.0255
	(0.472)	(0.291)	(0.434)	-0.31	-0.155	-0.409
Downstream privatization	-0.544	-0.171	-0.325	-0.0201	0.771**	-0.66
	(0.714)	(0.508)	(0.384)	-0.427	-0.335	-0.428
Upstream privatization	0.226	0.625	-0.427	0.433	-0.307	0.669*
	(0.307)	(0.427)	(0.321)	-0.427	-0.378	-0.376
Log demand	-0.0871	-0.0377	-0.0478	0.159*	0.193**	-0.0781
	(0.133)	(0.124)	(0.142)	-0.0906	-0.0891	-0.118
HHI	-0.00730	1.312	0.484	2.691	-0.32	2.008
	(1.670)	(1.525)	(2.193)	-3.741	-0.889	-3.837
Year trend	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	307	307	286	324	324	321
Number of industry x market's	53	53	50	54	54	54

Notes: Markets are defined based on three large regions (i.e. North, Central, South). Private firms are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. All privatization measures are calculated for each industry x market x year using this definition of private firms. Aggregate TFP measures are calculated for each industry x market x year from log TFPs and revenue shares of firms in the relevant sample. Standard error in parentheses are clustered at industry x market level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

Table 3.11: Spillover on incumbent vs. entrant firms

	(1)	(2)	(3)	(4)	(5)	(6)
Sample used for aggregation	Incumbent private firms			Entrant private firms		
Dependent variable	*Weighted average log TFP*	*Unweighted average log TFP*	*Olley-Pakes covariance (TFP, share)*	*Weighted average log TFP*	*Unweighted average log TFP*	*Olley-Pakes covariance (TFP, share)*
Lagged dependent variable	0.334**	0.426***	0.426***	-0.294***	-0.0542	-0.000220
	(0.150)	(0.0901)	(0.0901)	(0.100)	(0.116)	(0.0783)
Same-industry privatization	-0.153	-0.193	-0.0876	-1.479	-0.792	-1.422
	(0.279)	(0.177)	(0.298)	(0.988)	(0.602)	(0.998)
Downstream privatization	0.319	0.524	-0.124	1.861*	1.767**	0.318
	(0.423)	(0.340)	(0.408)	(1.047)	(0.734)	(1.020)
Upstream privatization	0.162	0.000156	0.142	1.716*	-0.766	2.311**
	(0.453)	(0.308)	(0.312)	(1.009)	(0.678)	(0.945)
Log demand	0.130	0.145**	-0.0373	0.913**	0.766***	-0.00582
	(0.0933)	(0.0737)	(0.110)	(0.417)	(0.270)	(0.353)
HHI	3.379	-0.875	3.334	-0.823	-0.362	-3.438
	(3.902)	(0.955)	(3.892)	(4.166)	(2.125)	(4.396)
Year trend	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	322	322	320	301	301	279
Number of industry x market's	54	54	54	53	53	48

	(7)	(8)	(9)	(10)	(11)	(12)
Sample used for aggregation	Incumbent & entrant private firms					Ent. priv. firms
Dependent variable	*OP cov. (W.Avg., share)[a]*	*OP cov. (Uw.Avg., share)[b]*	*Difference in W.Avg. log TFP (Inc. – Ent.)*	*Difference in Uw.Avg. log TFP (Inc. – Ent.)*	*Difference in revenue share (Inc. – Ent.)*	*Log TFP*
Lagged dependent variable	-0.196	-0.0146				
	(0.128)	(0.110)				
Same-industry privatization	0.875**	0.390	1.010	0.245	0.109	0.813*
	(0.435)	(0.368)	(0.833)	(0.583)	(0.123)	(0.479)
Downstream privatization	-0.793*	-0.718*	-1.473	-1.652**	-0.277*	0.818**
	(0.454)	(0.376)	(0.935)	(0.720)	(0.157)	(0.415)
Upstream privatization	-0.803*	0.402	-1.352	1.066*	0.0855	0.485
	(0.463)	(0.330)	(1.045)	(0.582)	(0.150)	(0.756)
Log demand	-0.324*	-0.198	-0.792**	-0.465*	-0.133***	0.291
	(0.188)	(0.126)	(0.401)	(0.240)	(0.0512)	(0.230)
HHI	1.969	0.235	2.161	-1.306	0.0531	-1.516
	(3.219)	(1.089)	(7.839)	(2.673)	(0.502)	(1.178)
Year trend	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	301	301	301	301	301	25,067
Number of industry x market's	53	53	53	53	53	

Notes: Markets are defined based on three large regions (i.e. North, Central, South). Private firms are defined as those with at least 50% domestic private ownership or those with non-zero foreign ownership. All privatization measures are calculated for each industry x market x year using this definition of private firms. Aggregate TFP measures are calculated for each industry x market x year from log TFPs and revenue shares of firms in the relevant sample. Column (12) reports results from a firm-level OLS regression with industry x province fixed effects among entrant private firms. Standard error in parentheses are clustered at industry x market level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

[a] OP cov. (W.Avg., share) = (Incumbents' revenue share – Entrants' revenue share) * (Incumbents' weighted average log TFP – Entrants' weighted average log TFP)

[b] OP cov. (Uw.Avg., share) = (Incumbents' revenue share – Entrants' revenue share) * (Incumbents' unweighted average log TFP – Entrants' unweighted average log TFP)

Table 3.12: Provincial institution quality and spillover

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	*Log TFP*									
Sample	Less SOE bias	More SOE bias	All private firms	All private firms	Low entry costs	High entry costs	All private firms	All private firms	More bias & high costs	All private firms
Lagged dependent variable	0.308***	0.298***	0.307***	0.306***	0.302***	0.304***	0.307***	0.306***	0.321***	0.307***
	(0.0393)	(0.0196)	(0.0241)	(0.0242)	(0.0384)	(0.0180)	(0.0241)	(0.0242)	(0.0195)	(0.0241)
Same-industry privatization	0.352	-0.607	-0.0262	-0.00792	0.574	-0.766**	-0.0247	-0.0161	-0.911**	-0.0374
	(0.412)	(0.396)	(0.309)	(0.309)	(0.396)	(0.345)	(0.311)	(0.308)	(0.450)	(0.316)
Downstream privatization	0.424*	0.450***	0.556***	2.364***	0.143	0.628***	0.554***	2.926**	0.632***	0.644***
	(0.217)	(0.165)	(0.166)	(0.905)	(0.203)	(0.169)	(0.167)	(1.236)	(0.186)	(0.197)
Downstream privatization # Less SOE bias province			-0.404*							
			(0.212)							
Downstream privatization # Province SOE bias score				-0.300**						
				(0.135)						
Downstream privatization # Low entry cost province							-0.387*			
							(0.217)			
Downstream privatization # Province entry cost score								-0.336**		
								(0.164)		
Downstream privatization # More bias # Low costs										-0.555
										(0.385)
Downstream privatization # Less bias # High costs										-0.674**
										(0.310)
Downstream privatization # Less bias # Low costs										-0.467*
										(0.252)
Upstream privatization	-0.119	0.00866	-0.0729	-0.0770	0.0533	-0.170	-0.0727	-0.0791	-0.115	-0.0624
	(0.282)	(0.306)	(0.223)	(0.224)	(0.271)	(0.347)	(0.223)	(0.225)	(0.424)	(0.223)
Log demand	-0.00896	0.560***	0.266**	0.265**	0.0835	0.525***	0.267**	0.266**	0.612***	0.267**
	(0.134)	(0.148)	(0.109)	(0.109)	(0.131)	(0.158)	(0.109)	(0.109)	(0.191)	(0.109)
HHI	-0.0464	-0.832*	-0.493	-0.510	-0.445	-0.527	-0.482	-0.495	-0.630	-0.479
	(0.646)	(0.481)	(0.413)	(0.418)	(0.640)	(0.455)	(0.412)	(0.417)	(0.574)	(0.411)
Year trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	32,623	32,245	64,868	64,868	35,564	29,304	64,868	64,868	23,798	64,868
Number of firms	9,660	10,534	20,194	20,194	10,355	9,839	20,194	20,194	8,068	20,194

Notes: Private firms are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. All privatization measures are calculated for each industry x year using this definition of private firms. Less-SOE-bias provinces are those with above-median PCI SOE bias scores; more-SOE-bias provinces are the remaining ones. Low-entry-cost provinces are those with above-median PCI entry cost scores; high-entry-cost provinces are the remaining ones. Standard error in parentheses are clustered at industry x province level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

Bibliography

Acemoglu, Daron and Joshua Linn, "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry," *Quarterly Journal of Economics*, 2004, *119* (3), 1049–1090.

__ , **Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti**, "Technology, Information, and the Decentralization of the Firm," *Quarterly Journal of Economics*, 2007, *122* (4), 1759–1799.

__ , **Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr**, "Innovation, Reallocation, and Growth," *American Economic Review*, 2018, *108* (11), 3450–3491.

Ackerberg, Daniel, Lanier Benkard, Steven Berry, and Ariel Pakes, "Econometric Tools for Analyzing Market Outcomes," in James Heckman and Edward Leamer, eds., *Handbook of Econometrics*, Vol. 6, Part A of *Handbook of Econometrics*, Elsevier, 2007, pp. 4171–4276.

Aghion, Philippe and Jean Tirole, "Formal and Real Authority in Organizations," *Journal of Political Economy*, 1997, *105* (1), 1–29.

__ **and Peter Howitt**, "A Model of Growth through Creative Destruction," *Econometrica*, 1992, *60* (2), 323–351.

__ , **John Van Reenen, and Luigi Zingales**, "Innovation and Institutional Ownership," *American Economic Review*, 2013, *103* (1), 277–304.

__ , **Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt**, "Competition and Innovation: An Inverted-U Relationship," *Quarterly Journal of Economics*, 2005, *120* (2), 701–728.

__ , **Richard Blundell, Rachel Griffith, Peter Howitt, and Susanne Prantl**, "The Effects of Entry on Incumbent Innovation and Productivity," *Review of Economics and Statistics*, 2009, *91* (1), 20–32.

Agrawal, Ajay, Carlos Rosell, and Timothy Simcoe, "Do Tax Credits Affect R&D Expenditures by Small Firms? Evidence from Canada," 2014. Working Paper.

Ahern, Kenneth, Daniele Daminelli, and Cesare Fracassi, "Lost in Translation? The Effect of Cultural Values on Mergers around the World," *Journal of Financial Economics*, 2015, *117* (1), 165–189.

Akcigit, Ufuk, Douglas Hanley, and Stephanie Stantcheva, "Optimal Taxation and R&D Policy," 2017. Working Paper.

_ , **John Grigsby, Tom Nicholas, and Stephanie Stantcheva**, "Taxation and Innovation in the 20th Century," 2018. Working Paper.

_ , **Salome Baslandze, and Stefanie Stantcheva**, "Taxation and the International Mobility of Inventors," *American Economic Review*, 2016, *106* (10), 2930–2981.

Alesina, Alberto and George-Marios Angeletos, "Fairness and Redistribution," *American Economic Review*, 2005, *95* (4), 960–980.

Algan, Yann and Pierre Cahuc, "Inherited Trust and Growth," *American Economic Review*, December 2010, *100* (5), 2060–2092.

_ **and** _ , "Trust and Growth," *Annual Review of Economics*, 2013, *5* (1), 521–549.

_ **and** _ , "Trust, Growth, and Well-Being: New Evidence and Policy Implications," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 2, Amsterdam: Elsevier, 2014, pp. 49–120.

Angrist, Joshua, "The Perils of Peer Effects," *Labour Economics*, 2014, *30*, 98–108.

_ **and Guido Imbens**, "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association*, 1995, *90* (430), 431–442.

_ **and Jörn-Steffen Pischke**, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, New Jersey: Princeton University Press, 2009.

Arellano, Manuel and Olympia Bover, "Another Look at the Instrumental Variable Estimation of Error-Components Models," *Journal of Econometrics*, 1995, *68* (1), 29–51.

_ **and Stephen Bond**, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 1991, *58* (2), 277–297.

Arnold, Jens, Beata Javorcik, and Aaditya Mattoo, "Does Services Liberalization Benefit Manufacturing Firms?: Evidence from the Czech Republic," *Journal of International Economics*, 2011, *85* (1), 136–146.

Arora, Ashish, Marco Ceccagnoli, and Wesley Cohen, "R&D and the Patent Premium," *International Journal of Industrial Organization*, 2008, *26* (5), 1153–1179.

Arrow, Kenneth, "Economic Welfare and the Allocation of Resources for Invention," in National Bureau of Economic Research, ed., *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton, New Jersey: Princeton University Press, 1962, pp. 609–626.

_ , "Gifts and Exchanges," *Philosophy and Public Affairs*, 1972, *1* (4), 343–362.

_ , *The Limits of Organizations*, New York: Norton, 1974.

Aw, Bee Yan, Mark Roberts, and Daniel Yi Xu, "R&D Investments, Exporting, and the Evolution of Firm Productivity," *American Economic Review*, 2008, *98* (2), 451–456.

Azoulay, Pierre, Joshua Graff Zivin, and Gustavo Manso, "Incentives and Creativity: Evidence from the Academic Life Sciences," *RAND Journal of Economics*, 2011, *42* (3), 527–554.

— , — , **Danielle Li, and Bhaven Sampat**, "Public R&D Investments and Private-Sector Patenting: Evidence from NIH Funding Rules," *Review of Economic Studies*, 2015, *86* (1), 117–152.

Bai, Jie, Seema Jayachandran, Edmund Malesky, and Benjamin Olken, "Does Economic Growth Reduce Corruption? Theory and Evidence from Vietnam," 2013. Working Paper.

Baker, George, Robert Gibbons, and Kevin Murphy, "Informal Authority in Organizations," *Journal of Law, Economics, and Organization*, 1999, *15* (1), 56–73.

— , — , **and** — , "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics*, 2002, *117* (1), 39–84.

Bandiera, Oriana, Luigi Guiso, Andrea Prat, and Raffaella Sadun, "Matching Firms, Managers, and Incentives," *Journal of Labor Economics*, 2015, *33* (3), 623–681.

— , **Stephen Hansen, Andrea Prat, and Raffaella Sadun**, "CEO Behavior and Firm Performance," 2017. Working Paper.

Banfield, Edward, *The Moral Basis of a Backward Society*, New York: Free, 1958.

Barro, Robert and Rachel McCleary, "Religion and Economic Growth across Countries," *American Sociological Review*, 2003, *68* (5), 760–781.

— **and** — , *Religion and Economy*, Princeton, New Jersey: Princeton University Press, 2018.

Barsh, Joanna, Marla Capozzi, and Jonathan Davidson, "How Companies Approach Innovation: A McKinsey Global Survey," *McKinsey Quarterly*, 2007.

— , — , **and** — , "Leadership and Innovation," *McKinsey Quarterly*, 2008.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, "Cross-Country Differences in Productivity: The Role of Allocation and Selection," *American Economic Review*, 2013, *103* (1), 305–334.

Becker, Bettina, "Public R&D Policies and Private R&D Investment: A Survey of the Empirical Evidence," *Journal of Economic Surveys*, 2015, *29* (5), 917–942.

Becker, Sascha and Ludger Woessmann, "Was Weber Wrong? A Human Capital Theory of Protestant Economic History," *Quarterly Journal of Economics*, 2009, *124* (2), 531–596.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen, "Who Becomes an Inventor in America? The Importance of Exposure to Innovation," 2018. Working Paper.

Bennedsen, Morten, Francisco Pérez-González, and Daniel Wolfenzon, "Do CEOs Matter?," 2010. Working Paper.

Bennis, Warren and Burt Nanus, *Leaders: Strategies for Taking Charge*, New York: Harper Business, 1997.

Bertrand, Marianne, "CEOs," *Annual Review of Economics*, 2009, *1* (1), 121–150.

— **and Antoinette Schoar**, "Managing with Style: The Effect of Managers on Firm Policies," *Quarterly Journal of Economics*, 2003, *118* (4), 1169–1208.

Bérubé, Charles and Pierre Mohnen, "Are Firms That Receive R&D Subsidies More Innovative?," *Canadian Journal of Economics*, 2009, *42* (1), 206–225.

Bisin, Alberto and Thierry Verdier, "'Beyond the Melting Pot': Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits," *Quarterly Journal of Economics*, 2000, *115* (3), 955–988.

_ **and** _ , "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory*, April 2001, *97* (2), 298–319.

Bloom, Nicholas and John Van Reenen, "Patents, Real Options and Firm Performance," *Economic Journal*, 2002, *112* (478), C97–C116.

_ **and** _ , "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

_ **, Mark Schankerman, and John Van Reenen**, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

_ **, Rachel Griffith, and John Van Reenen**, "Do R&D Tax Credits Work? Evidence from a Panel of Countries 1979-1997," *Journal of Public Economics*, 2002, *85* (1), 1–31.

_ **, Raffaella Sadun, and John Van Reenen**, "The Organization of Firms Across Countries," *Quarterly Journal of Economics*, 2012, *127* (4), 1663–1705.

Blundell, Richard, Rachel Griffith, and John Van Reenen, "Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms," *Review of Economic Studies*, 1999, *66* (3), 529–554.

Bøler, Esther Ann, Andreas Moxnes, and Karen Ulltveit-Moe, "R&D, International Sourcing, and the Joint Impact on Firm Performance," *American Economic Review*, 2015, *105* (12), 3704–3739.

Bottazzi, Laura, Marco Da Rin, and Thomas Hellmann, "The Importance of Trust for Investment: Evidence from Venture Capital," *The Review of Financial Studies*, 2016, *29* (9), 2283–2318.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang, "Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing," *Journal of Development Economics*, 2012, *97* (2), 339–351.

_ **, Trevor Tombe, and Xiaodong Zhu**, "Factor Market Distortions across Time, Space and Sectors in China," *Review of Economic Dynamics*, 2013, *16* (1), 39–58.

Branstetter, Lee, "Exploring the Link between Academic Science and Industrial Innovation," *Annals of Economics and Statistics*, 2005, (79-80), 119–142.

_ **and Mariko Sakakibara**, "When Do Research Consortia Work Well and Why? Evidence from Japanese Panel Data," *American Economic Review*, 2002, *92* (1), 143–159.

Bronzini, Raffaello and Eleonora Iachini, "Are Incentives for R&D Effective? Evidence from a Regression Discontinuity Approach," *American Economic Journal: Economic Policy*, 2014, *6* (4), 100–134.

_ **and Paolo Piselli**, "The Impact of R&D Subsidies on Firm Innovation," *Research Policy*, 2016, *45* (2), 442–457.

Butler, Jeffrey, Paola Giuliano, and Luigi Guiso, "The Right Amount of Trust," *Journal of the European Economic Association*, 2016, *14* (5), 1155–1180.

Calonico, Sebastian, Matias Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Campbell, John, "Household Finance," *Journal of Finance*, 2006, *61* (4), 1553–1604.

Cappelen, Ådne, Arvid Raknerud, and Marina Rybalka, "The Effects of R&D Tax Credits on Patenting and Innovations," *Research Policy*, 2012, *41* (2), 334–345.

Card, David and Stefano DellaVigna, "What do Editors Maximize? Evidence from Four Leading Economics Journals," 2017. Working Paper.

Caselli, Francesco, "Accounting for Cross-Country Income Differences," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 1A, Amsterdam: Elsevier, 2005, pp. 679–741.

Cassiman, Bruno, Reinhilde Veugelers, and Pluvia Zuniga, "In Search of Performance Effects of (In)direct Industry Science Links," *Industrial and Corporate Change*, 2008, *17* (4), 611–646.

Chang, Andrew, "Tax Policy Endogeneity: Evidence from R&D Tax Credits," *Economics of Innovation and New Technology*, 2018, pp. 1–25.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu, "Notching R&D Investment with Corporate Income Tax Cuts in China," 2018. Working Paper.

Chuang, Yating and Laura Schechter, "Stability of Experimental and Survey Measures of Risk, Time, and Social Preferences: A Review and Some New Results," *Journal of Development Economics*, 2015, *117* (C), 151–170.

Cohen, Wesley, "Fifty Years of Empirical Studies of Innovative Activity and Performance," in Bronwyn Hall and Nathan Rosenberg, eds., *Handbook of The Economics of Innovation*, Vol. 1, Amsterdam: North-Holland, 2010, pp. 129–213.

Coleman, James, *Foundations of Social Theory*, Cambridge, Massachusetts: Harvard University Press, 1990.

Commander, Simon and Jan Svejnar, "Business Environment, Exports, Ownership, And Firm Performance," *Review of Economics and Statistics*, 2011, *93* (1), 309–337.

Czarnitzki, Dirk, Petr Hanel, and Julio Miguel Rosa, "Evaluating the Impact of R&D Tax Credits on Innovation: A Microeconometric Study on Canadian Firms," *Research Policy*, 2011, *40* (2), 217–229.

Dahl, Gordon B, Katrine V Løken, and Magne Mogstad, "Peer Effects in Program Participation," *American Economic Review*, 2014, *104* (7), 2049–2074.

David, Paul A, Bronwyn Hall, and Andrew Toole, "Is Public R&D a Complement or Substitute for Private R&D? A Review of the Econometric Evidence," *Research Policy*, 2000, *29* (4-5), 497–529.

DeBacker, Jason, Bradley Heim, and Anh N. Tran, "Importing Corruption Culture from Overseas: Evidence from Corporate Tax Evasion in the United States," *Journal of Financial Economics*, 2015, *117* (1), 122–138.

Dechezleprêtre, Antoine, Elias Einiö, Ralf Martin, Kieu-Trang Nguyen, and John Van Reenen, "Do Tax Incentives for Research Increase Innovation: An RD Design for R&D," 2018. Working Paper.

Dernis, Hélène, Mafini Dosso, Fernando Hervás, Valentine Millot, Mariagrazia Squicciarini, Antonio Vezzani et al., "World Corporate Top R&D Investors: Innovation and IP Bundles," 2015.

Djankov, Simeon and Peter Murrell, "Enterprise Restructuring in Transition: A Quantitative Survey," *Journal of Economic Literature*, 2002, *40* (3), 739–792.

Do, Quoc-Anh, Kieu-Trang Nguyen, and Anh N. Tran, "One Mandarin Benefits the Whole Clan: Hometown Favoritism in an Authoritarian Regime," *American Economic Journal: Applied Economics*, 2017, *9* (4), 1–29.

Doepke, Matthias and Fabrizio Zilibotti, "Culture, Entrepreneurship, and Growth," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 2, Amsterdam: Elsevier, 2014, pp. 1–48.

Doraszelski, Ulrich and Jaumandreu Jordi, "R&D and Productivity: Estimating Endogenous Productivity," *Review of Economic Studies*, 2013, *80* (4), 1338–1383.

Durlauf, Steven, Paul Johnson, and Jonathan Temple, "Growth Econometrics," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 1A, Amsterdam: Elsevier, 2005, pp. 555–677.

Ederer, Florian and Gustavo Manso, "Incentives for Innovation: Bankruptcy, Corporate Governance, and Compensation Systems," in Robert Litan, ed., *Handbook on Law, Innovation and Growth*, Edward Elgar Publishing, 2011, chapter 5.

_ and _ , "Is Pay for Performance Detrimental to Innovation?," *Management Science*, 2013, *59* (7), 1496–1513.

Einiö, Elias, "R&D Subsidies and Company Performance: Evidence from Geographic Variation in Government Funding Based on the ERDF Population-Density Rule," *Review of Economics and Statistics*, 2014, *96* (4), 710–728.

Ericson, Richard and Ariel Pakes, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 1995, *62* (1), 53–82.

Eriksson, Katherine and Zachary Ward, "The Ethnic Segregation of Immigrants in the United States from 1850 to 1940," 2018. Working Paper.

Estrin, Saul, Jan Hanousek, Evžen Kočcenda, and Jan Svejnar, "The Effects of Privatization and Ownership in Transition Economies," *Journal of Economic Literature*, 2009, *47* (3), 699–728.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global Evidence on Economic Preferences," 2018. Working Paper.

_ , _ , _ , David Huffman, and Uwe Sunde, "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences," 2016. Working Paper.

Fee, Edward, Charles Hadlock, and Joshua Pierce, "Managers with and without Style: Evidence Using Exogenous Variation," *Review of Financial Studies*, 2013, *26* (3), 567–601.

Fernández, Raquel, "Does Culture Matter?," in Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, eds., *Handbook of Social Economics*, Vol. 1A, Amsterdam: North-Holland, 2011, pp. 481–510.

_ **and Alessandra Fogli**, "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 146–77.

Foley, Fritz and William Kerr, "Ethnic Innovation and U.S. Multinational Firm Activity," *Management Science*, 2013, *59* (7), 1529–1544.

Fowkes, Rigmor K, Joao Sousa, and Neil Duncan, "Evaluation of Research and Development Tax Credit," Technical Report, HMRC 2015.

Fukuyama, Francis, *Trust: The Social Virtues and the Creation of Prosperity*, New York: Free, 1995.

Gambetta, Diego, "Can We Trust Trust?," in Diego Gambetta, ed., *Trust: Making and Breaking Cooperative Relations*, Oxford: University of Oxford, 1988, pp. 213–237.

Ganguli, Ina, "Saving Soviet Science: The Impact of Grants When Government R&D Funding Disappears," *American Economic Journal: Applied*, 2017, *9* (2), 165–201.

Gelman, Andrew and Guido Imbens, "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," 2014. Working Paper.

Giannetti, Mariassunta and Yishay Yafeh, "Do Cultural Differences Between Contracting Parties Matter? Evidence from Syndicated Bank Loans," *Management Science*, 2012, *58* (2), 365–383.

Gibbons, Robert and Robert Kaplan, "Formal Measures in Informal Management: Can a Balanced Scorecard Change a Culture?," *American Economic Review: Paper & Proceedings*, 2015, *105* (5), 447–451.

Giuliano, Paola, "Living Arrangements in Western Europe: Does Cultural Origin Matter?," *Journal of the European Economic Association*, 2007, *5* (5), 927–952.

_ **and Antonio Spilimbergo**, "Growing Up in a Recession: Beliefs and the Macroeconomy," *Review of Economic Studies*, 2013, *81* (2), 787–817.

Glaeser, Edward, David Laibson, Jose Scheinkman, and Christine Soutter, "Measuring Trust," *Quarterly Journal of Economics*, 2000, *115* (3), 811–846.

González-Uribe, Juanita and Moqi Groen-Xu, "CEO Contract Horizon and Innovation," 2017. Working Paper.

González, Xulia, Jordi Jaumandreu, and Consuelo Pazó, "Barriers to Innovation and Subsidy Effectiveness," *RAND Journal of Economics*, 2005, *36*, 930–950.

Goodridge, Peter, Jonathan Haskel, Alan Hughes, and George Wallis, "The Contribution of Public and Private R&D to UK Productivity Growth," 2015. Working Paper.

Goolsbee, Austan, "Does Government R&D Policy Mainly Benefit Scientists and Engineers?," *American Economic Review*, 1998, *88* (2), 298–302.

Görg, Holger and Eric Strobl, "The Effect of R&D Subsidies on Private R&D," *Economica*, 2007, *74* (294), 215–234.

Gorodnichenko, Yuriy and Gerard Roland, "Culture, Institutions, and the Wealth of Nations," *Review of Economics and Statistics*, 2017, *99* (3), 402–416.

_ , Jan Svejnar, and Katherine Terrell, "Globalization and Innovation in Emerging Markets," *American Economic Journal: Macroeconomics*, 2010, *2* (2), 194–226.

Gow, Ian, Steven Kaplan, David Larcker, and Anastasia Zakolyukina, "CEO Personality and Firm Policies," 2016. Working Paper.

Graham, John, Jillian Grennan, Campbell Harvey, and Shivaram Rajgopal, "Corporate Culture: Evidence from the Field," 2018. Working Paper.

Grennan, Jillian, "A Corporate Culture Channel: How Increased Shareholder Governance Reduces Firm Value," 2014. Working Paper.

Griliches, Zvi, "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *Bell Journal of Economics*, 1979, *10* (1), 92–116.

_ and Jacques Mairesse, "Production Functions: The Search for Identification," in S. Strøm, ed., *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, Cambridge, UK: Cambridge University Press, 1998.

Gruber, John, *Public Finance and Public Policy*, Vol. 25, Worth Publishers, 2011.

Guceri, Irem, "Will the Real R&D Employees Please Stand Up? Effects of Tax Breaks on Firm Level Outcomes," *International Tax and Public Finance*, 2018, *25* (1), 1–63.

_ and Li Liu, "Effectiveness of Fiscal Incentives for R&D: Quasi-Experimental Evidence," 2017. Working Paper.

Guellec, Dominique and Bruno Van Pottelsberghe de la Potterie, "From R&D to Productivity Growth: Do the Institutional Settings and the Source of Funds of R&D Matter?," *Oxford Bulletin of Economics and Statistics*, 2004, *66* (3), 353–378.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "The Role of Social Capital in Financial Development," *American Economic Review*, June 2004, *94* (3), 526–556.

_ , _ , and _ , "Does Culture Affect Economic Outcomes?," *Journal of Economic Perspectives*, June 2006, *20* (2), 23–48.

_ , _ , and _ , "Trusting the Stock Market," *The Journal of Finance*, 2008, *63* (6), 2557–2600.

_ , _ , and _ , "Cultural Biases in Economic Exchange?," *Quarterly Journal of Economics*, 2009, *124* (3), 1095–1131.

_ , _ , and _ , "The Value of Corporate Culture," *Journal of Financial Economics*, 2015, *117* (1), 60–76.

_ , _ , and _ , "Long-Term Persistence," *Journal of the European Economic Association*, 2016, *14* (6), 1401–1436.

Gurmu, Shiferaw and Fidel Pérez-Sebastián, "Patents, R&D and Lag Effects: Evidence from Flexible Methods for Count Panel Data on Manufacturing Firms," *Empirical Economics*, 2008, *35* (3), 507–526.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "Market Value and Patent Citations," *RAND Journal of Economics*, 2005, *36* (1), 16–38.

— **and John Van Reenen**, "How Effective are Fiscal Incentives for R&D? A Review of the Evidence," *Research Policy*, 2000, *29* (4), 449–469.

— **and Rosemarie Ziedonis**, "An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979-1995," *RAND Journal of Economics*, 2001, *32* (1), 101–128.

— **, Christian Helmers, Mark Rogers, and Vania Sena**, "The Choice between Formal and Informal Intellectual Property: A Review," *Journal of Economic Literature*, 2014, *52* (2), 375–423.

— **, Jacques Mairesse, and Pierre Mohnen**, "Measuring the Returns to R&D," in Bronwyn Hall and Nathan Rosenberg, eds., *Handbook of the Economics of Innovation, Volume 2*, Vol. 2 of *Handbook of the Economics of Innovation*, North-Holland, 2010, pp. 1033–1082.

— **, Zvi Griliches, and Jerry Hausman**, "Patents and R&D: Is There a Lag?," *International Economic Review*, 1986, *27* (2), 265–284.

Hall, Browyn, Christian Helmers, Mark Rogers, and Vania Sena, "The Importance (or Not) of Patents to UK Firms," *Oxford Economic Papers*, 2013, *65* (3), 603–629.

Hall, Robert and Charles Jones, "Why do Some Countries Produce So Much More Output Per Worker than Others?," *Quarterly Journal of Economics*, 1999, *114* (1), 83–116.

— **and Dale Jorgenson**, "Tax Policy and Investment Behavior," *American Economic Review*, 1967, *57* (3), 391–414.

Harhoff, Dietmar, Francis Narin, Frederic Scherer, and Katrin Vopel, "Citation Frequency and the Value of Patented Inventions," *Review of Economics and Statistics*, 1999, *81* (3), 511–515.

— **, Frederic Scherer, and Katrin Vopel**, "Citations, Family Size, Opposition and the Value of Patent Rights," *Research Policy*, 2003, *32* (8), 1343–1363.

HMRC, "Research and Development Tax Credit Statistics," Technical Report, London 2014.

HMSO, *Finance Act*, London: HMSO, 2000.

— , *Finance Act*, London: HMSO, 2002.

— , *Finance Act*, London: HMSO, 2007.

Hofstede, Gerrt, Geert Jan Hofstede, and Michael Minkov, *Cultures and Organizations: Software of the Mind*, London: McGraw-Hill, 1991.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey Rosen, "Estimating Vector Autoregressions with Panel Data," *Econometrica*, 1988, *56* (6), 1371–1395.

Hopenhayn, Hugo, "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), 1127–1150.

Howell, Sabrina, "Financing Innovation: Evidence from R&D Grants," *American Economic Review*, April 2017, *107* (4), 1136–1164.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Imbens, Guido and Karthik Kalyanaraman, "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *Review of Economic Studies*, 2011, *79* (3), 933–959.

Jacob, Brian and Lars Lefgren, "The Impact of Research Grant Funding on Scientific Productivity," *Journal of Public Economics*, 2010, *79* (3), 1168–1177.

Jaffe, Adam, "Technological Opportunities and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value," *American Economic Review*, 1986, *95* (9-10), 1168–1177.

_ **and Trinh Le**, "The Impact of an R&D Subsidy on Innovation: A Study of New Zealand Firms," 2015. Working Paper.

_ **, Manuel Trajtenberg, and Rebecca Henderson**, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 1993, *108* (3), 577–598.

Javorcik, Beata, "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages," *American Economic Review*, 2004, *94* (3), 605–627.

Jovanovic, Boyan, "Selection and the Evolution of Industry," *Econometrica*, 1982, *50* (3), 649–670.

Kaplan, Steven and Morten Sørensen, "Are CEOs Different? Characteristics of Top Managers," 2017.

_ **, Mark Klebanov, and Morten Sørensen**, "Which CEO Characteristics and Abilities Matter?," *The Journal of Finance*, 2012, *67* (3), 973–1007.

Kerr, William and William Lincoln, "The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention," *Journal of Labor Economics*, 2010, *28* (3), 473–508.

Klein Woolthuis, Rosalinde, Bas Hillebrand, and Bart Nooteboom, "Trust, Contract and Relationship Development," *Organization Studies*, 06 2005, *26*, 813–840.

Kleven, Henrick and Mazhar Waseem, "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *Quarterly Journal of Economics*, 2013, *128* (2), 669–723.

Knack, Stephen and Philip Keefer, "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *Quarterly Journal of Economics*, 1997, *112* (4), 1251–1288.

Knights, David, Faith Noble, Theo Vurdubakis, and Hugh Willmott, "Chasing Shadows: Control, Virtuality and the Production of Trust," *Organization Studies*, 2001, *22* (2), 311–336.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological Innovation, Resource Allocation, and Growth," *Quarterly Journal of Economics*, 2017, *132* (2), 665–712.

La Porta, Rafael, Florencio Lopez de Silanes, Andrei Shleifer, and Robert Vishny, "Trust in Large Organizations," *American Economic Review*, 1997, *87* (2), 333–38.

Lach, Saul, "Do R&D Subsidies Stimulate or Displace Private R&D? Evidence from Israel," *Journal of Industrial Economics*, 2002, *50*, 369–390.

— , **Zvika Neeman, and Mark Schankerman**, "Government Financing of R&D: A Mechanism Design Approach," 2017. Working Paper.

Lambert, Richard, "Executive Effort and Selection of Risky Projects," *RAND Journal of Economics*, 1986, *17* (1), 77–88.

Lanjouw, Jean, Ariel Pakes, and Jonathan Putnam, "How to Count Patents and Value Intellectual Property: The Uses of Patent Renewal and Application Data," *Journal of Industrial Economics*, 1998, *46* (4), 405–432.

Lee, David, "Randomized Experiments from Non-Random Selection in U.S. House Elections," *Journal of Econometrics*, 2008, *142* (2), 675–697.

— **and Thomas Lemieux**, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2010, *48* (2), 281–355.

Lerner, Josh, "The Importance of Patent Scope: An Empirical Analysis," *RAND Journal of Economics*, 1994, *25* (2), 319–333.

Lin, Ping and Kamal Saggi, "Multinational Firms, Exclusivity, and Backward Linkages," *Journal of International Economics*, 2007, *71* (1), 206–220.

Liu, Xiaoding, "Cultural Origins and CEO Incentive Contracts," 2013. Working Paper.

— , "Corruption Culture and Corporate Misconduct," *Journal of Financial Economics*, 2016, *122* (2), 307–327.

Macchiavello, Rocco, "Development Uncorked: Reputation Acquisition in the New Market for Chilean Wines in the UK," 2010. Working Paper.

Malesky, Edmund, "The Vietnam Provincial Competitiveness Index 2006," *Vietnam Competitiveness Initiative Policy Papers*, 2006, *11.* Available at SSRN: `http://ssrn.com/abstract=1669715.`

— , "Straight Ahead on Red: How Foreign Firect Investment Empowers Subnational Leaders," *The Journal of Politics*, 2008, *70* (01), 97–119.

Malmendier, Ulrike and Geoffrey Tate, "CEO Overconfidence and Corporate Investment," *Journal of Finance*, 2005, *60* (6), 2661–2700.

— **and** — , "Superstar CEOs," *Quarterly Journal of Economics*, 2009, *124* (4), 1593–1638.

Manski, Charles, "Identification of Endogenous Social effects: The Reflection Problem," *Review of Economic Studies*, 1993, *60* (3), 531–542.

Manso, Gustavo, "Motivating Innovation," *Journal of Finance*, 2011, *66* (5), 1823–1860.

March, James, "Exploration and Exploitation in Organizational Learning," *Organization Science*, 1991, *2* (1), 71–87.

Martinez, Elizabeth, Nancy Beaulieu, Robert Gibbons, Peter Pronovost, and Thomas Wang, "Organizational Culture and Performance," *American Economic Review: Papers & Proceedings*, 2015, *105* (5), 331–335.

McCaig, Brian and Nina Pavcnik, "Export Markets and Labor Allocation in a Poor Country," 2012. Working Paper.

Megginson, William and Jeffry Netter, "From State to Market: A Survey of Empirical Studies on Privatization," *Journal of Economic Literature*, 2001, *39* (2), 321–389.

Moretti, Enrico and Daniel Wilson, "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists," *American Economic Review*, 2017, *107* (7), 1858–1903.

__ , **Claudia Steinwender, and John Van Reenen**, "The Intellectual Spoils of War? Defense R&D, Productivity and International Technology Spillovers," 2016. Working Paper.

Moser, Petra, Joerg Ohmstedt, and Paul Rhode, "Patent Citations and the Size of the Inventive Step: Evidence from Hybrid Corn," 2015. Working Paper.

Nguyen, Bang Dang and Kasper Nielsen, "The Value of Independent Directors: Evidence from Sudden Deaths," *Journal of Financial Economics*, 2010, *98* (3), 550–567.

Nickell, Stephen John, "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 1981, *49* (6), 1417–26.

Nooteboom, Bart and Erik Stam, *Micro-foundations for Innovation Policy*, Amsterdam: Amsterdam University Press, 2008.

Nunn, Nathan, "Culture and the Historical Process," *Economic History of Developing Regions*, 2012, *27* (1), 108–126.

OECD, *Survey of R&D tax credits*, Paris: OECD, 2013.

__ , *OECD Science, Technology and Industry Outlook*, Paris: OECD, 2014.

Olley, Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

O'Reilly, Charles, David Caldwell, Jennifer Chatman, and Bernadette Doerr, "The Promise and Problems of Organizational Culture: CEO Personality, Culture, and Firm Performance," *Group* & *Organization Management*, 2014, *39* (6), 595–625.

O'Reilly, Charles, Jennifer Chatman, and David Caldwell, "People and Organizational Culture: A Profile Comparison Approach to Assessing Person-Organzation Fit," *Academy of Management Journal*, 1991, *34* (3), 487–516.

Pakes, Ariel and Richard Ericson, "Empirical Implications of Alternative Models of Firm Dynamics," *Journal of Economic Theory*, 1998, *79* (1), 1–45.

Pan, Yihui, Stephan Siegel, and Tracy Yue Wang, "Corporate Risk Culture," *Journal of Financial and Quantitative Analysis*, 2017, *52* (06), 2327–2367.

__ , __ , **and** __ , "The Cultural Origin of CEOs' Attitudes Towards Uncertainty: Evidence from Corporate Acquisitions," 2018. Working Paper.

Putnam, Robert, *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon and Schuster, 2000.

__ , **Robert Leonardi, and Raffaella Nanetti**, *Making Democracy Work*, Princeton, New Jersey: Princeton University Press, 1993.

Rajan, Raghuram and Luihi Zingales, "Financial Dependence and Growth," *American Economic Review*, 1998, *88* (3), 559–586.

Rao, Nipurama, "Do Tax Credits Stimulate R&D Spending? Revisiting the Effect of the R&D Tax Credit in Its First Decade," *Journal of Public Economics*, 2016, *140*, 1–12.

Rodriguez-Clare, Andres, "Multinationals, Linkages, and Economic Development," *American Economic Review*, 1996, *86* (4), 852–73.

Roland Berger Market Research, "The Cost of a Sample European Patent," Technical Report, Munich.

Romer, Paul, "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), 71–102.

Sapienza, Paola, Anna Toldra Simats, and Luigi Zingales, "Understanding Trust," *Economic Journal*, 2013, *123* (12), 1313–1332.

Schein, Edgar, *Organization Culture and Leadership*, San Francisco, California: Jossey-Bass Publishers, 1985.

Schmoch, Ulrich, "Concept of a Technology Classification for Country Comparisons," in "Final Report to the World Intellectual Property Organisation (WIPO)" 2008.

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick, "Capitalists in the 21st Century," 2017. Working Paper.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Growing Like China," *American Economic Review*, 2011, *101* (1), 196–233.

Spenkuch, Jörg, "Religion and Work: Micro Evidence from Contemporary Germany," *Journal of Economic Behavior & Organization*, 2017, *135* (C), 193–214.

Spolaore, Enrico and Romain Wacziarg, "Ancestry, Language and Culture," in Victor Ginsburgh and Shlomo Weber, eds., *The Palgrave Handbook of Economics and Language*, London: Palgrave Macmillan, 2016, pp. 174–211.

Squicciarini, Mariagrazia, Hélène Dernis, and Chiara Criscuolo, "Measuring Patent Quality: Indicators of Technological and Economic Value," 2013. Working Paper.

Sull, Donald, "Measuring Corporate Culture Using Glassdoor Reviews," MIT Organizational Economics Lunch Seminar 2018.

Sutton, John, "Quality, Trade and the Moving Window: The Globalisation Process," *The Economic Journal*, 2007, *117* (524), F469–F498.

Syverson, Chad, "What Determines Productivity?," *Journal of Economic Literature*, 2011, *49* (2), 326–365.

Tabellini, Guido, "The Scope of Cooperation: Values and Incentives*," *Quarterly Journal of Economics*, 2008, *123* (3), 905–950.

_ , "Culture and Institutions: Economic Development in the Regions of Europe," *Journal of the European Economic Association*, 2010, *8* (4), 677–716.

Takalo, Tuomas, Tanja Tanayama, and Otto Toivanen, "Estimating the Benefits of Targeted R&D Subsidies," *Review of Economics and Statistics*, 2013, *95*, 255–272.

Tian, Xuan and Tracy Yue Wang, "Tolerance for Failure and Corporate Innovation," *Review of Financial Studies*, 2014, *27* (1), 211–255.

Tibshirani, Robert, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society*, 1996, *58* (1), 267–288.

Tirole, Jean, "A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality)," *Review of Economic Studies*, 1996, *63* (1), 1–22.

Trajtenberg, Manuel, "A Penny for Your Quotes: Patent Citations and the Value of Innovation," *RAND Journal of Economics*, 1990, *21* (1), 172–187.

_ , **Rebecca Henderson, and Adam Jaffe**, "University Versus Corporate Patents: A Window On The Basicness Of Invention," *Economics of Innovation and New Technology*, 1997, *5* (1), 19–50.

Tran, Anh N., "The Hidden Hand of Private Ownership: Tax Avoidance in 3,650 Privatized Vietnamese Enterprises," 2013. Working Paper.

Van den Steen, Eric, "On the Origin of Shared Beliefs (and Corporate Culture)," *RAND Journal of Economics*, 2010, *41* (4), 617–648.

Wallsten, Scott, "The Effects of Government-Industry R&D Programs on Private R&D: The Case of the Small Business Innovation Research Program," *RAND Journal of Economics*, 2000, *31*, 82–100.

Weber, Max, *The Protestant Ethic and the Spirit of Capitalism*, London: Allen & Unwin, 1905.

Wilson, Daniel, "Beggar thy neighbor? The in–state, out–of–state and aggregate effects of R&D tax credits," *Review of Economics and Statistics*, 2009, *91* (2), 431–436.

Wooldridge, Jeffrey, "On Estimating Firm-Level Production Functions using Proxy Variables to Control for Unobservables," *Economics Letters*, 2009, *104* (3), 112–114.

Xu, Guo, "How Does Collective Reputation Affect Hiring? Selection and Sorting in an Online Labour Market," 2015. Working Paper.

Zwick, Eric and Mahon James, "Tax Policy and Heterogeneous Investment Behavior," *American Economic Review*, 2017, *107* (1), 217–248.

Appendices

Appendix A

Appendices to Chapter 1: Trust and Innovation within the Firm

A.1 Theory appendices

A.1.1 Proof of Proposition 1

Recall from subsection 1.2.1 that in period 1, the manager sets b^L and b^M to zero and chooses b^H to maximize her expected payoff from hiring a good researcher:

$$\int_0^{\bar{\pi}(D)} \left[s^M + V_2^P\right] d\pi + \int_{\bar{\pi}(D)}^1 \left\{\pi \left[s^H - b^H + V_2^P\right] + (1-\pi)\left[s^L + DV_2^P\right]\right\} d\pi \tag{A.1}$$

where $\bar{\pi}(D) = \frac{c+(1-D)V_2^A}{b^H+(1-D)V_2^A}$. To further reduce notation burden, I omit the outcome superscript H from b^H and the period subscript 2 from V_2^P and V_2^A for the rest of this subsection.

Expression (A.1)'s first order condition with respect to b is:

$$\frac{\partial\bar{\pi}}{\partial b}\left[s^M + V^P\right] - \frac{\partial\bar{\pi}}{\partial b}\bar{\pi}\left[s^H - b + V^P\right] - \frac{\partial\bar{\pi}}{\partial b}(1-\bar{\pi})(s^L + DV^P) - \int_{\bar{\pi}}^1 \pi d\pi \tag{A.2}$$

$$\iff \left(-\frac{\partial\bar{\pi}}{\partial b}\right)\left\{\bar{\pi}\left[s^H - b - s^L + (1-D)V^P\right] - \left[s^M - s^L + (1-D)V^P\right]\right\} - \int_{\bar{\pi}}^1 \pi d\pi \tag{A.3}$$

where $\left(-\frac{\partial\bar{\pi}}{\partial b}\right) = \frac{c+(1-D)V^A}{(b+(1-D)V^A)^2}$ and $\int_{\bar{\pi}}^1 \pi d\pi = \frac{1-\bar{\pi}^2}{2}$.

Notice that (i) $\left(-\frac{\partial\bar{\pi}}{\partial b}\right)$, (ii) $\bar{\pi}$, (iii) $\left[s^H - b - s^L + (1-D)V^P\right]$, and (iv) $\int_{\bar{\pi}}^1 \pi d\pi$ are decreasing in b. In addition, (v) $\bar{\pi}\left[s^H - b - s^L + (1-D)V^P\right] - \left[s^M - s^L + (1-D)V^P\right]$ is non-negative as any value of b that makes (v) negative cannot be the manager's optimal choice. The first order condition therefore is also decreasing in b in the relevant range of b. This implies that for a given set of parameters (i.e., s^L, s^M, s^H, c, D, and corresponding V^P, V^A), equation (A.3) has a unique solution b^* that maximizes the manager's expected payoff in equation (A.1), which induces the good researcher to explore when π is above threshold $\bar{\pi}^* = \frac{c+(1-D)V^A}{b^*+(1-D)V^A}$.

Comparative static of $\bar{\pi}^*$ with respect to $(1-D)V^P$. As equation (A.3) is decreasing in b and $(1-D)V^P$, its unique solution b^* is also decreasing in $(1-D)V^P$. It then follows that $\bar{\pi}^*$ is increasing in $(1-D)V^P$ (as $\bar{\pi}^*$ is decreasing in b^*).

Comparative static of $\bar{\pi}^*$ with respect to $(1-D)V^A$. Let's rewrite b and $\frac{\partial\bar{\pi}}{\partial b}$ in terms of $\bar{\pi}$:

$$b = \frac{c+(1-D)V^A}{\bar{\pi}} - (1-D)V^A \quad = \quad \frac{c}{\bar{\pi}} + \left(\frac{1}{\bar{\pi}} - 1\right)(1-D)V^A, \tag{A.4}$$

$$\frac{\partial\bar{\pi}}{\partial b} = \frac{c+(1-D)V^A}{(b+(1-D)V^A)^2} \quad = \quad \frac{\bar{\pi}^2}{c+(1-D)V^A}. \tag{A.5}$$

Next, let's also rewrite the first order condition with respect to b (equation A.3) in terms of $\bar{\pi}$:

$$\frac{\bar{\pi}^2}{c+(1-D)V^A}\left\{\bar{\pi}\left[\Delta_{HL}-\frac{c}{\bar{\pi}}-\left(\frac{1}{\bar{\pi}}-1\right)(1-D)V^A\right]-\Delta_{ML}\right\}-\frac{1-\bar{\pi}^2}{2} = 0 \quad (A.6)$$

$$\Longleftrightarrow \bar{\pi}\Delta_{HL}-\Delta_{ML}-\left[\left(\frac{1}{\bar{\pi}^2}-1\right)\frac{c+(1-D)V^A}{2}+c+(1-\bar{\pi})(1-D)V^A\right] = 0 \quad (A.7)$$

where $\Delta_{HL}=s^H-s^L+(1-D)V^P$ and $\Delta_{ML}=s^M-s^L+(1-D)V^P$, both are positive. As equation (A.7) is increasing in $\bar{\pi}$, it has a unique solution $\bar{\pi}^*$. Furthermore, as equation (A.7) is decreasing in $(1-D)V^A$, this unique solution $\bar{\pi}^*$ is increasing in $(1-D)V^A$.

Comparative static of $\bar{\pi}^*$ with respect to D. As $(1-D)V^P$ and $(1-D)V^A$ are both decreasing in D, it follows that $\bar{\pi}^*$ is also decreasing in D (as $\bar{\pi}^*$ is increasing in $(1-D)V^P$ and $(1-D)V^A$). That is, for a given set of parameters (i.e., s^L, s^M, s^H, c, and corresponding V^P, V^A), $\bar{\pi}^*(1)<\bar{\pi}^*(0)$.

A.1.2 No credible commitment to tolerance of failure

This subsection considers the case when the manager cannot credibly commit to being tolerant of failure, the key decision that drives the model's results. In this setting, her decision whether to tolerate period 1's bad outcome is based on her updated belief at the end of period 1, with the aim to maximize her period-2 payoff. As bad outcome happens with probability $1-\theta^P$ due to a bad researcher and with probability $\theta^P\frac{(1-\bar{\pi})^2}{2}$ due to good researcher's bad luck ($\bar{\pi}$ is the good researcher's exploration threshold), the manager's updated belief after observing period 1's bad outcome is:

$$\theta^U=\frac{\theta^P(1-\bar{\pi})^2}{2(1-\theta^P)+\theta^P(1-\bar{\pi})^2}. \quad (A.8)$$

She then chooses to rehire the researcher if her period-2 expected payoff is larger than her outside option of zero. That is, when:

$$\frac{\theta^P(1-\bar{\pi})^2}{2(1-\theta^P)+\theta^P(1-\bar{\pi})^2}V_2^P+\frac{2(1-\theta^P)}{2(1-\theta^P)+\theta^P(1-\bar{\pi})^2}s^L>0 \quad (A.9)$$

$$\Longleftrightarrow \theta^P>\frac{-2s^L}{(1-\bar{\pi})^2V_2^P-2s^L}\equiv\bar{\theta}_{post}(\bar{\pi}). \quad (A.10)$$

In an equilibrium in which the manager chooses to tolerate failure (i.e., $D=1$), the good researcher's exploration threshold in period 1 is $\bar{\pi}=\bar{\pi}^*(1)$. This is then an equilibrium only when $\theta^P>\bar{\theta}_{post}(\bar{\pi}^*(1))$. Vice versa, in an equilibrium in which the manager chooses not to tolerate failure (i.e., $D=0$), the good researcher's exploration threshold is $\bar{\pi}=\bar{\pi}^*(0)$. This is then an equilibrium only when $\theta^P\leq\bar{\theta}_{post}(\bar{\pi}^*(1))$. In addition, as $\bar{\pi}^*(1)) < (\bar{\pi}^*(0)$ (Proposition 1) and $s^L < 0$, it follows that $\bar{\theta}_{post}(\bar{\pi}^*(1)) < \bar{\theta}_{post}(\bar{\pi}^*(0))$. The game's equilibrium can be summarized up as follows.

Proposition 4. *When the manager cannot credibly commit to being tolerant of failure, the game's equilibrium depends on her prior belief θ^P.*
(i) If $\theta^P > \bar{\theta}_{post}(\bar{\pi}^*(0))$, the manager credibly chooses to tolerate failure (i.e., $D = 1$) and the good researcher chooses exploration in period 1 when $\pi_1 > \bar{\pi}^*(1)$.
(ii) If $\theta^P \leq \bar{\theta}_{post}(\bar{\pi}^*(1))$, the manager chooses not to tolerate failure (i.e., $D = 0$) and the good researcher chooses exploration in period 1 when $\pi_1 > \bar{\pi}^*(0)$.
(iii) If $\bar{\theta}_{post}(\bar{\pi}^*(1)) < \theta^P \leq \bar{\theta}_{post}(\bar{\pi}^*(0))$, there are two equilibria: one in which the manager tolerates failure (i.e., $D = 1$) as in (i), and one in which she does not (i.e., $D = 0$) as in (ii).

Comparison with the baseline model. Recall from subsection 1.2.1 that in the baseline model when the manager can credibly commit to tolerance of failure, she chooses so (i.e., $D = 1$) when her prior belief θ^P is above threshold $\bar{\theta}$ (equation 1.2):

$$\theta^P > \frac{-2s^L}{2\left[V_1^P(1) - V_1^P(0)\right] + \left[1 - \bar{\pi}^*(0)\right]^2 V_2^P - 2s^L} \equiv \bar{\theta}.$$

As $V_1^P(1) > V_1^P(0)$ and $s^L < 0$, it follows that $\bar{\theta} < \bar{\theta}_{post}(\bar{\pi}^*(0))$ (see equation A.10). The intuition is that the manager's *ex ante* cutoff $\bar{\theta}$ takes into consideration $V_1^P(1) - V_1^P(0)$, the gain from optimal exploration in period 1 under $D = 1$, and therefore $\bar{\theta}$ is lower than her *ex post* cutoff $\bar{\theta}_{post}(\bar{\pi}^*(0))$, which does not internalize this gain.

As a result, for $\theta^P \in (\bar{\theta}, \bar{\theta}_{post}(\bar{\pi}^*(0)))$, without the capacity to commit, there always exists an equilibrium in which the manager does not tolerate period 1's bad outcome (i.e., $D = 0$), even though it is *ex ante* optimal for her to do so (i.e., $D = 1$) (Proposition 4).[1] Furthermore, if it is also the case that $\bar{\theta} < \bar{\theta}_{post}(\bar{\pi}^*(1))$,[2] then for $\theta^P \in (\bar{\theta}, \bar{\theta}_{post}(\bar{\pi}^*(1)))$, this non-tolerant equilibrium is the unique equilibrium, and the manager cannot at all implement the *ex ante* desirable policy of tolerance of failure. These problems are alleviated only if the manager can credibly commit to her *ex ante* decision, as in the baseline model, or if she is high trusting with $\theta^P > \bar{\theta}_{post}$. This result implies that trust acts as a substitute for commitment.

[1] In this case, there is another equilibrium in which she does tolerate period 1's failure.
[2] The relationship between $\bar{\theta}$ and $\bar{\theta}_{post}(\bar{\pi}^*(1))$ is ambiguous and depends on the parameter set.

A.2 Data construction

A.2.1 Firm sample construction

BoardEx to Computstat. I start with BoardEx dataset which contains detailed data on the background of CEOs and top officers for a large set of firms worldwide and select all firms that are both listed and headquartered in the US.[3] I then match the selected BoardEx firms to Compustat using ticker. To ensure that the matching is correct, I manually check all cases in which (i) the matching is not one to one,[4] or (ii) the company names in BoardEx and Compustat do not match. I then use CIK code to verify that the matching is indeed correct. Matched firms are larger than the remaining Compustat firms, with coverage of 55% in terms of firm counts and 85% in terms of total assets among Compustat firms with non-missing total assets between 2000 and 2011.

BoardEx-Compustat to Orbis. Next, I match BoardEx-Compustat firms to Orbis, a global company database provided by Bureau Van Dijk, to obtain the linkage between firms and patents.[5] This patent-to-firm linkage is based on a matching procedure implemented by the OECD and is available as part of Orbis.[6] The matching between BoardEx-Compustat and Orbis firms is done via ISIN/CUSIP. I also manually check all cases in which (i) the matching is not one to one, or (ii) the company names in BoardEx/Compustat and Orbis do not match. In addition, I use Orbis' manual search function to look for BoardEx-Compustat firms that cannot be identified in Orbis using ISIN/CUSIP. This results in a close to full match (above 99%) and allows me identify all patent applications owned by the matched firms. As Orbis also contains information on firm's ownership structure, I additionally identify patent applications by subsidiaries that are above 50% owned by one of these firms.

Sample restriction. Finally, I exclude all firms in finance, insurance, and real estate (SIC2 between 60 and 67), as these sectors make up a considerable share of the firm sample but traditionally do not patent their innovations. This resulting sample includes 4,345 firms during the study period between 2000 and 2011 (Table A.3), which yields in a final baseline sample of 3,598 firms after conditioning on firms having at least one CEO (i) whose ethnic origins could be inferred from her last name, and (ii) whose data on gender, age, and education are non-missing (see appendix A.2.3).

[3] BoardEx and Compustat data were retrieved through the Wharton Research Data Services (WRDS) in May 2017.

[4] This can happen when a firm undergoes a major merger and acquisition (M&A), in which case BoardEx considers it to be two different firms before and after the M&A while Compustat considers it to be the same firm if the ticker does not change. I follow BoardEx's approach to ensure that within-firm identification strategy is valid.

[5] I accessed Orbis platform through the LSE Library Services. The linkage between firms and patents provided by Orbis was retrieved in July 2017.

[6] The matching is done based on the names and addresses of patent applicants on patent records, which data are available from PATSTAT. In an UK setting, Dechezleprêtre et al. (2018) find that the matching quality is excellent with about 95% of UK and EPO patents being matched to their owning companies.

A.2.2 Patent and inventor data

My patent data are drawn from the 2016 Autumn Edition of the World Patent Statistical Database (PATSTAT) maintained by the European Patent Office (EPO). PATSTAT is the world's largest international patent database with nearly 70 million patent documents from over 60 patent offices, including all the major ones such as the United States Patent and Trademark office (USPTO), the European Patent Office (EPO), the Japan Patent Office (JPO), and the Chinese Patent and Trademark Office (SIPO). PATSTAT data cover close to the population of all worldwide patents between 1900 and 2015 and contain comprehensive information on patent application and publication dates, applicants and inventors, patent family, technology classification, and backward and forward citations.

Baseline patent counts. Each patent application to a patent office is uniquely identified in both PATSTAT and Orbis by its unique EPODOC application number. In addition to this identifier, PATSTAT reports a unique patent family indicator (DOCDB) which is the same for all patent applications (in different countries) related to the same invention. For the purpose of measuring innovation (i.e., to avoid double-counting inventions that are protected in several countries), I count all patent applications in the same family, irrespective of where they are filed, as one patent and assign this patent to its earliest application year. I only consider patents applications classified as "patent of invention" in PATSTAT, which is equivalent to USPTO's utility patents. The number of patent families filed by a firm (or a group of inventors in a firm) is my primary measure of innovation.[7] Over 2000-2012, 2,230 out of 3,598 baseline firms filed at least one patent, and together owned 1.8 million patent applications in 700,000 patent families over this period. In addition, I also construct alternative measures of innovation counting only patent families filed to or granted by the USPTO, which yield similar results (e.g., Table 1.5, Table A.14).

Patent quality measures. While not all innovations are patented and patenting norms vary across industries, it is reasonable to assume that within the same industry, the most valuable inventions are patented and therefore counting patents screens out the low-value ones.[8] In addition, I utilize various measures of patent quality to adjust for quality variation among patents and, more importantly, directly study this variation as an outcome of interest. The most well-known patent quality measure is forward citation counts (i.e., the number of future citations a given patent receives), which has been shown to be positively correlated with patent quality (Trajtenberg, 1990; Harhoff et al., 1999; Moser et al., 2015) and also firm's market value (Hall et al., 2005; Kogan et al.,

[7] I do not use fractional count to account for multiple applicants as this requires obtaining patent-firm linkage for the universe of firms. However, in practice, only a small share of patents are filed jointly by at least two firms (Dechezleprêtre et al., 2018).

[8] Note that all specifications in this paper include fixed effects below industry level, thereby sufficiently accounting for across-industry heterogeneity.

2017). In this paper, I use forward citation counts to compute: (i) a patent's quality decile relative to its technology field by application year cohort (e.g., Figure 1.7, Table A.13),[9] (ii) firm's quality-adjusted patent counts, and (iii) firm's average patent quality. Given that forward citations take time to accumulate, I restrict my sample to patents filed before 2012 to allow for observation windows of at least 5 years.

Besides forward citations, I also employ a range of alternative patent quality measures (Squicciarini et al., 2013), as listed below.

- First, the number of scientific papers a patent cite reflects how close the patent is to scientific knowledge and is an indicator of more complex and fundamental knowledge contained in the patent (Branstetter, 2005; Cassiman et al., 2008).
- Second, the scope of a patent, defined as the number of distinct technology classes (at IPC4 level) the patent is allocated to, has been shown to be associated with the patent's technological and economic value (Lerner, 1994).
- Third, the generality index, defined as one minus the technology-class Herfindahl–Hirschman Index (HHI) (at IPC4 level) of a patent's forward citations, measures the range of technology fields and industries influenced by the patent (Trajtenberg et al., 1997).
- Fourth, the originality index, defined as one minus the technology-class HHI (at IPC4 level) of a patent's backward citations, captures the breath of technology fields on which the patent relies, thereby reflecting its knowledge diversification (Trajtenberg et al., 1997).

Inventor data. PATSTAT also contains information on patent inventors' names and addresses as they appear on patent records. Based on data on inventors' countries extracted or inferred from their addresses (as provided by PATSTAT), 1,554 firms and 30% of patents in my sample have at least one non-US-based inventors. The share non-US-based inventors in each of these patents ranges from 0.5 to 1, with 1 being the median. That is, 60% of these patents are exclusively by non-US-inventors; furthermore, in almost all cases the inventors are based in the same country, consistent with the interpretation that the patents are by overseas R&D labs of multinational firms. The 10 most common locations of these labs, based on their patent contributions, are (in order) Germany, Great Britain, India, Canada, Japan, China, France, Israel, Switzerland, and Italy.[10] As it is possible for one patent to have inventors based in different countries, I use fractional count to calculate the number of patents at firm by inventor country level.

Data on patent inventors' names in PATSTAT come in much less standardized format as they are extracted from patent records from many different patent offices worldwide.

[9]Schmoch (2008) classifies patents into 35 technology fields of balanced size in 6 technology sectors of based on the International Patent Classification (IPC). This classification has subsequently been used in the innovation literature by Squicciarini et al. (2013) and Dechezleprêtre et al. (2018), among others.

[10]Eurobarometer bilateral trust measure is available for 7 out of these 10 countries.

To correctly separate out an inventor's last name (from first and middle names and even addresses), I supplement an algorithmic procedure with manual data cleaning. This allows me to identify 200,000 unique inventor last names from 1.8 million unique inventor name strings with reasonable confidence. Next, I match these last names to ethnic origins using the census-based mapping detailed in A.2.4, and further manually clean the remaining unmatched ones (as described in appendix A.2.3). By this process, I am able to identify the inventors' countries of origin for 90% of the patent sample based on either their non-US address or last names. I also use fractional count to calculate the number of patent at firm by inventor country level, as one patent could have multiple inventors and one inventor count be probabilistically mapped to multiple countries of origin.

A.2.3 CEO biographical data

I identify a firm's CEO in BoardEx from her position title, that is, if (i) it includes either one of the following phrases: *"CEO," "Chief Executive,"* or *"Principal Executive,"* and (ii) the phrase is not preceded by terms such as deputy, vice, division, group, regional, emeritus, etc. I verify if each firm has one CEO at a point in time, unless there are co/joint-CEOs, and manually check all exceptions. CEO transitions are inferred from the start and end dates of each CEO position.

CEOs' trust measures are computed from their ethnic origins as inferred from their last names (see subsection 1.3.2). I first map CEOs' last names to ethnic origins using the census-based last name-ethnic origin mapping detailed in appendix A.2.4. For the remaining CEOs whose last names do not appear with reasonable frequency in the censuses, I handpick out cases in which the last names distinctively belong to an ethnic group,[11] and manually search for the origins of unmatched CEO last names that appear with high frequency. This results in a final match rate of 83% at CEO level (77% from census-based mapping, 4% from manual mapping, and 2% from non-US citizenship). Panel A of A.3 shows that there are no significant differences between these name-matched 83% and the remaining non-matched 17% across all observable characteristics.

Besides using CEOs' names to infer their ethnic origins, I also employ data on their nationality, gender, age, education, and employment history. First, I exclude all CEOs who are explicitly not US citizens. They comprise only 4.8% of the 54% of CEOs for whom BoardEx contains nationality information. A quick check reveals that the other 46% represent cases in which the CEOs are obviously US citizens, so that the firm's website does not state their nationality. They are thus counted as US citizens. Second, I classify all degrees associated with each CEO into four different categories: below bachelor, bachelor, masters, and doctorate, and separately identify if a CEO has an MBA degree. I further supplement this classification with relevant information contained in CEOs' titles, such as "Doctor," "JD," or "MBA." The education variable is equal to the

[11] For example, compound last names including "Van" are likely Dutch, "Von" German, or "Les" French, etc. Further details are available upon request.

highest degree level a CEO has attained, and CEOs with no education information are dropped. Third, I use information on CEOs' employment history to impute their tenure in the respective firms, and to identify whether they have held an R&D related position prior to becoming the CEO. R&D-related positions are those whose title contains either one of the following words (or their derivations): "research," "innovation," "scientific," or "technology."

Table A.7 employ special subsamples of CEO retirement and death events. I define retirement as cases in which the CEOs (i) leave office at the age of or around 65, and (ii) do not have any executive positions afterward, as observed in BoardEx. 65 is the official Social Security retirement age and the traditional retirement age used in the related literature (e.g, Fee et al., 2013). In the data, I also observe a spike in CEO's leaving executive positions for good around 65. CEO deaths are identified from CEO's year of death as provided by BoardEx. They could be unexpected or the result of long-term health decline. Given that there are very few CEO deaths while still in position in my sample (only 34 cases), I do not further narrow them down to only sudden deaths as is the standard in the related literature (Nguyen and Nielsen, 2010; Bennedsen et al., 2010).

A.2.4 Mapping last names to ethnic origins

Sample of foreign-origin individuals. I start with de-anonymized full population samples of four US censuses between 1910 and 1940, which contain information on names and birthplaces of the US population. These restricted-access de-anonymized censuses are provided by the Minnesota Population Center through a formal application process. I keep all observations that meet the following criteria: (i) the individual is either male or never-married female; (ii) his last name is non missing; and (iii) either he or one of his parents was born outside of the US. This results in a sample of 79 million individuals with foreign (i.e., non-US) birthplace or ancestry across four censuses.

Each individual's origin is defined as: (i) his birthplace if it is outside of the US, (ii) his father's birthplace if his own birthplace is in the US or missing, or (iii) his mother's birthplace if both his own and his father's birthplaces are in the US or missing. I further refine this mapping by (i) dropping foreign-born individuals to both US-born parents, (ii) assigning individuals who were born outside of the US and Europe (e.g., Canada, Australia) to his parents' birthplaces if they were born in Europe, and (iii) refining coarse birthplace information (e.g., Central Europe) with additional information on mother tongue. However, these adjustments affect less than 1% of the sample. Among the 79 million foreign-origin individuals in the censuses, 87% are originally from Europe, 7% from Canada, 3% from Central America (mostly Mexico), 2% from Asia, and 1% from other parts of the world.

Last name-GSS ethnic origin mapping. Birthplace data in the census are coded mostly at country level, while ethnic-specific trust measure derived from the GSS is available for 36 most common ethnicities in the US (Table A.1). To address this, I construct a mapping between these two different classifications as follows.

- First, I map a country of origin in the census to an ethnic origin in the GSS if they represent the same country (e.g., Germany, Sweden, Italy) or region (e.g., England and Wales, Scotland).
- Second, I create new "aggregate" GSS ethnic groups (mostly for different regions within Europe) and map the remaining census countries of origin to their corresponding aggregate ethnic groups if possible. For example, Bulgarian, which is not an ethnicity included in the GSS, is mapped to a new ethnic group labeled as Eastern European, which is the aggregate of GSS ethnic groups Czechoslovakian, Hungarian, Polish, Romanian, and Russian.
- Third, I map the remaining countries in the census to existing coarse ethnic groups in the GSS such as African, Arabic, other Asian, other Spanish, or missing.

While this mapping may seem coarse, the fact that the GSS' ethnic classification is designed to cover the most common ethnicities in the US implies that a large share (at least 80%) of foreign-origin individuals in the censuses could be mapped to a GSS ethnic origin under the first step. On the other hand, the remaining ones still need to be accounted for systematically, as dropping them could introduce unwanted selection into the final last name-ethnic origin mapping. The exact correspondence between the census' country of origin and the GSS' ethnic origin classifications is available upon request.

79 million foreign-born individuals in the censuses share among them five million unique last names, the majority of which appear fewer than 10 times. To improve precision, I first filter out aberrant observations by dropping ethnic origins that occur less than 10% of the times for a given last name. I then consider only 75,000 last names that appear for at least 100 times in the remaining sample, which constitute 66% of this sample. The probabilistic mapping between last names and GSS ethnic origins is constructed from the resulting sample. Specifically, I compute w_{se}, the probability that a person with last name s is of ethnicity e as $w_{se} = \frac{n_{se}}{N_s}$, in which n_{se} is the number of individuals with last name s from ethnic origin e, and N_s is the total number of individuals with last name s in the sample. For example, based on this mapping, the last name Johnson is 78% Swedish and 22% Norwegian; the last name Smith is 32% English, 26% German, 24% Irish, and 18% Canadian.

This last name-GSS ethnic origin mapping is used in computing CEO's inherited generalized trust measure,[12] and other measures of inherited cultural traits. In sensitivity tests, I find that lowering the aforementioned 100 observation and 10% share cutoffs to retain more observations does not significantly improve the match rate of CEO last names and slightly reduces the precision of the key estimates.

[12] $w_{de} = w_{se}$ if CEO d's last name is s.

Most common last name supplements. One concern is that data from historical censuses do not capture more recent waves of migration to the US. However, it has been shown that CEOs in the US are predominantly "WASP" (White Anglo-Saxon Protestant), which groups arrived in the US well before the 1940s. Furthermore, to address the concern, I supplement the census-based mapping with lists of most common last names in 50 different countries collected from online sources such as `forebear.com` or `wikipedia.com`. These lists also provide me a way to cross-check the quality of the census-based mapping. First, I develop a list of all last names that could account for at least 0.01% of immigrants in the US. Each last name's predicted share is computed as the share of the last names in its respective country times the share of immigrants from that country in the US foreign-origin population (based on census data between 1960-2015). I then convert countries to GSS ethnic origins and add last names from the list to the census-based mapping. For last names that are already included in the mapping, I find that the census-based mapping is generally consistent with the information from the list.

Last name-Eurobarometer origin mapping. To compute CEOs' bilateral trust measure, I construct the mapping between last names and countries of origin covered in the Eurobarometer in similar steps to above. As the Eurobarometer provides bilateral trust measure for only 16 trust-originating and 28 trust-receiving countries, foreign-origin individuals in the census sample who are not from one of these countries are assigned to the "not-covered" category. Furthermore, I construct two separate mappings. The first one considers the 16 trust-originating countries and is used for mapping CEOs' last names. The second one considers the 28 trust-receiving countries and is used for mapping inventors' last names.

After constructing the first mapping that is used for CEOs, I drop all last names with above 20% probability of being from a "not-covered" country of origin, then drop this "not-covered" category and rescale the remaining w_{se}'s so that $\sum_{e \in \mathbf{E}} w_{se} = 1$ where $\mathbf{E}$ is the set of 16 trust-originating countries. That is, the final mapping contains only last names with at least 80% probability of being from countries for which bilateral trust measure is available. The rescaling is required as the CEO's bilateral trust measure is the weighted average of country pairwise bilateral trust.

A.3 Trust measurement error and bias

A.3.1 Relative magnitude of trust measurement error

Because of the indirect nature of my measure of inherited trust, it is important to gauge the relative magnitude of measurement error, and its impact on the estimate. In what follows, I propose a method to evaluate the extent of measurement error of inherited trust, using existing results from trust game experiments such as Glaeser et al.'s (2000). Denote a person i's trust as T_i, the major ingredient in my theory. As remarked in the literature, the GSS's trust survey question produces a measurement error ϵ_i, so that we only observe surveyed trust as $TS_i = T_i + \epsilon_i$.The empirical ethnic component of trust, as calculated from trust survey, is $TEth_c = \mathbb{E}(TS_i|c) = \mathbb{E}(T_i|c) + \mathbb{E}(\epsilon_i|c)$. In case of an independent error ϵ_i, $TEth_c = \mathbb{E}(T_i|c)$.

My first question is on the relative magnitude of the discrepancy between $TEth_c$ and T_i, namely $R_{TEth} = \frac{\mathrm{Var}(TEth_i)}{\mathrm{Var}(T_i)}$. As $\frac{\mathrm{Var}(TEth_c)}{\mathrm{Var}(TS_i)} = 0.06$ comes straight from the GSS sample, it remains to find $R_T = \frac{\mathrm{Var}(T_i)}{\mathrm{Var}(TS_i)}$.

Consider the experimental setting in Glaeser et al. (2000) in which subjects play a trust game, and their decisions are then linked to their answers to a GSS trust question. Based on the literature on the stability of trust experiments, I suppose that the trust game decision TG_i (a number between 0 and 15 in that context) contains an idiosyncratic error η_i: $TG_i = \gamma T_i + \eta_i$, with a ratio of signal to total variation $R_{TG} = \frac{\mathrm{Var}(\gamma T_i)}{\mathrm{Var}(TG_i)}$. According Falk et al. (2016), this ratio is around 60%. We learn from Glaeser et al. (2000) that the regression of TG_i on TS_i yields a coefficient of $\hat{b}_G$ with a standard error of $\hat{\sigma}_G$. I will make use of those two numbers and R_{TG} to compute R_T.[13]

Using formulae of regressions with measurement errors, I can write $\hat{b}_G = \gamma \frac{\mathrm{Var}(T_i)}{\mathrm{Var}(TS_i)} = \gamma R_T$. Its standard error can also be written as:

$$\begin{aligned}
\hat{\sigma}_G^2 &= \frac{\mathrm{Var}(TG_i - \hat{b}_G TS_i)}{\mathrm{Var}(TS_i)} = \frac{\mathrm{Var}[(\gamma - \gamma R_T)T_i + \eta_i - \gamma R_T \epsilon_i]}{\mathrm{Var}(TS_i)} \\
&= \gamma^2(1-R_T)^2 R_T + \gamma^2 \frac{1-R_{TG}}{R_{TG}} R_T + \gamma^2 R_T^2 (1-R_T) \\
&= \gamma^2 R_T(1-R_T) + \gamma^2 \frac{1-R_{TG}}{R_{TG}} R_T.
\end{aligned}$$

Replacing $\gamma = \hat{b}_G / R_T$, we obtain:

$$\hat{\sigma}_G^2 R_T = \hat{b}_G^2 \left(1 - R_T + \frac{1-R_{TG}}{R_{TG}}\right) \Rightarrow R_T = \frac{\hat{b}_G^2}{(\hat{b}_G^2 + \hat{\sigma}_G^2) R_{TG}} = \frac{t^2}{t^2+1} \frac{1}{R_{TG}},$$

[13] There is a debate following Glaeser et al. (2000) on the validity of different trust measures. In defense of trust surveys, Sapienza et al. (2013) argue that the sender's behavior in the trust game, Glaeser et al.'s (2000) preferred measure of trust, is not necessarily a good measure of trust, because it is confounded by other-regarding preferences. In contrast, WVS/GSS trust questions better capture the belief-based component of the trust game, which corresponds better to the concept of trust as defined in Gambetta (1988) and in my model.

with $t = \frac{\hat{b}_G}{\hat{\sigma}_G}$ the t-statistic of the test $b_G = 0$. As there are two potential outcomes from trust games in Glaeser et al. (2000), I compute the average of t over the two potential outcomes from trust games in Glaeser et al. (2000) at around 0.50,[14] mapping into $R_T = 0.33$. I thus deduce $R_{TEth} = 0.18$. That is, the ethnic component of trust measures about 18% of the variation in individual trust. Finally, when I use a LASSO model to predict trust using all observables and their interactions, the ratio of predicted variation $\frac{\text{Var}(TEth_c)}{\text{Var}(TS_i)}$ rises to about 0.11, corresponding to $R_{TEth} = 0.33$.

Discussion. A few remarks can be drawn from those exercises. First, one can argue that $TEth_c$ is a much better measure of trust than a simple survey answer TS_i, as the variance of the survey noise ϵ_i far outweighs the variance of individual components $T_i - TEth_c$ (the ratio of variance is $\frac{0.67}{0.27}$, or about 2.5 times). Therefore, it would not have added value even if we could administer a trust survey among CEOs.[15] Second, even if we could run a trust game among CEOs, the ratio of the variance of the experimental noise η_i to the variance of ν_i is about $\frac{0.33 \times \frac{100\% - 60\%}{60\%}}{0.27} \sim 0.81$. That is, using my inherited trust measure is 81% as precise as using trust game results from CEOs. Third, as shown in appendix A.3.1, while ethnic specific inherited trust likely represents only 18% of inherent individual trust, the benchmark regression likely produces an unbiased of the true effect (there is no attenuation bias as in the case of classical measurement errors). The main intuition from this exercise is that both methods of elicitation of individual trust, either via surveys or via trust games, produce a considerable amount of measurement error, as has been shown throughout the literature. While my method of averaging trust survey answers by ethnic origin misses the individual-specific component of trust, it also helps in smoothing out those measurement errors. Quantitatively, the latter effect can more than compensate the former.

A.3.2 Bias due to trust measurement error

The second question regarding measurement error is how much does the discrepancy between $TEth_c$ and T_i affect the estimate of the effect of trust. Let us assume the true relationship between innovation outcome Y_{ft} of firm f in year t and its current CEO d's individual trust T_{dt} as $Y_{ft} = \beta T_{dt} + \theta_f + u_{fdt}$, with a firm fixed effect θ_f, and an independent error term u_{fdt}. When current CEO's ethnic-specific inherited trust $TEth_{ct}$ is used in place of individual trust T_{dt}, the fixed effect estimator is $\hat{\beta}_{TE} = \frac{\text{Cov}(M.Y_{ft}, M.TEth_{ct})}{\text{Var}(M.TEth_{ct})}$, given the linear de-mean operator $M.X_{it} = X_{it} - \mathbb{E}_t(X_{it}|i)$.

[14] The outcomes are the amount sent by the first player, and the reservation price that the first player considers equivalent to the value of the game. As discussed in Sapienza et al. (2013), those measures should be considered with caution, as they may also include effects due to social preferences, not just beliefs

[15] Of course, if we can administer *many* trust surveys on the same individual, we can average out much more precisely individual trust. I consider this possibility highly infeasible though.

Also observe that:

$$\begin{aligned}
\text{Cov}(M.Y_{ft}, M.TEth_{ct}) &= \text{Cov}(\beta M.T_{dt} + M.u_{fdt}, M.TEth_{ct}) \\
&= \beta\text{Cov}(M.\mathbb{E}(T_{dt}|c) + M.(T_{dt} - \mathbb{E}(T_{dt}|c)), M.TEth_{ct}) \\
&= \beta\text{Cov}(M.TEth_{ct} - M.\mathbb{E}(\epsilon_{dt}|c), M.TEth_{ct}).
\end{aligned}$$

In case of independent survey measurement error ϵ_{dt}, the expression above is reduced to $\beta\text{Var}(M.TEth_{ct})$. Therefore, using the ethnic component of trust $TEth_{ct}$ in place of individual trust T_{it} does not create any bias in the firm fixed effect specification. In essence, this exercise is similar to taking a cell-average of the right hand side variable, and then use it as a new regressor, a procedure that is very useful especially when one can only observe cell averages (see also Angrist and Pischke, 2009, c. 2.).

If the survey measurement error ϵ_{dt} is not mean-independent of the respondent's country, $\hat{\beta}_{TE}$ with be biased from β by $-\frac{\text{Cov}(M.\mathbb{E}(\epsilon_{dt}|c), M.TEth_{ct})}{\text{Var}(M.TEth_{ct})}$. Based on the empirical results, we can assume that there is little autocorrelation over time between different CEOs at the same firm, in which case we can get rid of the operator M to rewrite the bias as $-\frac{\text{Cov}(\mathbb{E}(\epsilon_{dt}|c), TEth_{ct})}{\text{Var}(M.TEth_{ct})}$.

The bias' sign is that of $-\text{Cov}(\mathbb{E}(\epsilon_{dt}|c), TEth_{ct})$, or the opposite of the covariance across countries between ethnic-based inherited trust, and individual survey measurement errors. It is likely negative if, for example, high-trust countries' respondents tend to push their answers higher, and low-trust countries' respondents tend to lower theirs. There is a technical reason to expect this pattern: Surveyed trust TS_i is a yes-no answer, which naturally exaggerates the variation in the individual trust component T_i. For example, two individuals' beliefs at 60% and 40% will map into two opposite answers of value 1 and 0, respectively.[16] Consequently, the estimator $\hat{\beta}_{TE}$ likely underestimates the true effect of individual CEO's trust on innovation.

[16] There is, however, another reason to expect the covariance to be negative and the bias positive: if the individual belief ranges mostly on one side of 50%, say, from 50% to 100%, then they all correspond to a survey answer of 1. As a stronger belief entails a smaller error term, the covariance is negative. When both effects are taken into account, based on the empirical distribution of survey answers, I can show that under mild conditions, and in most simple simulations, the positive-covariance effect largely dominates, therefore the bias is very probably negative.

A.4 Framework for separating mechanisms

A.4.1 Proof of Proposition 3

As the count of patents within quality range $[c_1, c_2]$ is $N\left[F(\frac{c_2-\tilde{b}(T)}{a(T)}) - F(\frac{c_1-\tilde{b}(T)}{a(T)})\right] \stackrel{def}{\equiv} Y(a(T), b(T))$, the effects of changes in a and b on patent counts within quality range $[c_1, c_2]$ are:

$$\frac{\partial Y}{\partial a} = N\left[f\left(\frac{c_2-\tilde{b}}{a}\right)\left(-\frac{c_2-\tilde{b}}{a^2}\right) - f\left(\frac{c_1-\tilde{b}}{a}\right)\left(-\frac{c_1-\tilde{b}}{a^2}\right)\right], \tag{A.11}$$

$$\frac{\partial Y}{\partial b} = N\left[f\left(\frac{c_2-\tilde{b}}{a}\right)\left(\frac{-1}{a}\right) - f\left(\frac{c_1-\tilde{b}}{a}\right)\left(\frac{-1}{a}\right)\right]. \tag{A.12}$$

Recall the assumptions from subsection 1.6.1 that better quality patents are always rarer (i.e., $F_T'(x)$ is decreasing on $[0, \infty)$ $\forall T$). This assumption implies that $f\left(\frac{c_2-\tilde{b}}{a}\right) < f\left(\frac{c_1-\tilde{b}}{a}\right)$ as $c_1 < c_2$. As a result, $\frac{\partial Y}{\partial b}$ in expression (A.12) is always positive, indicating that higher $b(T)$ increases the count of patents within the quality range $[c_1, c_2] \subset [0, \infty)$.

A.4.2 Patent quality under mean-preserving spread

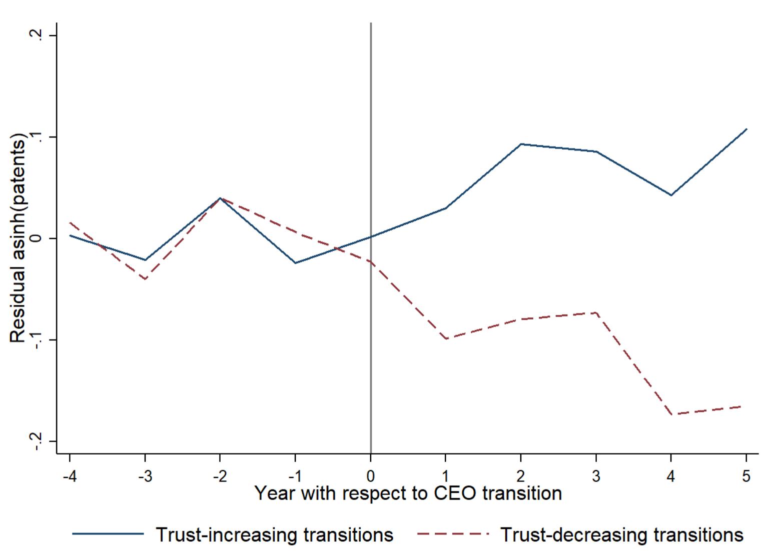
Unlike expression (A.12) which is always positive, expression (A.11) does not have an unambiguous sign: while $f\left(\frac{c-\tilde{b}}{a}\right)$ is decreasing in c over the $[c_1, c_2]$ interval as shown in appendix A.4.1, $c - \tilde{b}$ is increasing. It is thus possible to identify the mechanism at work (i.e., through $a(T)$ or $b(T)$) under conditions that warranty $\frac{\partial Y}{\partial a} < 0$, namely, for ranges of c where $f\left(\frac{c-\tilde{b}}{a}\right)\left(\frac{c-\tilde{b}}{a^2}\right)$ is increasing in c. This condition is quite easy to satisfy for small c, at least among distributions in the exponential family, such that when c is small, $f\left(\frac{c-\tilde{b}}{a}\right)$ decreases less fast than $c - \tilde{b}$ increases. The following proposition illustrates a special case:

Proposition 5. *Consider a normal distribution $\mathcal{N}(\bar{x}, \sigma)$ with density f, and $b = 0$ (i.e., no mean-shifting mechanism at work). Higher $a(T)$ decreases the count of patents of quality within any range $[c_1, c_2] \subset [0, a\sigma + \bar{x}]$.*

Proof. The proposition's statement is equivalent to $\frac{\partial Y}{\partial a}$ in expression (A.12) being negative, itself equivalent to $f\left(\frac{c-\tilde{b}}{a}\right)\left(\frac{c-\tilde{b}}{a^2}\right)$ being increasing in c over the $[c_1, c_2]$ interval. This happens when its derivative with respect to c: $\frac{1}{a^2}\left[\frac{1}{a}f'\left(\frac{c-\tilde{b}}{a}\right)(c-\tilde{b}) + f\left(\frac{c-\tilde{b}}{a}\right)\right]$, is nonnegative.

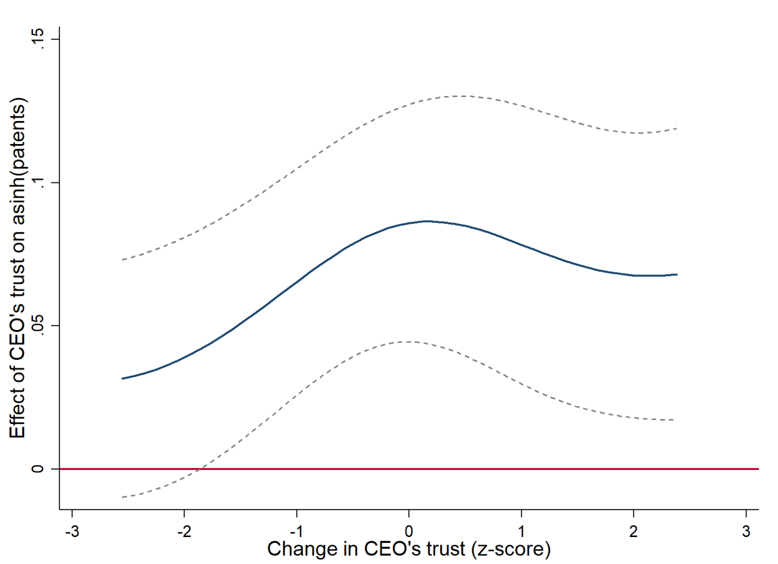
In case f is a normal distribution density, this derivative being nonnegative is equivalent to $-\frac{1}{a^2\sigma^2}(c-\tilde{b})^2 + 1 \geq 0 \Leftrightarrow |c - \tilde{b}| = |c - \bar{x} - b| \leq a\sigma$. Let $b = 0$, this condition is equivalent to $c \in [0, a\sigma + \bar{x}]$. □

Figure A.1: Patents by change in CEO's trust (non-matched sample)



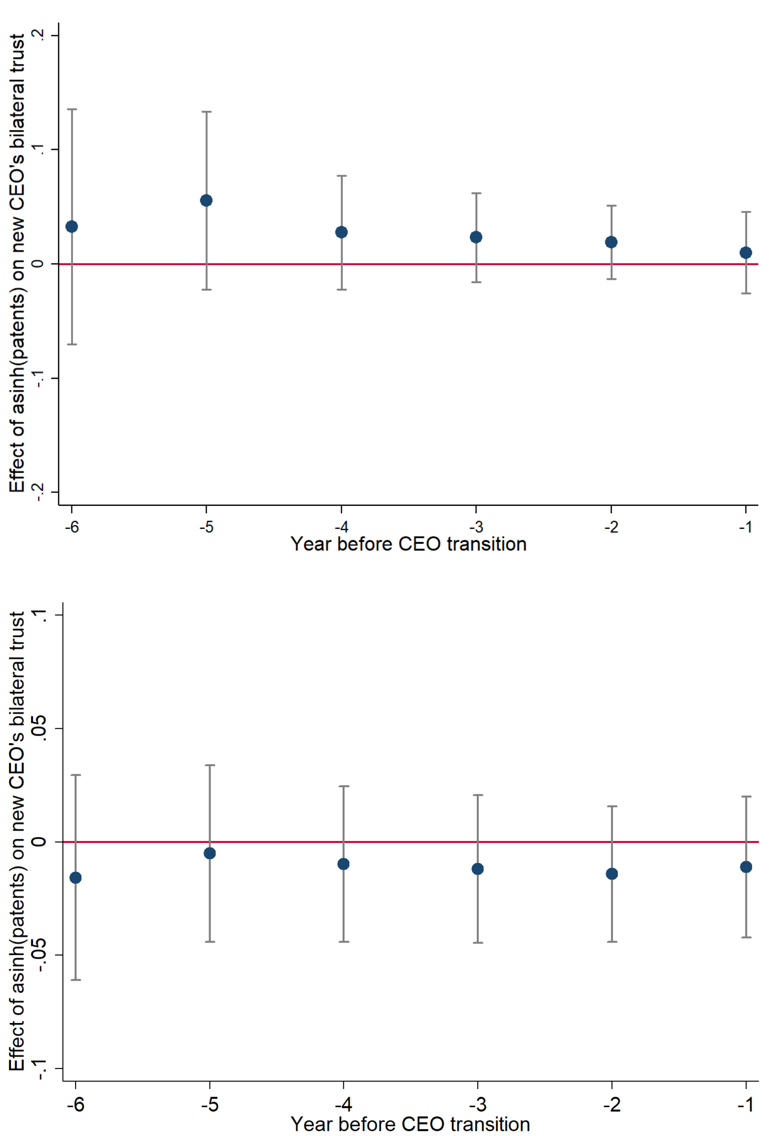
Notes: This figure plots firms' average residual patent application counts (after partialling out the covariates) by year with respect to CEO transition year (i.e., year 0). The solid blue line groups together all CEO transitions in which the new CEOs are *more* trusting than their predecessors (i.e., trust-increasing transitions), and the dotted red line corresponds to those in which the new CEOs are *less* trusting (i.e.,trust-decreasing transitions). Each group's annual average residual patent counts are plotted relative to the group's pre-transition mean, which is normalized to 0. The sample includes CEO transitions in which both predecessor's and successor's tenures are at least 5 years.

Figure A.2: CEO's trust effect by change in CEO's trust



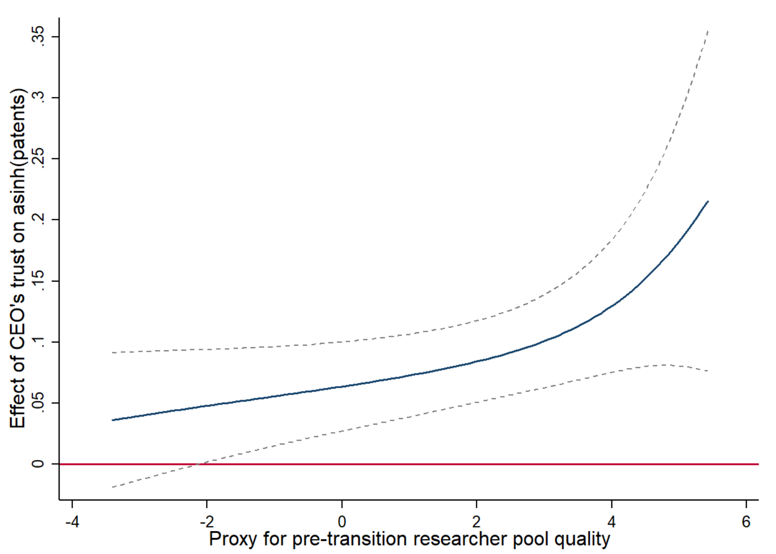
Notes: This figure plots semi-parametric estimates of the CEO's trust coefficient on firm's patents as a function of the change in CEO's trust after the corresponding transition (the X-axis variable). The semiparametric estimation is based on equation (1.5), using a Gaussian kernel function of the X-axis variable and a bandwidth of 20% of the range. The dashed lines indicate the 95% confidence intervals for the CEO's trust coefficients. Standard errors are clustered by CEO's main ethnicity.

Figure A.3: Pre-change patents and new CEO's bilateral trust



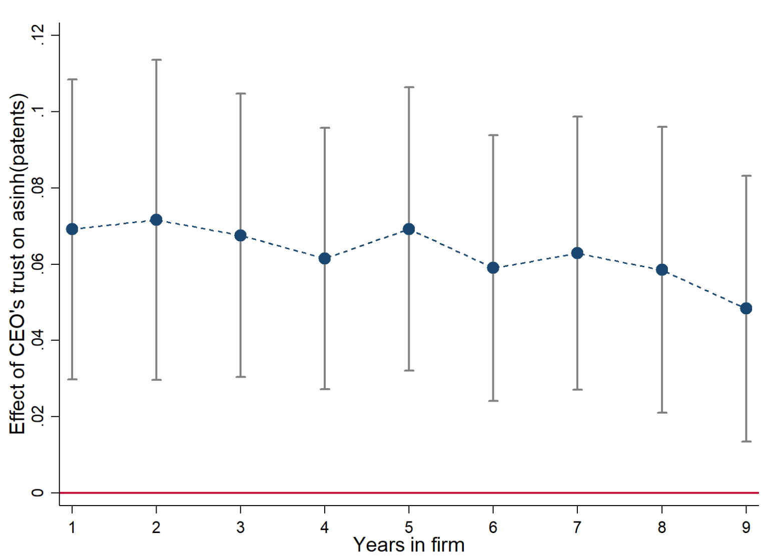
Notes: This figure plots the coefficients $\hat{\gamma}_k$ for $k \in [-6, -1]$ from estimating: $\Delta bitrust_{fdct} = \sum_{k=-7}^{-1} \gamma_k (\text{asinh}(pat_{fdct}) \times event_{t-k}) + \beta bitrust_{fdct} + \mathbf{X}_{ft} + \mathbf{Z}_{dt} + \theta_f + \kappa_c + \omega_t + \varepsilon_{fdct}$, in which (i) $\Delta bitrust_{fdct}$ is the difference between CEO d's and her successor's bilateral trust measures towards individuals from country c, and (ii) $event_{t-k}$ is an indicator equal to 1 if the transition happens in year $t-k$. Estimates are shown with their 95% confidence intervals. Standard errors are clustered by firm. The upper plot corresponds to the bilateral trust sample in which an inventor's country is inferred from his patent-listed address for non-US-based inventors. The lower plot corresponds to the bilateral trust sample in which an inventor's country is additionally inferred from his last name for US-based inventors.

Figure A.4: CEO's trust effect by pre-transition researcher pool quality



Notes: This figure plots semi-parametric estimates of the CEO's trust coefficient on firm's patents as a function of pre-transition researcher pool quality (the X-axis variable). Firm-level proxy for researcher pool quality is computed from the residuals from regressing patents on observable firm and CEO characteristics, controlling for SIC2 industry and year fixed effects (subsection 1.6.3) over a 2-year pre-transition window. The semiparametric estimation is based on equation (1.5), using a Gaussian kernel function of the X-axis variable and a bandwidth of 20% of the range. The dashed lines indicate the 95% confidence intervals for the CEO's trust coefficients. Standard errors are clustered by CEO's main ethnicity.

Figure A.5: CEO's trust effect by tenure in firm



Notes: This figure plots the coefficients $\hat{\beta_k}$ for $k \in [1,9]$ from estimating: $\text{asinh}(pat_{fd,t+1}) = \sum_{k=1}^{9}(trust_{fdt} \times tenure_{dk} \times successor_d) + \mathbf{X}_{ft} + \mathbf{Z}_{dt} + \theta_E + \omega_t + \varepsilon_{fdt}$ using the transition-event sample, in which (i) $tenure_{dk}$ is an indicator equal to 1 if the CEO d starts working in firm f in year $t-k+1$, and (ii) $sucessor_d$ is an indicator equal to 1 if CEO d is the successor in transition E.

Table A.1: GSS inherited trust measure by ethnic origin

Rank	Ethnic origin	Trust measure	Rank	Ethnic origin	Trust measure
1	Belgium	0.727	19	Japan	0.500
2	Sweden	0.629	20	Romania	0.500
3	Switzerland	0.622	21	India	0.494
4	Norway	0.619	22	Arabic	0.478
5	Denmark	0.603	23	Other Asian	0.478
6	Canada	0.600	24	Italy	0.470
7	England and Wales	0.593	25	China	0.468
8	Hungary	0.587	26	Greece	0.467
9	Lithuania	0.577	27	Austria	0.465
10	Ireland	0.565	28	Spain	0.423
11	Russia and former USSR	0.565	29	Finland	0.419
12	Scotland	0.553	30	Portugal	0.368
13	Germany	0.553	31	Mexico	0.368
14	Netherlands	0.551	32	Philippines	0.356
15	Czechslovakia	0.551	33	West Indies (Hispanic)	0.353
16	Yugoslavia	0.533	34	Africa	0.265
17	France	0.529	35	Other Spanish	0.246
18	Poland	0.523	36	West Indies (non-Hispanic)	0.200

Notes: This table reports inherited trust measure by ethnic origin, $ethtrust_e$, computed as the average trust attitude (0 – low trusting, 1 – high trusting) of GSS respondents whose (i) self-reported ethnic origin is e and (ii) GSS occupation prestige score is at least 50 (subsection 1.3.2). The standard deviation of this inherited trust measure at ethnicity level is 0.115.

Table A.2: GSS ethnic origins of CEOs

Baseline sample			Name-matched sample		
Rank	**Ethnic origin**	**Share of CEOs**	**Rank**	**Ethnic origin**	**Share of CEOs**
1	Ireland	19.5%	1	Ireland	18.8%
2	Germany	18.7%	2	Germany	18.0%
3	England and Wales	16.6%	3	England and Wales	17.2%
4	Canada	10.0%	4	Canada	10.1%
5	Russia and former USSR	8.1%	5	Russia and former USSR	8.3%
6	Italy	6.7%	6	Italy	6.6%
7	Scotland	3.3%	7	Scotland	3.2%
8	Sweden	2.7%	8	Sweden	2.6%
9	Poland	2.2%	9	Poland	2.2%
10	Austria	1.6%	10	Australia	1.6%
11	Norway	1.6%	11	Norway	1.5%
12	China	1.2%	12	China	1.1%
13	Mexico	0.9%	13	Mexico	0.9%
14	India	0.7%	14	India	0.9%
15	Netherlands	0.7%	15	Netherlands	0.8%
16	Denmark	0.7%	16	Denmark	0.7%
17	Czechslovakia	0.6%	17	Czechslovakia	0.6%
18	Hungary	0.5%	18	Hungary	0.5%
N = 5,753			N = 7,027		

Notes: This table reports the distribution of CEOs' ethnic origins as inferred from their last names (subsection 1.3.2) for (i) 5,753 CEOs in the baseline sample, and (ii) 7,027 name-matched CEOs.

Table A.3: Baseline sample's descriptive statistics

Panel A. CEO's characteristics

Sample:	**Baseline**		**Name-matched**		**Unmatched**	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Inherited generalized trust (baseline)	0.56	(0.04)	0.56	(0.04)		
Inherited generalized trust (LASSO)	0.55	(0.09)	0.55	(0.09)		
Inherited generalized trust (full GSS)	0.45	(0.04)	0.45	(0.04)		
Inherited generalized trust (full WVS)	0.37	(0.08)	0.36	(0.08)		
Gender (1 – male, 0 – female)	0.97	(0.18)	0.97	(0.17)	0.98	(0.15)
CEO's age in 2000	48.2	(9.1)	48.6	(9.3)	48.1	(9.12)
Highest degree: Bachelor	0.37	(0.48)	0.36	(0.48)	0.34	(0.47)
Highest degree: Masters	0.43	(0.49)	0.43	(0.49)	0.42	(0.49)
Highest degree: Doctor	0.18	(0.38)	0.18	(0.38)	0.21	(0.41)
Has MBA degree	0.34	(0.48)	0.34	(0.47)	0.35	(0.48)
Has non-MBA postgrad degree	0.26	(0.44)	0.26	(0.44)	0.28	(0.45)
Has prior R&D experience	0.02	(0.14)	0.02	(0.13)	0.02	(0.14)
Age when becoming CEO	50.1	(8.5)	50.3	(8.7)	49.9	(8.9)
Prior tenure in firm (yrs)	6.44	(8.18)	6.59	(8.36)	6.70	(8.57)
Tenure as CEO (yrs)	7.23	(6.14)	7.23	(6.32)	7.22	(6.11)
# CEOs	5,753		7,027		1,466	

Panel B. Firm's characteristics

Sample:	**Baseline**		**Name-matched**		**Unmatched**	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Patent applications p.a.	18.0	(149.7)	15.9	(138.8)	5.2	(25.2)
asinh(patent applications)	1.00	(1.62)	0.93	(1.56)	0.86	(1.36)
Citation-weighted patents p.a.	202.9	(1,953)	176.1	(1,764)	59.7	(230.5)
asinh(citation-weighted patents)	1.72	(2.56)	1.61	(2.49)	1.55	(2.34)
Firm's age in 2000	11.3	(15.4)	12.1	(15.2)	8.7	(13.7)
# years in sample	7.5	(3.6)	9.5	(3.1)	8.4	(3.8)
# CEOs in sample	1.68	(0.88)	2.13	(1.13)	1.17	(0.41)
# matched CEOs in sample	1.68	(0.88)	1.85	(1.00)	0.00	(0.00)
Is R&D performing firm	0.60	(0.49)	0.59	(0.49)	0.63	(0.48)
Is patenting firm	0.55	(0.50)	0.57	(0.49)	0.55	(0.50)
Total assets p.a. ($mil)	3,472	(14,957)	3,460	(19,541)	1,676	(9,198)
Total sales p.a. ($mil)	2,924	(12,765)	2,824	(13,163)	2,179	(12,373)
Employment p.a. ('000)	10.76	(51.72)	9.94	(46.68)	5.20	(15.10)
R&D stock p.a. ($mil)	297.1	(2,003)	272.5	(1,895)	82.0	(271.4)
R&D expenditure p.a. ($mil)	71.9	(466.4)	66.1	(442.8)	23.7	(90.6)
# Firms	3,598		4,000		345	

Notes: **Panel A** reports the descriptive statistics of CEO's characteristics for (i) 5,753 CEOs in the baseline sample, (ii) 7,027 name-matched CEOs, and (iii) 1,466 unmatched CEOs. **Panel B** reports the descriptive statistics of firm's characteristics for (i) 3,598 firms in the baseline sample (considering only firm × year observations that correspond to name-matched CEOs), (ii) 4,000 firms having at least one name-matched CEOs (considering all firm × year observations in the study period), and (iii) 345 firms having no name-matched CEO. Inherited generalized trust measure ranges from 0 – low trusting to 1 – high trusting.
p.a. stands for per annum.

Table A.4: Bilateral trust samples' descriptive statistics

Panel A. CEO's characteristics

Sample: Based on inventors'	**Non-US addresses**		**Addresses/last names**	
# associated inventor countries	4.8	(5.2)	6.8	(6.1)
Bilateral trust (towards inventor country)	2.69	(0.31)	2.67	(0.32)
Inherited generalized trust (baseline)	0.55	(0.04)	0.55	(0.04)
Gender (1 – male, 0 – female)	0.97	(0.16)	0.97	(0.16)
CEO's age in 2000	48.4	(8.8)	48.3	(8.8)
Highest degree: Bachelor	0.35	(0.48)	0.36	(0.48)
Highest degree: Masters	0.45	(0.50)	0.44	(0.50)
Highest degree: Doctor	0.19	(0.39)	0.18	(0.38)
Has MBA degree	0.37	(0.48)	0.37	(0.48)
Has non-MBA postgrad degree	0.26	(0.44)	0.25	(0.44)
Has prior R&D experience	0.03	(0.16)	0.02	(0.15)
# CEOs	960		1,654	

Panel B. Firm's characteristics

Sample: Based on inventors'	**Non-US addresses**		**Addresses/last names**	
# associated inventor countries	4.8	(5.1)	6.8	(6.1)
Patent applications p.c. p.a.	1.5	(10.7)	3.0	(21.3)
asinh(patent applications)	0.39	(0.85)	0.68	(1.03)
Citation-weighted patents p.c. p.a.	11.9	(69.9)	30.5	(222.5)
asinh(citation-weighted patents)	0.86	(1.59)	1.59	(1.90)
Firm's age in 2000	14.0	(16.6)	12.6	(15.7)
# years in sample	6.4	(3.7)	6.4	(3.7)
# CEOs in sample	1.4	(0.6)	1.4	(0.6)
Total assets p.a. ($mil)	4,840	(20,729)	3,983	(17,037)
Total sales p.a. ($mil)	3,960	(12,921)	3,294	(11,490)
Employment p.a. ('000)	12.4	(36.8)	10.7	(35.0)
R&D stock p.a. ($mil)	803.3	(3,234)	479.3	(2,491)
R&D expenditure p.a. ($mil)	189.1	(735.4)	113.4	(567.0)
# Firms	730		1,263	

Notes: This table reports the descriptive statistics of CEO's and firms' characteristics for (i) CEOs and firms in the bilateral trust sample based on inventors' patent-listed non-US addresses, and (ii) CEOs and firms in the bilateral trust sample based on inventors' addresses (for non-US based inventors) or last names (for US-based inventors). Bilateral trust measure ranges from 1 – least trusting to 4 – most trusting. Inherited generalized trust measure ranges from 0 – low trusting to 1 – high trusting.
p.c. stands for per country; p.a. stands for per annum.

Table A.5: Average patents before and after CEO transitions

Variable:	**Average residual asinh(patents)**		
Sample:	Before transition	After transition	Difference
Trust-increasing CEO transitions	-0.146	-0.012	0.135*
	(0.056)	(0.052)	(0.076)
Trust-decreasing CEO transitions	-0.088	-0.216	-0.128*
	(0.053)	(0.051)	(0.073)
Difference	-0.058	0.204***	0.262**
	(0.077)	(0.073)	(0.106)

Notes: This table reports the average residual patent application counts (after partialling out the covariates) in the 5 years before and after CEO transitions, separately for trust-increasing and trust-decreasing transitions as described in the notes to Figure 1.4. There are 61 trust-increasing CEO transitions, each of which is matched to a trust-decreasing CEO transition based on their average pre-transition residual patent counts (resulting in a total of 44 unique matched trust-decreasing CEO transitions). Pre-transition period covers years -5 to 0; post-transition period covers years 1 to 5.

Table A.6: Robustness checks for CEO's trust effect on firm's patents

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	**asinh(Future patent applications)**							
Specification:	Baseline	Alt. clusterings		Alt. samples				Poisson
CEO's trust	0.063***	0.063***	0.063***	0.063***	0.061***	0.075***	0.087***	0.168**
	(0.019)	(0.022)	(0.022)	(0.019)	(0.020)	(0.020)	(0.026)	(0.069)
Sample excluding				Single-tons	Female CEOs	Interim CEOs	Transition years	
Clustering scheme		Firm	Two-way					Robust
Firm & Year FEs	X	X	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X	X	X
Observations	29,384	29,384	29,384	29,195	28,523	28,909	26,202	17,536
Firms	3,598	3,598	3,598	3,409	3,550	3,558	3,552	1,915

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	**Future patent applications**							
Forward:	1-year						2-year	3-year
Transformation:	asinh(.)			log(1+.)	win.	raw	asinh(.)	
Specification:	Additional controls			Alt. transformations			Alt. forwards	
CEO's trust	0.066***	0.063***	0.060***	0.053***	1.462**	4.464***	0.046*	0.039*
	(0.019)	(0.019)	(0.018)	(0.015)	(0.584)	(1.293)	(0.027)	(0.023)
log(employment)	0.101***							
	(0.012)							
asinh(R&D stock)		0.015*						
		(0.008)						
asinh(R&D exp.)			0.092***					
			(0.011)					
Dep. var. mean					*13.28*	*18.02*		
Firm & Year FEs	X	X	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X	X	X
Observations	28,506	29,384	29,384	29,384	29,384	29,384	29,384	29,384
Firms	3,548	3,598	3,598	3,598	3,598	3,598	3,598	3,598

Panel C.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**asinh(Future patent applications)**					**D(asinh(pat))**
CEO's trust	0.070***	0.071***		0.041	0.035	
	(0.018)	(0.019)		(0.028)	(0.025)	
Trust × Change in trust		0.013				
		(0.012)				
Post-transition			-0.056***			
			(0.018)			
Post-transition × Trust-increasing			0.087***			
			(0.021)			
Predecessor CEO's trust						-0.016***
						(0.005)
Event sample				Trust increasing	Trust decreasing	
Event & Year FEs	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X
Observations	20,389	20,389	19,504	9,764	9,740	2,444
Events	2,446	2,446	2,343	1,191	1,152	2,444

Notes: This table report the robustness checks for the baseline effect of CEO's inherited trust on firm's patents using equation (1.5). **Panel A:** Column (1) reports the baseline specification in which (i) the sample includes all observations of firm f × year t × its current CEO d; (ii) the dependent variable is the inverse hyperbolic sine of firm f's patent application counts in year $t+1$; (iii) the explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2); (iv) baseline controls include firm's age, age squared, ln(total assets), ln(sale), and CEO's age, age squared, gender, education dummies, tenure in firm; and (v) standard errors are clustered by CEO's main ethnicity. Column (2) clusters standard errors by firm. Column (3) clusters standard errors two-way by CEO's main ethnicity and firm. Column (4) excludes singletons; column (5) female CEOs; column (6) interim CEOs; column (7) CEO transition years. Column (8) estimates a semi-log Poisson count model with winsorized $pat_{f,t+1}$ as the dependent variable. **Panel B:** Column (1) additionally controls for ln(employment); column (2) asinh(R&D stock); column (3) asinh(R&D expenditure). Columns (4)-(6) use $ln(1+pat_{f,t+1})$, winsorized $pat_{f,t+1}$, and raw $pat_{f,t+1}$ as the dependent variable. Columns (7)-(8) use $asinh(pat_{f,t+2})$ and $asinh(pat_{f,t+3})$ as the dependent variable. **Panel C:** This panel employs a sample constructed from CEO transition events and event fixed effects (instead of firm fixed effects). For each event, I include all firm f × year t × its current CEO d observations that correspond to the predecessor's and successor's terms. Column (1) reports the baseline CEO's trust effect using this sample. Column (2) interacts CEO's trust measure with $\Delta trust_E$, the difference between successor and predecessor CEOs' trust measures. Column (3) presents a difference-in- differences specification in which the post-transition dummy is interacted with a dummy indicating the transition is a trust-increasing event. Columns (4) and (5) employ subsamples of trust-increasing and trust-decreasing CEO transition events. Column (6) reports $\hat{\beta}$ from estimating: $\Delta\,\text{asinh}(pat_E) = \beta trust_E^{pre} + \Delta\mathbf{X}_E + \Delta\mathbf{Z}_E + \varepsilon_E$, in which (i) each observation E is a CEO transition event, (ii) $\Delta\,\text{asinh}(pat_E)$, $\Delta\mathbf{X}_E$, and $\Delta\mathbf{Z}_E$ are the differences between post- and pre-transition average patents, firm's, and CEO's characteristics respectively, and (iii) $trust_E^{pre}$ is the trust measure of the predecessor CEO.

*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table A.7: CEO's retirement and death events

Panel A. Including all years in each event

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent var:	**asinh(Next period's patent applications)**						**D(.)**
Sample:	Retired 64-65	Retired 64-66	Retired 63-67	Died transition yr	Died transition yr+1	Retired or died	Retired or died
CEO's trust	0.281***	0.104**	0.083*	0.410	0.400	0.095**	
	(0.095)	(0.042)	(0.047)	(0.308)	(0.279)	(0.045)	
Predecessor CEO's trust							-0.024**
							(0.012)
Observations	913	2,285	3,440	253	353	3,756	386
Events	92	230	346	34	46	386	386

Panel B. Excluding transition years

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent var:	**asinh(Next period's patent applications)**						**D(.)**
Sample:	Retired 64-65	Retired 64-66	Retired 63-67	Died transition yr	Died transition yr+1	Retired or died	Retired or died
CEO's trust	0.328***	0.137**	0.124**	0.736**	0.780***	0.142**	
	(0.103)	(0.060)	(0.056)	(0.363)	(0.300)	(0.054)	
Predecessor CEO's trust							-0.029**
							(0.012)
Event & Year FEs	X	X	X	X	X	X	
Baseline controls	X	X	X	X	X	X	
Observations	825	2,073	3,126	217	306	3,400	377
Events	91	228	342	29	40	377	377

Notes: This table reports CEO's trust effect in subsamples of transitions following CEO's retirements or deaths. Columns (1)-(6) estimate equation (1.5). Each subsample includes all firm f $\times$ year t $\times$ its current CEO d observations that correspond to the predecessor's and successor's terms of the relevant transitions (Panel A) and that are not the transition years (Panel B). Columns (1)-(3)'s subsamples include transitions in which the predecessor CEO retired at 65, between 64 and 66, or between 63 or 67 respectively. Columns (4)-(5)'s subsamples include transitions in which the predecessor CEO died in or within one year of the transition year. Column (6) combines column (3)'s and column (5)'s subsamples. The dependent variable is the inverse hyperbolic sine of firm f's patent application counts in year $t+1$. The explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2). Baseline controls include (i) firm's age, age squared, ln(total assets), ln(sale), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. Column (7) estimates $\Delta\, \text{asinh}(pat_E) = \beta trust_E^{pre} + \Delta\mathbf{X}_E + \Delta\mathbf{Z}_E + \varepsilon_E$, in which (i) each observation E is a CEO transition event included in column (6)'s subsample, and (ii) $trust_E^{pre}$ is the trust measure of the departing CEO. Standard errors are clustered by CEO's main ethnicity in columns (1) to (3) and (6). Robust standard errors are reported for columns (4), (5), and (7).

*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table A.8: Alternative measures of other cultural traits

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**asinh(Future patent applications)**					
CEO's trust	0.067***	0.059***	0.063***	0.073***	0.069***	0.063***
	(0.022)	(0.020)	(0.018)	(0.023)	(0.017)	(0.018)
Self-reported upper class	-0.131					
	(0.196)					
Occupation prestige		0.015				
		(0.015)				
Alt. work ethic			0.008			
			(0.021)			
Alt. risk preference				-0.027		
				(0.026)		
Confidence in government					0.011	
					(0.013)	
Confidence in science						0.032*
						(0.017)
Firm & Year FEs	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X
Observations	29,384	29,384	29,384	29,384	29,384	29,384
Firms	3,598	3,598	3,598	3,598	3,598	3,598

Notes: This table explores alternative GSS-based measures of CEO's other inherited cultural traits as additional controls in equation (1.5). Baseline sample includes all observations of firm f $\times$ year t $\times$ its current CEO d. The dependent variable is the inverse hyperbolic sine of firm f's patent application counts in year $t+1$. The explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2). Baseline controls include (i) firm's age, age squared, ln(total assets), ln(sale), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. Column (1) controls for the share of self-reported upper class in CEO's ethnic groups. Column (2) controls for average GSS occupational prestige score among those in CEO's ethnic groups. Column (3) controls for CEO's inherited work ethic, derived from the GSS question: *"If you were to get enough money to live as comfortably as you would like for the rest of your life, would you continue to work or would you stop working?"*. Column (4) controls for CEO's inherited risk preference, proxied by the share of GSS respondents in CEO's ethnic groups who consider job security as the least important feature of a job. Columns (5) and (6) control for CEO's inherited confidence in the government and in the scientific community, derived from the GSS question: *"I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?"*. Cultural trait controls in columns (2)-(6) are standardized by their standard deviations at ethnicity level. Standard errors are clustered by CEO's main ethnicity.

*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table A.9: Global Preference Survey's trust and other cultural traits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	**asinh(Future patent applications)**						
CEO's trust (GPS)	0.040**	0.039***	0.036**	0.041**	0.046***	0.041**	0.043***
	(0.015)	(0.014)	(0.015)	(0.016)	(0.013)	(0.016)	(0.013)
Risk preference		0.003					0.019
		(0.017)					(0.027)
Patience			0.010				-0.016
			(0.014)				(0.026)
Positive reciprocity				-0.005			0.003
				(0.018)			(0.033)
Negative reciprocity					-0.023***		-0.031***
					(0.007)		(0.010)
Altruism						-0.003	0.004
						(0.014)	(0.028)
Firm & Year FEs	X	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X	X
Observations	29,384	29,384	29,384	29,384	29,384	29,384	29,384
Firms	3,598	3,598	3,598	3,598	3,598	3,598	3,598

Notes: This table employs inherited trust measure and other cultural trait measures constructed in the same way as described in subsection 1.3.2 but using the Global Preference Survey (GPS) (Falk et al., 2018). Baseline sample includes all observations of firm f $\times$ year t $\times$ its current CEO d. The dependent variable is the inverse hyperbolic sine of firm f's patent application counts in year $t+1$. The explanatory variable is CEO d's GPS-based inherited trust measure, standardized by its standard deviation at ethnicity level. Baseline controls include (i) firm's age, age squared, ln(total assets), ln(sale), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. Column (1) reports the baseline effect of CEO's GPS-based inherited trust. Column (2) controls for CEO's inherited risk preference; column (2) patience; column (3) positive reciprocity; column (4) negative reciprocity; column (5) altruism. Column (6) controls for all those variables. Standard errors are clustered by CEO's main ethnicity.

*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table A.10: CEO's trust effect in US-only bilateral trust sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	**asinh(Future patent applications)**						
Sample:	Based on last names of US-based inventors						
CEO's bilateral trust	0.019*	0.019*	0.011	0.005	0.017	0.018	0.020*
	(0.011)	(0.011)	(0.007)	(0.012)	(0.012)	(0.011)	(0.011)
Common language dummy					0.017		
					(0.024)		
Geographical distance (1000km)						-0.008	
						(0.010)	
Genetic distance (z-score)							-0.019
							(0.049)
Excl. same-country pairs				X			
Firm × Year FEs	X	X		X	X	X	X
CEO FEs	X		X	X	X	X	X
Inventor country FEs	X			X	X	X	X
CEO × Year FEs		X					
Inv. country × Year FEs		X					
Firm × Inv. country's FEs			X				
Year FEs			X				
Observations	53,967	53,967	53,967	49,497	52,769	52,769	51,334
Firm × Inv. country's	8,240	8,240	8,240	7,661	8,051	8,051	7,828
Firms	1,186	1,186	1,186	970	997	997	997

Notes: This table reports the effect of CEO's bilateral trust towards a country on patents by inventors from that country using equation (1.6). Samples include all observations of firm f × year t × its current CEO d × country c such that firm f has patents by inventors from country c during 2000-2012. An inventor's country is inferred from his last name only for US-based inventors. The explanatory variable is CEO d's bilateral trust towards individuals from country c, standardized by its standard deviation at country pair level. The dependent variable is firm f's total patent application counts by inventors from country c in year $t+1$. Column (4) excludes same-country CEO-inventor country pairs. Columns (5) to (7) control for CEO-inventor country pairwise distances, including: (i) whether the countries share a common language (column 5), (ii) weighted geographical distance between the countries (column 6), and (iii) weighted genetic distance between the countries' populations (column 7) (Spolaore and Wacziarg, 2016). Standard errors are clustered by CEO's main ethnicity × inventor country.
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table A.11: Margins of CEO's bilateral trust effect

Panel A. Bilateral trust sample based on inventors' non-US addresses

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**asinh(Future patent applications)**					
	By firms in bila-		By inventor groups			
Sample:	teral trust sample		Both margins		Intensive margin	
CEO's generalized trust	0.101***	0.126***				
	(0.037)	(0.027)				
CEO's bilateral trust			0.047*	0.037*	0.053	0.072
			(0.024)	(0.020)	(0.049)	(0.049)
Observations	7,356	6,915	22,450	22,450	9,065	9,065
Firm × Inv. country's			3,383	3,383	1,673	1,673
Firms	724	724	700	700	437	437

Panel B. Bilateral trust sample based on inventors' addresses and last names

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**asinh(Future patent applications)**					
	By firms in bila-		By inventor groups			
Sample:	teral trust sample		Both margins		Intensive margin	
CEO's generalized trust	0.071**	0.095***				
	(0.031)	(0.028)				
CEO's bilateral trust			0.025**	0.015**	0.031**	0.020**
			(0.011)	(0.007)	(0.013)	(0.009)
Firm & Year FEs	X	X				
Baseline controls		X				
Firm × Year FEs			X		X	
CEO FEs			X	X	X	X
Inventor country FEs			X		X	
Firm × Inv. country FEs				X		X
Year FEs				X		X
Observations	12,764	11,931	53,967	53,967	38,467	38,467
Firm × Inv. country's			8,240	8,240	6,418	6,418
Firms	1,256	1,256	1,186	1,186	925	925
	724	724	700	700	437	437

Notes: This table reports CEO's generalized and bilateral trust effects on patents among the samples of firms included in bilateral trust analyses. In Panel A, an inventor's country is inferred from his patent-listed address for non-US-based inventors; in Panel B, an inventor's country is additionally inferred from his last name for US-based inventors. In each panel, columns (1) and (2) report CEO's generalized trust effect on firm's patents among the sample of firms included in the corresponding panel's bilateral trust sample, using equation (1.5) (i.e., observation unit is firm f × year t × its current CEO d, see notes to Table 1.1 for further details). Columns (3)-(6) report CEO's bilateral trust effect on inventors' patents using equation (1.6) (i.e., observation unit is firm f × year t × its current CEO d × country c, see notes to Table 1.3 for further details). Columns (3) and (4) employ all observations such that firm f has patents by inventors from country c during 2000-2012. The resulting coefficients capture both intensive and extensive margins of CEO's bilateral trust effect. Columns (5) and (6) employ only observations such that firm f has patents by inventors from country c before CEO d assumes position. The resulting coefficients capture only the intensive margin of CEO's bilateral trust effect. Standard errors are clustered by CEO's main ethnicity in columns (1) and (3) and by CEO's main ethnicity × inventor country in columns (3)-(6).
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table A.12: Directions of CEO-inventor bilateral trust

Panel A. Bilateral trust sample based on inventors' non-US addresses

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	**asinh(Future patent applications)**				
CEO-toward-inventors bilateral trust	0.102***		0.100**	0.138***	
	(0.025)		(0.044)	(0.052)	
Inventors-toward-CEO bilateral trust		0.076**	0.003		-0.016
		(0.032)	(0.047)		(0.055)
Observations	12,863	12,863	12,863	12,863	12,863
Firm × Inventor country's	2,009	2,009	2,009	2,009	2,009
Firms	580	580	580	580	580

Panel B. Bilateral trust sample based on inventors' addresses and last names

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	**asinh(Future patent applications)**				
CEO-toward-inventors bilateral trust	0.041**		0.030	0.049*	
	(0.016)		(0.026)	(0.028)	
Inventors-toward-CEO bilateral trust		0.037***	0.013		0.029
		(0.014)	(0.023)		(0.026)
Firm × Year FEs	X	X	X	X	X
CEO FEs	X	X	X	X	X
Inventor country FEs	X	X	X	X	X
Inventors-to-CEO trust decile FEs				X	
CEO-to-inventors trust decile FEs					X
Observations	32,648	32,648	32,648	32,648	32,648
Firm × Inventor country's	5,005	5,005	5,005	5,005	5,005
Firms	1,072	1,072	1,072	1,072	1,072

Notes: This table explores the effects of different directions of bilateral trust on patents using equation (1.6). Samples include all observations of firm f × year t × its current CEO d × country c such that (i) firm f has patents by inventors from country c during 2000-2012, and (ii) both bilateral trust variables are non-missing. An inventor's country is inferred from his patent-listed address for non-US-based inventors in Panel A, and additionally from his last name for US-based inventors in Panel B. The explanatory variables are (i) CEO d's bilateral trust towards individuals from country c, and (ii) individuals from country c's bilateral trust towards CEO d, both standardized by their same standard deviations at country pair level. The dependent variable is firm f's total patent application counts by inventors from country c in year $t+1$. Decile dummies in columns (4) and (5) are computed with respect to the relevant bilateral trust sample. Standard errors are clustered by CEO's main ethnicity × inventor country.

*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table A.13: Bilateral trust effect by patent quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	**asinh(Future patents in each quality quartile)**							
Sample:	Based on non-US addresses				Based on addresses/last names			
Quality quartile:	1st	2nd	3rd	4th	1st	2nd	3rd	4th
CEO's bilateral trust	0.017	0.014	0.017	0.027*	0.004	0.006	0.008	0.024***
	(0.012)	(0.011)	(0.015)	(0.016)	(0.006)	(0.005)	(0.007)	(0.009)
Firm × Year FEs	X	X	X	X	X	X	X	X
CEO FEs	X	X	X	X	X	X	X	X
Inventor country FEs	X	X	X	X	X	X	X	X
Observations	23,284	23,284	23,284	23,284	56,942	56,942	56,942	56,942
Firm × Inv. country's	3,481	3,481	3,481	3,481	8,554	8,554	8,554	8,554
Firms	730	730	730	730	1,263	1,263	1,263	1,263

Notes: This table reports the heterogenous effects of CEO's bilateral trust on patents in different quality quartiles using equation (1.6). Samples include all observations of firm f × year t × its current CEO d × country c such that firm f has patents by inventors from country c during 2000-2012. An inventor's country is inferred from his patent-listed address for non-US-based inventors in columns (1)-(4), and additionally from his last name for US-based inventors in columns (5)-(8). The explanatory variable is CEO d's bilateral trust towards individuals from country c, standardized by its standard deviation at country pair level. The dependent variable is firm f's total patent application counts by inventors from country c in year $t+1$ in each patent quality quartile, with 1 being the bottom quartile and 4 the top. A patent's quality quartile is computed based on its forward citation counts with respect to its technology field × year cohort. Standard errors are clustered by CEO's main ethnicity × inventor country.

*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table A.14: Bilateral trust effect on quality-weighted patents

Panel A. Bilateral trust sample based on inventors' non-US addresses

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	**asinh(Future quality-weighted patents)**						
Quality measure:	Forward cites	Backward NPL cites	Tech scope	Gene-rality	Origi-nality	Granted all	Granted USPTO
CEO's bilateral trust	0.095**	0.053*	0.100***	0.023	0.032*	0.040*	0.038*
	(0.039)	(0.032)	(0.034)	(0.015)	(0.018)	(0.021)	(0.020)
Observations	23,284	23,284	23,284	23,284	23,284	23,284	23,284
Firm × Inv. country's	3,481	3,481	3,481	3,481	3,481	3,481	3,481
Firms	730	730	730	730	730	730	730

Panel B. Bilateral trust sample based on inventors' addresses and last names

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	**asinh(Future quality-weighted patents)**						
Quality measure:	Forward cites	Backward NPL cites	Tech scope	Gene-rality	Origi-nality	Granted all	Granted USPTO
CEO's bilateral trust	0.051***	0.046***	0.051***	0.020**	0.022**	0.023**	0.027**
	(0.018)	(0.016)	(0.016)	(0.008)	(0.009)	(0.010)	(0.011)
Firm × Year FEs	X	X	X	X	X	X	X
CEO FEs	X	X	X	X	X	X	X
Inventor country FEs	X	X	X	X	X	X	X
Observations	56,942	56,942	56,942	56,942	56,942	56,942	56,942
Firm × Inv. country's	8,554	8,554	8,554	8,554	8,554	8,554	8,554
Firms	1,263	1,263	1,263	1,263	1,263	1,263	1,263

Notes: This table reports CEO's bilateral trust effect on quality-weighted patents using equation (1.6). Samples include all observations of firm f × year t × its current CEO d × country c such that firm f has patents by inventors from country c during 2000-2012. An inventor's country is inferred from his patent-listed address for non-US-based inventors in Panel A, and additionally from his last name for US-based inventors in Panel B. The explanatory variable is CEO d's bilateral trust towards individuals from country c, standardized by its standard deviation at country pair level. The dependent variable is firm f's total patent application counts by inventors from country c in year $t+1$, weighted by: forward citations (column 1); backward citations to non-patent (i.e., scientific) literature (column 2); patent technological scope (column 3); generality index (i.e., technological diversity of forward citations) (column 4); originality index (i.e., technological diversity of backward citations) (column 5); granted patents (column 6); and USPTO patents (column 7). Standard errors are clustered by CEO's main ethnicity × inventor country.
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table A.15: Effect of CEO's trust on R&D

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**asinh(R&D expenditure)**			**asinh(R&D stock)**		
Forward :	0 year	1 year	2 year	0 year	1 year	2 year
CEO's trust	0.028	0.018	0.020	-0.014	0.001	0.018
	(0.032)	(0.027)	(0.029)	(0.019)	(0.018)	(0.019)
Firm & Year FEs	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X
Observations	29,384	28,125	26,710	29,384	28,125	26,710
Firms	3,598	3,558	3,487	3,598	3,558	3,487

Notes: This table reports the baseline effect of CEO's inherited trust on R&D expenditure and stock using equation (1.5). Baseline sample includes all observations of firm f × year t × its current CEO d. The dependent variable is the inverse hyperbolic sine of firm f's R&D expenditure or stock in year $t+k$m for $k \in [1,3]$. The explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2). Baseline controls include (i) firm's age, age squared, ln(total assets), ln(sale), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. Standard errors are clustered by CEO's main ethnicity.
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Table A.16: Effect of CEO's trust on firm future performance

Panel A.

Dependent variable: Future	(1) **ln(sales)**	(2) **ln(employment)**	(3) **ln(capital)**	(4) **TFP (KL)**	(5) **TFP (KLM)**
CEO's trust	-0.036**	-0.032**	-0.022	0.006	0.000
	(0.017)	(0.013)	(0.025)	(0.013)	(0.026)
Trust × Proxy for pre-transition researcher quality	0.048***	0.034***	-0.000	0.015	0.031
	(0.014)	(0.010)	(0.013)	(0.011)	(0.018)
Firm & Year FEs	X	X	X	X	X
Baseline controls	X	X	X	X	X
Observations	18,019	17,873	16,782	17,238	7,719
Events	2,237	2,224	2,149	2,177	1,421

Panel B.

Dependent variable: Future	(1) **ln(sales)**	(2) **ln(employment)**	(3) **ln(capital)**	(4) **TFP (KL)**	(5) **TFP (KLM)**
Trust × Quality quintile 1	-0.153***	-0.154***	-0.068	0.007	-0.008
	(0.040)	(0.035)	(0.047)	(0.036)	(0.041)
Trust × Quality quintile 2	-0.119**	-0.044	-0.063*	-0.046	-0.076
	(0.046)	(0.030)	(0.036)	(0.030)	(0.052)
Trust × Quality quintile 3	0.039	0.002	0.006	0.035	-0.005
	(0.058)	(0.043)	(0.049)	(0.060)	(0.106)
Trust × Quality quintile 4	0.058**	0.069**	0.104**	-0.008	-0.026
	(0.024)	(0.026)	(0.051)	(0.032)	(0.052)
Trust × Quality quintile 5	0.034	-0.000	-0.056	0.051*	0.117***
	(0.045)	(0.027)	(0.050)	(0.027)	(0.041)
Firm & Year FEs	X	X	X	X	X
Baseline controls	X	X	X	X	X
Observations	18,019	17,873	16,782	17,238	7,719
Events	2,237	2,224	2,149	2,177	1,421

Notes:
This table explores the heterogeneous effects of CEO's trust on firm's patents by pre-transition researcher pool quality using equation (1.5) and the sample constructed from CEO transition events. For each event, I include all firm f × year t × its current CEO d observations that correspond to the predecessor's and successor's terms. The explanatory variable is CEO d's GSS-based inherited trust measure, standardized by its standard deviation at ethnicity level (subsection 1.3.2). Firm-level proxy for researcher pool quality is computed from averaging the residuals from regressing patents on observable firm and CEO characteristics, controlling for SIC2 industry and year fixed effects (subsection 1.6.3) over a 2-year pre-transition window. The dependent variable is firm f's performance in year $t+2$, including: ln(sales) (column 1); ln(employment) (column 2); ln(capital) (column 3); TFP computed from value added, employment, and capital following Olley and Pakes (1996) (column 4); and TFP computed from sales, employment, capital, and material following Olley and Pakes (1996) (column 5). Baseline controls include (i) firm's age, age squared, asinh(R&D expenditure), and (ii) CEO's age, age squared, gender, education dummies, tenure in firm. Panel A interacts CEO's trust measure with firm-level proxy for pre-transition pool quality. Panel B interacts CEO's trust measure with researcher pool quality quintile dummies (computed based on firm-level proxy for pre-transition researcher pool quality). Standard errors are clustered by CEO's main ethnicity.
*** denotes statistical significance at 1% level, ** 5% level, * 10% level.

Appendix B

Appendices to Chapter 2: Do Tax Incentives Increase Firm Innovation?

B.1 Institutional details of policy and tax-adjusted user cost

B.1.1 SME definition

The UK R&D Tax Relief Scheme's SME (Small and Medium Sized Enterprise) definition is based on total assets ("balance sheet total"), employment ("staff headcount"), and sales ("turnover") as described in Section 2.2. We summarize the key elements of the definition rules below but for further technical details on these rules see `http://www.hmrc.gov.uk/manuals/cirdmanual/CIRD91400.htm`.

Measurements of staff headcount, assets, and sales turnover for ceiling tests: Assets is the gross amount of assets shown in the company accounts. The staff headcount of an enterprise represents the number of full-time person-years attributable to people who have worked within or for the enterprise during the year under consideration.[1] The staff headcount and financial data used for the "ceiling tests" (the maximum values possible for a firm to be eligible for SME status) are those relating to the latest accounting year. Assets and sales are converted to Euros using the exchange rate on the last day of the relevant accounting period, or the average exchange rate throughout that accounting period (whichever is more beneficial for the enterprise). An enterprise passes the ceiling tests if its staff headcount and either its aggregated assets or its aggregated turnover fall below the respective ceilings. An enterprise loses (acquires) its SME status if it fails (passes) the ceiling tests over two consecutive accounting periods.

Account aggregation rules for different enterprise types: In the case of an autonomous enterprise, the staff headcount and financial data are determined exclusively on the basis of the consolidated account of the enterprise itself.[2] In the case of a "linked" enterprise, the ceiling tests are applied to the aggregates of the figures in its own accounts and those from the accounts of all other enterprises to which it is linked (including non-UK ones), unless the linked enterprises' account data are already included through account consolidation.[3]

B.1.2 UK R&D Tax Relief Scheme

The R&D Tax Scheme includes a SME Scheme and a Large Company ("LCO") component.[4] Between its introduction in 2000 and 2012, more than 28,500 different

[1] The contributions of part-time workers, or those who work on a seasonal or temporary basis count as appropriate fractions of a full-time person-year. The term staff includes employees, persons seconded to the enterprise, owner-managers, partners (other than sleeping partners); it excludes apprentices or students engaged in vocational training with an apprenticeship or vocational training contract, and any periods or maternity or parental leave.

[2] An autonomous enterprise is one that is not a linked enterprise or a partner enterprise. Generally, an enterprise is autonomous if it has holding of less than 25% of the capital or voting rights in one or more enterprises and/or other enterprises do not have a stake of 25% or more of the capital voting rights in the enterprise.

[3] Linked enterprises are those in which one enterprise is able to exercise control, directly or indirectly, over the affairs of the other.

[4] For further details, see `http://www.hmrc.gov.uk/manuals/cirdmanual/CIRD90000.htm` (SME Scheme) and `http://www.hmrc.gov.uk/manuals/cirdmanual/CIRD85050.htm` (Large Company Scheme).

companies had made claims under the SME Scheme, and over 7,000 under the Large Company Scheme, claiming more than £9.5bn in total R&D support. The annual amount of R&D support had risen to over £1bn by 2008, reaching £1.4bn in 2012, and covered qualifying R&D expenditure worth £13.2bn HMRC (2014).

Both SME and Large Company Schemes are volume-based, i.e., the tax relief accrues on the total R&D spending rather than the incremental R&D over a prior base (the main US R&D tax relief scheme is incremental). It works mostly through enhanced deduction of current R&D expenditure from taxable income, thus reducing R&D-performing companies' corporate tax liabilities. For example, if a company is allowed an enhancement rate of 75%, for a £10,000 spend on R&D, it can deduct an additional £7,500 from its taxable income before calculating its tax liability. In addition, under the SME Scheme, a company that has taxable loss after the additional deduction can also claim payable tax credit up to the amount of payable credit rate × enhanced qualifying R&D expenditure[5] This payable tax credit can only be used to reduce the company's employers' payroll tax (National Insurance Contributions, NIC) liabilities. Alternatively, the company (either as an SME or as a large company) can choose to carry the loss forward as normal. Qualifying R&D expenditure must be allowable as a deduction in calculating trading profits, which includes all flow costs, employee costs, materials, utilities, software, or subcontracted R&D expenditure (but only if the contractor is an SME).[6] To be eligible for R&D tax relief, a company must also spend at least £10,000 a year on qualifying R&D expenditure in an accounting period. If an SME works as a subcontractor for a large company, only the subcontractor SME can claim R&D tax relief, under the Large Company Scheme.[7] There is also an upper limit of €7.5m on the total amount of aid a company can receive for a R&D project under the SME Scheme.

The evolution of the UK R&D Tax Relief Scheme is summarized in Table A1. It was first introduced in April 2000 only for SMEs (Finance Act 2000),[8] then later extended to

[5]For example, if a company is allowed an enhancement rate of 75% and payable credit rate of 14%, spends £10,000 in R&D, and has no taxable income before the additional deduction, it can claim payable tax credit of 0.14 ×£10,000 × (1+0.75)=£2,450. If instead the company has £1,500 in taxable income before the additional deduction, it can first use £2,000 of its R&D to reduce its taxable income to zero (i.e., £1,500 = 75% x £2000, then claim payable tax credit of 0.14 × £8,000×(1+0.75)=£1,960. This latter case is called a combination claim.

[6]A large company that has taxable loss before the additional deduction therefore may still benefit from R&D tax relief by carrying the "enhanced" loss forward to further reduce its taxable income in the next period. However, this reduction is only meaningful when the company has enough taxable income in this next period.

[7]An SME already receiving another form of notified state aid for a project cannot claim R&D tax relief for that same project under the SME Scheme (which is also a notified state aid), as total state aid intensity cannot exceed 25% under European Commission's State Aid rules. However, from April 2003 onward, SMEs are allowed to claim R&D tax relief for such projects under the Large Company Scheme.

[8]Qualifying R&D expenditure could include R&D performed outside of the UK by foreign branches of UK holding companies, as foreign branches' revenues and costs are directly consolidated into their UK holding companies' tax revenues and costs for UK tax purpose. Qualifying R&D expenditure is unlikely to include R&D performed outside of the UK by foreign subsidiaries of UK holding companies, as foreign subsidiaries' net profits are indirectly incorporated into their UK holding companies' tax revenues as dividends for UK tax purpose instead.

large companies starting from April 2002 (Finance Act 2002).[9] Between April 2000 and December 2004 the ceilings for staff headcount, assets, and sales were 249, €27m, and €40m respectively. From January 2005, they were raised to 249, €43m, and €50m. This followed European Union guidelines for SME definitions. Throughout the period from April 2000 (April 2002) to March 2008, the enhancement rates were set at 50% for SMEs and 25% for large companies, and the payable credit rate for SMEs was 16%.[10]

As discussed in the main paper, various changes to the scheme became effective at different points in 2008. First, from April 2008, the enhancement rate for large companies was increased from 25% to 30%. Then from August 2008, the enhancement rate for SMEs was increased from 50% to 75% and the payable credit rate for SMEs was reduced from 16% to 14% (to ensure that state aid intensity stays below the EU imposed limit of 25%). Also from August 2008, the SME Scheme was extended to "larger" SMEs as the SME ceilings were doubled to 499, €86m, and €100m for staff headcount, assets, and sales respectively. This change in SME definition is applicable only for the purpose of the R&D tax relief and therefore is the main focus of our paper, as it allows us to separate the impacts of the R&D Tax Relief Scheme from other programs. It should also be noted that while these new SME ceilings were announced in Finance Act 2007, the date on which they became effective (August 1st, 2008) was announced much later, in July 2008.[11]

There were tweaks to the system in 2011 and 2012. From April 2011, the SME enhancement rate was increased to 100% and the SME payable credit rate was reduced to 12.5%. From April 2012, the SME enhancement rate was again increased to 125%. However, the SME definition as announced in Finance Act 2007 and the large company enhancement rate of 30% remained unchanged throughout this period.

The formal definition of R&D has been stable. To qualify for tax relief the costs must be consistent with the UK accounting definition of R&D under GAAP (accounting standards FRS102 s18, IAS38, FRS105 s13 and SSAP13). "To quality for R&D, a company must be undertaking a project to seek an advance in science or technology through the resolution of scientific or technological uncertainties. The advance being sought must constitute an advance in the overall knowledge or capability in a field of science or technology, not a company's own state of knowledge or capability alone."

B.1.3 A Simple Model of patents and R&D demand

Consider a CES production function in R&D capital (G) and non-R&D capital (Z). If input markets are competitive we can write the long-run static first order condition for

[9] Finance Act 2000 (Chapter 17, Schedule 20) and Finance Act 2002 (Chapter 23, Schedule 12).

[10] One exception to this differential treatment of SMEs and large companies was the Vaccine Research Relief Scheme (VRR) launched in April 2003, which extended the higher 50% additional allowance to cover specific areas of vaccine and drug research conducted in large companies (Finance Act 2003, Chapter 14, Schedule 31). The VRR enhancement rate was later reduced to 40% from August 2008 onward.

[11] Finance Act 2007, Section 50 (Appointed Day) Order 2008 of July 16th, 2008.

factor demand of the firm as:

$$lnG = -\sigma ln\rho + \sigma lnU + lnZ + B \tag{B.1}$$

where ρ is the user cost of R&D capital, U is is the user cost of non-R&D capital and B is a technological constant reflecting factor bias terms in the production function. Assume that G can be described by the perpetual inventory formula $G_t = (1-\delta)G_{t-1} + R_t$ where R is the R&D expenditure in period t. Since in steady state, the R&D just offsets the depreciated part of the R&D stock $\delta G = R$, we can re-write the first order condition in steady state as:

$$lnR = -\sigma ln\rho + \sigma lnU + lnZ + ln\delta + B \tag{B.2}$$

This is essentially the equation we estimate in equation (2.1). We also consider a knowledge production function:

$$lnPAT = \mu + \alpha lnG$$

Substituting the R&D first order condition into this "structural" patent equation generates our key reduced form patent equation:

$$\ln PAT = -\alpha\sigma ln\rho + \alpha lnZ + \alpha\sigma lnU + \alpha ln\delta + \alpha B - \mu$$

This is essentially what we estimate in equation (2.2). Around the R&D SME threshold the user cost of non-R&D capital and technology are assumed to be smooth. Non-R&D capital (assets) is the running variable so we have a polynomial approximation to lnZ.

The main departure from the R&D and patent equations above is that the presence of firms with zero patents and/or R&D means we cannot take logarithms. So we use levels instead of logs as dependent variables. To obtain the logarithmic (proportional) changes we use the empirical averages of the dependent variable in the pre-policy period. We also show that the calculations are robust to using a Poisson regression whose first moment is the exponential log-link function and so is equivalent to estimating in logarithms.

B.1.4 Estimating the instrument's sharpness using a subsample

Our approach is a fuzzy RD Design. Equations (2.1) and (2.3) are the first stage and structural form of a knowledge (patent) production function. But as discussed in subsection 2.7.2 we may also be interested in the elasticity of R&D with respect to its tax-adjusted use cost. To do this we need to scale the estimate in equation (2.1) by the "sharpness" of the IV. Consider equation (2.6):

$$SME_i = \alpha_6 + \lambda E_{i,2007} + f_6(z_i, 2007) + \epsilon_{6i}$$

Recall that $E_{i,2007}$ is a binary indicator of firm i's being below the new asset threshold in 2007 and SME_i is a binary indicator of the firm's true SME eligibility (which is observable only for R&D performing firms). Let $\lambda_E = Pr(SME = 1|E, Z)$ for $E \in \{0, 1\}$ in the full baseline sample of both R&D performing and non-R&D performing firms. For the sharpness of $E_{i,2007}$ as an instrument for firm's SME-scheme eligibility, we would like to estimate $\lambda \equiv \lambda_1 - \lambda_0$. The problem is that we only observe SME_i for the subsample of R&D performing firms as (a) this data is not in HMRC datasets for non-R&D performers and (b) we cannot calculate eligibility status with precision from the accounting variables. Thus we can only estimate equation (2.6) on the R&D performers subsample. Under the RD Design identification assumptions discussed in Section 2.3, the resulting $\hat{\lambda}$ from this regression is a consistent estimate for $\tilde{\lambda} \equiv \widetilde{\lambda_1} - \widetilde{\lambda_0}$, where $\widetilde{\lambda_E} = Pr(SME = 1|E, Z, R > 0)$ for $E \in \{0, 1\}$. When will $\tilde{\lambda}$ be equal to λ? We will prove that a sufficient condition for this is that SME-scheme eligibility does not change firm's likelihood of performing R&D, which is something we test (and find empirical support for) in the data.

Let p_S and p_L are the probabilities a firm will perform R&D if it is eligible for the SME scheme (p_S), and if it is not (p_L), and $\rho \equiv \frac{p_S}{p_L}$. Note that by RD Design, we can assume that p_S (and p_L) is the same for firms just below and above the threshold. In the subsample of R&D performing firms, we then have:

$$\widetilde{\lambda_E} = Pr(SME = 1|E, Z, R > 0) = \frac{\lambda_E p_S}{\lambda_E p_S + (1 - \lambda_E) p_L}$$

Expanding and rearranging $\widetilde{\lambda_1} - \widetilde{\lambda_0}$ gives:

$$\begin{aligned}
\widetilde{\lambda_1} - \widetilde{\lambda_0} &= (\lambda_1 - \lambda_0) \frac{p_S p_L}{[\lambda_1 p_S + (1 - \lambda_1) p_L][\lambda_0 p_S + (1 - \lambda_0) p_L]} \\
\Rightarrow \tilde{\lambda} &= \lambda \frac{\rho}{(\lambda_1 \rho + 1 - \lambda_1)(\lambda_0 \rho + 1 - \lambda_0)} \\
&= \lambda \{1 + \frac{(\rho - 1)[(1 - \lambda_1)(1 - \lambda_0) - \lambda_1 \lambda_0 \rho]}{[1 + \lambda_1(\rho - 1)][1 + \lambda_0(\rho - 1)]}\}
\end{aligned}$$

When SME-scheme eligibility does not change firm's likelihood of performing R&D $\rho = 1$ (i.e. $p_S = p_L$). In this case $\tilde{\lambda} = \lambda$. Table B.7 Panel A shows that the policy does not appear to increase firm's participation in R&D performance, suggesting that $p_S \approx p_L$ or $\rho \approx 1$ holds in our setting.[12] This implies that $\tilde{\lambda} \approx \lambda$ in a first-order approximation (as $\frac{(\rho-1)[(1-\lambda_1)(1-\lambda_0)-\lambda_1\lambda_0\rho]}{[1+\lambda_1(\rho-1)][1+\lambda_0(\rho-1)]} \approx 0$).[13]

Finally, consider the sign of the second-order bias when ρ is not exactly 1. If $\rho > 1$, the sign of the bias depends on $(1 - \lambda_1)(1 - \lambda_0) - \lambda_1 \lambda_0 \rho$ which can be either negative or positive. When $\lambda_1 + \lambda_0 \geq 1$ (i.e., sufficiently large share of SME firms in the full baseline

[12]Formally, the regressions in Table B.7 Panel A estimate $\Delta_p = Pr(R > 0|E = 1, Z) - Pr(R > 0|E = 0, Z) = [\lambda_1 p_S + (1 - \lambda_1) p_L] - [\lambda_0 p_S - (1 - \lambda_0) p_L] = (\lambda_1 - \lambda_0)(p_s - p_L)$. $\Delta_p = 0$ implies that $p_S - p_L = 0$ under the reasonable assumption that $\lambda_1 - \lambda_0$. In addition, Table B.8 provides further evidence that the policy effect on R&D is entirely driven by pre-policy R&D performing firms, whose decisions to engage in R&D performance in the pre-policy period did not depend on their post-policy SME status.

[13]Note that although $p_S = p_L$ is a sufficient condition, it is not a necessary condition. $\tilde{\lambda} \approx \lambda$ also if (i) $\lambda = 0$, (ii) $\lambda_1 = 1$ and $\lambda_0 = 0$ (or vice versa), or (iii) $\rho = \frac{(1-\lambda_1)(1-\lambda_0)}{\lambda_1 \lambda 0}$.

sample) $(1-\lambda_1)(1-\lambda_0) \leq \lambda_1\lambda_0 < \lambda_1\lambda_0\rho$, which implies that the bias is negative. However, when $\lambda_1 + \lambda_0 < 1$, the bias could still be either negative or positive.

B.1.5 Tax-adjusted user cost of R&D

The full formula for tax-adjusted user cost of R&D as described in sub-section 7.2 is:

$$\begin{aligned}\rho_{t,f} = \quad & (Pr(\textit{Has tax liability}) \times \frac{(1-\tau_t(1+e_{t,f}))}{(1-\tau_t)} \\ & + Pr(\textit{No tax liability}) \times (1 - c_{t,f}(1+e_{t,f}))) \times (r+\delta)\end{aligned}$$

where τ is the effective corporate tax rate, e is the enhancement rate, c is the payable credit rate, r is the real interest rate, δ is the depreciation rate, t denotes year, and f denotes the whether the company is an SME or a large company. Note that $\rho_{t,f}$ varies over time with τ_t, $e_{t,f}$, and $c_{t,f}$.

For simplicity, we do not consider the possibility that a loss-making large company may still benefit from R&D tax relief by carrying the "enhanced" loss forward to future years to reduce its taxable income, as this reduction is only meaningful if the company makes enough profits in this next period. This simplification may overestimate large companies' tax-adjusted user cost of R&D and, as a result, underestimate the R&D tax-price elasticity (by overestimating the difference in tax-adjusted user cost of R&D between SMEs and large companies). We also do not consider combination claims (cases in which an SME combines tax deduction with the payable tax credit) as there are almost none of them in our baseline sample.

The evolution of tax adjusted user costs of R&D for SMEs and large companies over time is summarized in Table B.2. For large companies (for which the payable credit rate is always zero), there are slight decreases in the corporate tax rate over 2006-12 (from 30% to 28% to 26%) coupled with slight increases in the enhancement rate (from 25% to 30%) over the same period. This resulted in a relatively stable tax-adjusted user cost of 0.190 throughout this period. It is therefore reasonable to use the baseline sample's average R&D over 2006-08 as a proxy for how much an average firm *i*n the baseline sample would spend on R&D if it remained a large company over 2009-11, after the policy change. For SMEs, large increases in enhancement rate (from 50% to 75% to 100%) more than offset the slight decrease in corporate tax rate and payable credit rate (from 16% to 14% to 12.5%), leading to a steady reduction in SMEs' tax-adjusted user cost of R&D from 0.154 in 2006 to 0.141 in 2011. This widens the difference in tax-adjusted user cost of R&D between SMEs and large companies over time, from an average percentage difference of -0.218 over 2006-08 to -0.269 over 2009-11.

Finally, as a robustness check, we also consider using the small firm profit rate (from 19% to 21% to 20% over 2006-11) instead of the main rate for corporate tax rate. As the tax deduction is less generous with a lower corporate tax rate, the resulting tax-adjusted user cost in the tax deduction case is higher for both SMEs and large companies and their gap is smaller in magnitude (average percentage difference over 2006-08 is -0.185 and over 2009-11 is -0.228).

B.1.6 Macro aspects of the R&D Tax Relief Scheme

A full welfare analysis of the R&D Tax Relief Scheme requires both an analysis of the benefits in terms of (say) the increased GDP generated by the R&D induced by the policy (including spillovers) and the deadweight cost of taxation. We would also need to take a position on other general equilibrium effects such as the increase in the wages of R&D workers due to increased demand (Goolsbee, 1998). As an interim step towards this we follow the convention in the literature which is to calculate a "value for money" ratio $\mu \equiv \frac{\Delta_R}{\Delta_{EC}}$ where Δ_R is the amount of R&D induced by the policy and Δ_{EC} is the total amount of additional taxpayer money needed to pay for the scheme (which we call "Exchequer Cost", EC).

We consider three policy-relevant experiments. First, we look at the 2008 extension of the SME Scheme. Second, we do a "value for money" calculation in our data period 2006-11. Finally, we do a simulation of what the path of UK business R&D to GDP would have been with and without the R&D Tax Relief Scheme.

2008 extension of the SME Scheme

With respect to the 2008 extension of the SME Scheme to cover "larger" SMEs, Δ_R measures the increase in R&D induced by more generous tax relief under the SME Scheme by a firm benefitting from the scheme thanks to the new thresholds. That is, $\Delta_R = R_{new} - R_{old}$ where R_{new} and R_{old} are the firm's R&D's under the new and old policies respectively. Similarly, $\Delta_{EC} = EC_{new} - EC_{old}$ where EC_{new} and EC_{old} are the firm's corresponding Exchequer costs due to the policy change.

Rearranging the R&D tax-price elasticity formula gives:

$$\eta = \frac{\frac{R_{new}-R_{old}}{(R_{new}+R_{old})/2}}{\frac{\rho_{new}-\rho_{old}}{(\rho_{new}+\rho_{old})/2}} = \frac{\frac{\Delta_R}{\bar{R}}}{\frac{\Delta_\rho}{\bar{\rho}}} \equiv \frac{\Delta_R}{\bar{R}} = \eta \times \frac{\Delta_\rho}{\bar{\rho}}$$

where ρ is the tax-adjusted user cost of R&D, $\Delta_A \equiv A_{new} - A_{old}$, and $\bar{A} \equiv (A_{new} + A_{old})/2$. For simplicity, we consider the tax deduction case and the SME payable tax credit case separately.

SME tax deduction case

In this case,

$$\begin{aligned} \rho^{deduction} &= \frac{(1-\tau(1+e))}{1-\tau}(r+\delta) \\ EC^{deduction} &= R \times e \times \tau \end{aligned}$$

where τ is the effective corporate tax rate, e is the enhancement rate, r is the real interest rate, and δ is the depreciation rate. As the above firm moves from being a large company pre-2008 to being an SME post-2008, its enhancement rate increases from 25% to 75%. At the same time, corporate tax rate decreases from 30% to 28%. Combining $e_{old} = 0.25$, $e_{new} = 0.75$, $\tau_{old} = 0.30$, $\tau_{new} = 0.28$ with estimated R&D tax-price elasticity of $\eta = -4.0$ gives $\frac{\Delta_\rho}{\bar{\rho}} = -0.23$ and $\frac{\Delta_R}{\bar{R}} = 0.92$, which implies $\frac{R_{new}}{R_{old}} = 2.70$.

On the cost side, we have:

$$
\begin{aligned}
EC_{old} =& R_{old} + e_{old} + \tau_{old} = R_{old} \times 0.075 \\
EC_{new} =& R_{new} + e_{new} + \tau_{new} = R_{new} \times 0.21
\end{aligned}
$$

Putting all the elements together gives

$$
\mu^{deduction} \equiv \frac{\Delta_R}{\Delta_{EC}} = \frac{R_{new} - R_{old}}{EC_{new} - EC_{old}} = \frac{(R_{old} \times 2.70) - R_{old}}{(R_{old} \times 2.70 \times 0.21) - (R_{old} \times 0.075)} = \frac{1.70}{0.49} = 3.46
$$

so the value for money ratio in the SME tax deduction case is 3.46. In other words, £1 of taxpayer money generates £3.46 in additional R&D.

Finally, note that Δ_{EC} could be rewritten as:

$$
\Delta_{EC} = EC_{new} - EC_{old} = R_{new} \times 0.21 - R_{old} \times 0.075 = \Delta_R \times 0.21 + R_{old} \times (0.21 - 0.075)
$$

where the first element represents the Exchequer costs associated with new R&D and the second term reflects additional Exchequer costs paid on existing R&D due to more generous tax relief. In this case, the majority of the additional costs are because of the new R&D generated, i.e., $\Delta_R \times 0.21 = R_{old} \times 0.36$ makes up close to 73% of $\Delta_{EC}(\Delta_{EC} = R_{old} \times 0.49)$.

SME payable tax credit case

In this case,

$$
\begin{aligned}
\rho^{credit} =& (1 - c(1+e))(r + \delta) \\
EC^{credit} =& R \times c \times (1 + e)
\end{aligned}
$$

where c — the payable credit rate -– is always zero for large companies and 14% for SMEs post-2008. Combining $c_{old} = 0$, $c_{new} = 0.14$, $e_{old} = 0.25$, $e_{new} = 0.75$, and $\eta = -4.0$ gives $\frac{\Delta_\rho}{\bar{\rho}} = -0.28$ and $\frac{\Delta_R}{\bar{R}} = 1.11$, which implies $\frac{R_{new}}{R_{old}} = 3.51$. On the cost side, $EC_{old} = 0$ and $EC_{new} = R_{new} \times c_{new} \times (1 + e_{new}) = R_{new} \times 0.25$. Putting all the elements together gives:

$$
\mu^{payable} \equiv \frac{\Delta_R}{\Delta_{EC}} = \frac{R_{new} - R_{old}}{EC_{new} - EC_{old}} = \frac{R_{old} \times 3.51 - R_{old}}{R_{old} \times 3.51 \times 0.25 - 0} = \frac{2.51}{0.86} = 2.92
$$

The value for money ratio in the payable tax credit case is 2.92. In this case, the amount of additional R&D's Exchequer costs due to newly-generated R&D $\Delta_R \times 0.25 = R_{old} \times 0.62$ constitutes close to 72% of $\Delta_{EC}(\Delta_{EC} = R_{old} \times 0.82)$.

R&D Tax Relief Scheme over 2006-11

To evaluate the overall R&D Tax Relief Scheme over 2006-11, we calculate:

$$\mu \equiv \frac{\Delta_R}{\Delta_{EC}} = \frac{R_{tax\ relief} - R_{no\ tax\ relief}}{EC_{tax\ relief} - EC_{no\ tax\ relief}} = \frac{R_{tax\ relief} - R_{no\ tax\ relief}}{EC}$$

separately for each of three sub-schemes, SME tax deduction scheme (Table B.17 Panel B), SME payable tax credit scheme (Panel C), and large company tax deduction scheme (Panel D), in each year, using the same approach as described in detail above. We generalize our estimated tax-price elasticity of 4.0 to the whole population of SMEs, but use a lower-bound tax-price elasticity of 1.1 for the population of large companies as these firms are less likely to be credit constrained and therefore less responsive to tax incentives. In addition, we use the small profits rate (19%-21%) instead of the regular corporate tax rate (26%-30%) for the population of SMEs as most of them are much smaller than the "larger" SMEs in our baseline sample and therefore most likely qualify for the small profits rate.

As reported in Table B.17, the SME tax deduction's value for money ratio decreases from 4.2 in 2006 to 3.6 in 2011 as SME tax deduction becomes significantly more generous over time. On the other hand, SME payable tax credits and large company tax deduction's value for money ratios are stable at around 2.9 and 1.5 respectively as these schemes do not change much over this period. The fact that all the value for money ratios are well above unity indicates that the R&D Tax Relief Scheme is effective in inducing additional R&D at relatively low cost to the Exchequer. Finally, we estimate the amount of additional R&D induced by the R&D Tax Relief Scheme as $\Delta_R = \mu \times EC$ using the calculated value for money ratios μ's and Exchequer costs national statistics (HMRC 2015). We do this for each of the three schemes in each year in Panels B, C and D, and then aggregate them together in Panel E.

To give an example, consider the SME tax deduction scheme in Panel B for 2009. The tax-adjusted user cost of R&D under this sub-scheme in 2009, calculated using the policy parameters, is $\frac{1-0.21\times(1+0.75)}{1-0.21}(0.05+0.15) = 0.16$ The counterfactual user cost in world without R&D tax relief is $0.05 + 0.15 = 0.20$ $(e = 0)$. The percentage difference between these user costs is then $\frac{\Delta_\rho}{\bar{\rho}} = \frac{0.16-0.20}{(0.16+0.20)/2} = -0.22$. The tax-price elasticity of R&D of SMEs as estimated in sub-section 7.2 is $\eta^{SME} = -4.0$.

The elasticity formula and Exchequer cost formulae give:

$$\begin{aligned}
\eta^{SME} &= \frac{\frac{\Delta_R}{\bar{R}}}{\frac{\Delta_\rho}{\bar{\rho}}} = \bar{R} \times \eta^{SME} \times \frac{\Delta_\rho}{\bar{\rho}} \\
\Delta_{EC} &= EC_{tax\ relief} - 0 = R_{tax\ relief} \times e \times \tau \\
&= (\bar{R} + \frac{\Delta_R}{2} \times e \times \tau = \bar{R} \times (1 + 0.5 + \frac{\Delta_R}{\bar{R}}) \times e \times \tau \\
\Rightarrow \mu^{SME\ deduction} &= \frac{\Delta_R}{\Delta_{EC}} = \frac{\eta^{SME} \times \frac{\Delta_\rho}{\bar{\rho}}}{(1 + 0.5 \times \frac{\Delta_R}{\bar{R}}) \times e \times \tau} \\
&= \frac{4.0 \times 0.22}{(1 + 0.5 \times 4.0 \times 0.22) \times 0.75 \times 0.21} = 3.89
\end{aligned}$$

We report this value for money ratio in the second row of Table B.17 Panel B. From HMRC data we know that £130m was paid out in the SME deduction in this year. Hence, we can calculate that the total amount of additional R&D induced $\Delta_R = \mu^{SME\ deduction} \times EC$=3.89×130=506 (£m), as shown in fourth row of Panel B.

As discussed in sub-section 2.7.3, our aggregate estimates in Panel E suggest that the overall impact of the R&D Tax Relief Scheme is large. Over 2006-11, the policy, which costs less than £6 billion in lost tax revenue, induced close to £12 billion in additional R&D. On an annualized basis, spending £0.96 billion produced £1.98 billion of additional R&D.

These calculations show our estimates of what the counterfactual path of R&D would have been in the absence of the R&D Tax Relief Scheme. The bottom row of Table B.17 gives the yearly breakdown. For example, the final column shows that on average 2006-11 we estimate that R&D would be a full 20% lower in the absence of the tax scheme.

Counterfactual R&D without the Tax Relief Scheme 2000-11

It is important to note that throughout our analysis we have been focusing on qualifying R&D, i.e., that part of business R&D that is eligible for tax relief. Aggregate qualifying R&D is lower than the figures for Business Enterprise R&D (BERD) reported in Figure 2.4. For example, in 2011 aggregate BERD was £17bn and aggregate qualifying R&D was £12bn. There are various reasons for this difference, including the fact that BERD includes R&D spending on capital investment whereas qualified R&D does not (only current expenses are liable). It is also the case that HMRC defines R&D more narrowly for tax purposes that BERD which is based on the Frascati definition.

We present counterfactual BERD to GDP ratios in Figure 2.4. To calculate the counterfactual (the dotted line "UK without tax relief" in Figure 2.4) we simply deduct the additional qualified R&D that we estimate were created by the R&D tax relief system (second row of Table B.17 Panel E) from the aggregate BERD numbers from OECD MSTI Dataset (`https://stats.oecd.org/Index.aspx?DataSetCode=MSTI_PUB`). Since BERD is greater than qualifying R&D, the 20% fall in qualifying R&D translates into a 13% fall in BERD.

B.2 Data

B.2.1 CT600 dataset

The CT600 dataset is constructed by the UK tax authority (HMRC) and is a confidential panel dataset of corporate tax returns or assessments made from the returns for the universe of companies that file a corporate tax return in the UK. We can only access the dataset from within an HMRC facility (similar to a US Census Bureau Research Data Center) and merging with other datasets requires approval from HMRC. It is currently not possible to merge CT600 with other government secured datasets available at different facilities.[14] The CT600 dataset covers all accounting periods whose end dates fall between April 1st, 2001 and March 31st, 2012 (we denote the fiscal year ending in March 31st, 2012 by "2011" as most of the data will fall in this calendar year) and consists of all information on the UK Company Tax Return form (which is called the CT600 form). Specifically, an extension of CT600, the Research and Development Tax Credits (RDTC) dataset, provides detailed information on tax relief claims. However, CT600 contains little information on financial statement variables (e.g., assets and employment are not included) as they are not directly required on corporate tax forms. [15]

We convert the original observation unit of firm by accounting period in CT600 to firm by financial year by aggregating all accounting periods the end dates of which fall in the same financial year.[16] This conversion affects a very small number of observations as only 3% of our firm by year observations are aggregates of multiple accounting periods. Our converted dataset then contains 15.7 million firm by year observations over 12 financial years from 2000 to 2011 (covering 3.2 million firms), including 9.1 million firm by year observations over our study period from 2006 to 2011 (covering 2.5 million firms). Our key variables of interest are those related to firms' R&D tax relief claims from CT600's RDTC dataset, which include the amount of qualifying R&D expenditure each firm has in each year and the scheme under which it makes the claim (SME vs. Large Company Scheme). These variables, originally self-reported by firms on their CT600 forms, have been further validated and corrected by HMRC staff using additional tax processing data available only within the tax authority. It should also be

[14] For example, it is currently not possible to merge CT600 with the BERD firm survey which is used to build the national estimate of R&D. Since BERD is a stratified random sample which puts large weight on the biggest R&D performers, we would likely only have a small overlap with firms around the threshold.

[15] The CT600 dataset was further extended to cover up to the end of financial year 2014 in late 2017. However, the corresponding RDTC dataset has not been made available as of the writing of this paper. As a result, we focus on the period between 2009 and 2011, for which we have reliable R&D data, as our post-policy period for R&D analyses. In addition, it is unlikely that our key running variable – total assets in 2007 – has strong predictive power of firm's SME status after 2011. We do use data on sales up to 2013 from this extended CT600 dataset in our firm performance analysis (see Table B.13).

[16] Financial year t begins on April 1st of year t and ends on March 31st of year $t + 1$.

noted that R&D tax relief variables are only available for R&D-tax-relief-claiming firms for the years in which they make the claims. While we believe it is reasonable to assume that non-claiming firms have zero qualifying R&D expenditure, it is not possible to construct their precise SME eligibility without full information on employment, assets (balance sheet total), sales, and ownership structure.

Table B.20 shows that over our study period of 2006-11, we observe claims in 53,491 firm by year observations (by 20,730 firms), 81% of which are under the SME Scheme. The total qualifying R&D expenditure and estimated Exchequer costs under the SME Scheme are in nominal terms £11.2bn and £1.8bn respectively; the corresponding figures under the Large Company Scheme are £48.5bn and £3.9bn (excluding claims by SME subcontractors). These figures are in line with the official R&D Tax Relief Scheme statistics released in HMRC (2014).

We also use the data on sales and on investment in plant and machinery from CT600. Sales are annualized to account for different accounting period lengths. CT600 tax-accounting sales, which is calculated using the cash-based method, is not the same as financial-accounting sales (reported in the FAME data – see below), which is calculated using the accrual method and used to determine SME eligibility.[17] However, CT600 sales provides a good measure for firms' growth and performance, given its wide coverage.

B.2.2 FAME dataset

FAME is a database of UK companies provided by Bureau Van Dijk (BVD), a private sector company. The panel dataset contains companies' balance sheet and income statement data from companies' annual accounts filed at the UK company registry (Companies House), together with additional information on addresses and industry codes. Like other countries, UK regulations for reporting accounting variables vary with company size, so some balance sheet and income statement variables are missing – we discuss the implications of this below.[18]

Our FAME dataset also covers 14 financial years from 2000 to 2013 and contains 23.9 million firm by year observations (covering 4.4 million firms), including 11.5 million firm by year observations over our study period from 2006-11 (covering 3.1 million firms). Our key SME-eligibility variable from FAME (for R&D tax relief purpose) is total assets (i.e., balance sheet total). As almost all UK companies are required by the Companies House to send in their balance sheets for their annual accounts regardless of their size, total assets coverage in FAME is close to complete, at 97% over our study

[17]The cash-based method focuses on actual cash receipts rather than their related sales transactions. The accrual methods records sale revenues when they are earned, regardless of whether cash from sales has been collected.

[18]All UK limited companies, public limited companies (PLC), and limited liability partnerships (LLP) are required to file annual accounts with the Companies House. An annual account should generally include a balance sheet, an income statement, a director's report, and an audit report. However, smaller companies may be exempt from sending in income statement, director's report, or audit report. All UK registered companies are required to file annual returns with the Companies House, which contain information on registered address and industry codes.

period of 2006-11. On the other hand, sales (financial-accounting sales used to determine SME eligibility) is reported by only 15%, as smaller firms are not required to provide their income statements.[19] The proportion of firms who report employment is even lower at 5%, as employment reporting is not mandatory. Even in our baseline sample of relatively larger firms (i.e., firms with total assets in 2007 between €61m and €111m), the proportion of firms who report sales is 67% and the proportion who report employment is 55%. For this reason, while we do use FAME sales and employment as running variables in some alternative specifications, our baseline sample and key results are derived using total assets as the running variable.

Besides total assets, sales, and employment, other FAME variables used in our paper include primary industry code (UK 4-digit SIC), address, and fixed assets as a proxy for capital stock.

B.2.3 PATSTAT dataset

Our patent data are drawn from the World Patent Statistical Database (PATSTAT) maintained by the European Patent Office (EPO).[20] PATSTAT is the largest international patent database available to the research community and includes nearly 70 million patent documents from over 60 patent offices, including all of the major offices such as the United States Patent and Trademark office (USPTO), the Japan patent office (JPO) and the Chinese Patent and Trademark Office (SIPO) in addition to the EPO. PATSTAT data cover close to the population of all worldwide patents between 1900-2015.

PATSTAT reports the name and address of patent applicants, which allows matching individual patents with company databases. The matching between PATSTAT and FAME is implemented by Bureau Van Dijk and is available as part of the ORBIS online platform through a commercial agreement. The quality of the matching is excellent: over our sample period, 94% of patents filed in the UK and 96% of patents filed at the EPO have been matched with their owning company.

A patent in country i grants a holder an exclusive right to commercially exploit the invention in that country. Accordingly, she will patent her invention in country i if she plans to either market there directly or license to another firm who will sell it there. The set of patents in different countries related to the same invention is called a patent family. The vast majority of patent families include only one patent (usually in the home country of the inventor). Importantly, PATSTAT reports not only the unique identifier of each patent application, it also indicates a unique patent family indicator for each patent (we use the DOCDB patent family indicator). This allows us to identify all patent applications filed worldwide by UK-based companies and to avoid double-counting inventions that are protected in several countries.

[19]Small companies (those having any 2 of the following: (1) sales of £6.5m or less, (2) assets of £3.26m or less, (3) 50 employees or less) are only required to send in balance sheets. Micro-entities (those having any 2 of the following: (1) sales of £632,000 or less, (2) assets of £316,000 or less, (3) 10 employees or less) are only required to send in simplified balance sheets.

[20]For further details see `http://www.epo.org/searching/subscription/raw/product-14-24.html`.

In this study, our primary measure of innovation is the number of patent families – irrespective of where the patents are filed. This proxies for the number of inventions a firm makes. This means that we count the number of patents filed anywhere in the world by firms in our sample, be it at the UK Intellectual Property Office, at the European Patent Office, at the USPTO or anywhere else, but we use information on patent families to make sure that any invention patented in several places is only counted once. Patents are sorted by the first year they were filed (the priority year). We use fractional counts to account for multiple applicants. For example, if two firms jointly apply for a patent, then each firm *i*s attributed one half of a patent. In practice, only 8% of patents filed by UK-based companies are filed jointly by at least two companies.

There are many well-known issues with patents as a measure of innovation. As noted above, not all inventions are patented, although it is reasonable to assume the most valuables ones are, so counting patents screens out many of the low value inventions. Nevertheless, since patents are of very heterogeneous importance we use several approaches to examine how our results change when looking at patent quality. First, we distinguish between patents filed at the UK patents office and patents files at the EPO and USPTO.[21] Since the financial and administrative cost is about six times higher at the EPO than UK patent office, EPO and USPTO patents will, on average be of higher private value. A second measure of patent quality is the size of patent families, the number of jurisdictions in which each patent is filed. There is evidence that the number of jurisdictions in which a patent is filed is an indicator of its economic value as patenting is costly (see Guellec and Van Pottelsberghe de la Potterie, 2004, and Harhoff et al., 2003). A third measure of quality is to distinguish by technology class, as some classes (e.g., pharmaceuticals) are likely to be more valuable than others (e.g., business process methods). Fourth, we know whether the patent filed was subsequently granted, with the reasonable presumption that granted patents are of higher value. Fifth, we use patent citations, also available from PATSTAT. For each patent in the database, we know how many times it was cited by subsequent patents (excluding self-citations). We use the number of subsequent citations (referred to as forward citations) as a measure of value. Again, this measure is well rooted in the patent literature (Hall et al., 2005, Lanjouw et al., 1998). The disadvantage for our purposes is that we only have a short finite window of time for future citations causing a truncation problem.

In PATSTAT, patents are categorized based on the International Patent Classification (IPC). We use IPC codes at three-digit level to construct measures of the technological distance between firms used to investigate spillover effects.

[21]Note that because of differences in the "technological scope" of patents across patent offices, two patents filed in the UK may be "merged" into a single patent filed at the EPO. In this case, these three patents will constitute a single patent family and the number of patent families is smaller than the number of UK patents. This configuration happens very rarely, however.

B.2.4 Sample construction: merging datasets

CT600 was merged with FAME using an HMRC-anonymized version of company registration number (CRN), which is a unique regulatory identifier in both datasets. 95% of CT600 firms between 2006 and 2011 also appear in FAME, covering close to 100% of R&D performing firms and 100% percent of patenting firms in this period.[22] Unmatched firms are slightly smaller but not statistically different from matched ones across different variables reported in CT600, including sales, gross trading profits, and gross and net corporate tax chargeable.[23] Furthermore, that the match rate is less than 100% is due to CRN entering error in FAME, which happens more often among firms that are much smaller than those around SME-eligibility thresholds.[24] For these reasons, we believe sample selection due to incomplete matching between CT600 and FAME is unlikely to be an issue for us.[25]

PATSTAT has been merged with FAME by BVD. As PATSAT comprehensively covers all UK patenting firms, we can safely infer that non-matched firms have zero patents. Over our study period of 2006-11, 9,420 out of 2.5 million CT600 firms claim a total of 46,405 patent families (in 17,293 firm by year observations), including 23,617 higher-quality EPO patents. These patents cover 90% of the total recorded in PATSTAT.

From the merged master dataset, we construct our baseline sample based on total assets in 2007, as it is our key running variable. Specifically, our baseline sample includes 5,888 firms that satisfy the two following conditions: (1) the firm's total assets in 2007 is between €61m and €111m (within €25m below and above the SME threshold of €86m), and (2) the firm appears in CT600 in 2008 (to exclude firms exiting before the policy change in 2008). Baseline sample descriptive statistics are summarized in Table 2.1 and discussed in detail in sub-section 2.4.2.

B.2.5 Variable construction

As FAME total assets and sales are reported in sterling while the corresponding SME ceilings are set in euros, we convert sterling to euros using the exact same rule used by HMRC for tax purposes. That is, the conversion should be done using the exchange rate on the last day of the relevant accounting period or the average daily exchange rate throughout that accounting period, whichever is more beneficial for the enterprise. The

[22]Out of 2,495,944 firms present in CT600 between 2006 and 2011, 2,358,948 firms are matched to FAME (94.5% match rate). Over the same period, 20,627 out of 20,730 R&D-performing firms and 9,376 out of 9,420 patenting firms are matched to FAME (99.5% match rate).

[23]Differences (standard errors) between matched and unmatched firms in sales (£'000), gross trading profits (£), gross corporate tax chargeable (£) and net corporate tax chargeable (£) are 970 (3,286), 8,969 (13,703), 3,497 (3,898) and 1,961 (2,291) respectively. None of these differences are statistically significant at conventional level.

[24]Because of confidentiality concerns, we do not get to work directly with CRNs but an anonymized version of CRNs provided by the HMRC Datalab for both FAME and CT600 datasets. This prevents us from further cleaning and matching of initially unmatched firms due to above issue.

[25]The correlation between ln(sales) from CT600 and ln(sales) from FAME is 0.90. As noted above, the variables are not measured in the same way, but the fact that their correlation is high is reassuring that the match is well performed.

daily exchange rate is obtained from the OECD, exactly the same method as used by HMRC.

For qualifying R&D expenditure, we do not include the amounts claimed by SME subcontractors, which do not benefit from more generous reliefs under the SME Scheme. Since SME subcontracting makes up only a small portion of the overall R&D Tax Relief Scheme, we confirmed excluding SME subcontracting does not materially affect our key findings. To account for price differences across years, we also convert nominal values of R&D expenditure to their real values in 2007 price, using UK annual CPI as reported in the World Bank Economic Indicators database.[26]

We address the presence of outliers in R&D spending or patenting by winsorizing our key outcome variables, which include qualifying R&D expenditure and number of all patents as well as number of EPO patents, UK patents, and US patents. Specifically, for each variable, the top 2.5% of non-zero values in each year within the sample of firms with 2007 total assets between €46m and €126m are set to the corresponding 97.5 percentile value (i.e., winsorization at 2.5% of non-zero values). This translates into "winsorizing" the R&D of top 5 to 6 R&D spenders and the number of patents of top 2 to 4 patenters in the baseline sample in each year. It should be noted that our key findings are robust to alternative choices of winsorization window (e.g., 1% or 5% instead of 2.5%), or to excluding outliers instead of winsorizing outcome variables (see Tables A3-A5). Construction of other variables is detailed in the notes to tables.

B.2.6 Running variable selection: SME criterion binding ratio

We chose total assets as the key running variable as it is the only SME criterion with close to complete coverage in FAME. In addition, as discussed in sub-section 2.7.6, we also find evidence that the asset criterion is more binding than the sales one. A firm is considered an SME if it meets either one of the criteria, thus the asset criterion is binding only when the firm already fails the sales one and vice versa.

We calculate the binding/non-binding ratio for the asset criterion as the number of firms with 2007 sales in [€100m, €180m] (i.e., firms for which the asset criterion binds), divided by the number of firms with 2007 sales in [€20m, €100m] (i.e, firms which also meet the sales criterion), conditioned on firms' 2007 total assets being in [€36m, €136m] (i.e., +/-€50m window around the asset threshold of €86m). Similarly, the same ratio for the sales criterion is the number of firms with 2007 assets in [€86m, €166m] (i.e., firms for which the sales criterion binds), divided by the number of firms with 2007 assets in [€6m, €86m] (i.e., firms which already meet the asset criterion), conditioned on firms' 2007 sales being in [€50m, €150m] (i.e., +/-€50m window around the sales threshold of €100m). The binding/non-binding ratio for the asset criterion is 0.36, considerably higher than the same ratio for the sales criterion of 0.20, as presented in Figure B.8.

[26]Ratios of current-£ to 2007-£ derived using UK annual CPI are 1.023 for 2006, 1.000 for 2007, 0.965 for 2008, 0.945 for 2009, 0.915 for 2010, and 0.875 for 2011.

This implies that the *below-asset-threshold indicator* is a more precise instrument for firm's SME status than the *below-sales-threshold indicator*, consistent with the results reported in Table B.14 Panel B. Finally, the qualitative results that the asset criterion is more binding than the sales criterion does not change when we pick different windows to calculate the binding/non-binding ratios.

B.3 R&D technology spillovers

B.3.1 Semi-parametric estimation of spillovers

We modify the spillover regression in equation (2.4) from section 2.6 to model the potentially heterogeneous effect of baseline firm i's likely-eligibility for the SME scheme on connected firm j's average patents over 2009-13 as a non-parametric function of the primary technology class size (measured in percentile and denoted as x):

$$PAT_{j,09-13} = \alpha_4(x) + \theta(x)E_{i,2007} + f_4(z_{i,2007}, x) + g_4(z_{k,2007}, x) + \epsilon_{4ij}$$

Figure 2.5 plots the estimated function $\theta(x)$ of the spillover effect based on primary technology class size percentile. It is estimated from semi-parametric local linear regressions of equation (2.4) at each value of x, weighted by a Gaussian kernel with a bandwidth of 20% (with x ranging from 1 to 100). The observed pattern is similar across a wide range of bandwidths.

B.3.2 Alternative approach to estimating R&D technology spillovers

In this sub-section, we discuss a complementary approach to estimating R&D technology spillovers using a monadic specification instead of the dyadic specification discussed in Section 2.6. Following the work of Jaffe (1986) we calculate the knowledge spillover pool available to firm j as $spilltechR_{j,09-11} = \sum_{i,i \neq j} \omega_{ij} R_{i,09-11}$ where $R_{i,09-11}$ is the average R&D of firm i over 2009-11 and ω_{ij} is measure of technological "proximity" between firms i and j, computed based on the distribution of technology classes in which the firms patent (e.g., if two firms have identical patent class distributions then proximity is 1; if they patent in entirely different patent classes then proximity is zero).[27] We follow our earlier approach of using $(E_{i,2007}$ as instrument for $R_{i,09-11}$ $((E_{i,2007}$ is firm i's below-asset-threshold indicator in 2007).[28] Consequently, we construct $spilltechR_{j,09-11} = \sum_{i,i \neq j} \omega_{ij} E_{i,2007}$ as instrument for $spilltechR_{j,09-11}$. The exclusion restriction requires that

[27]Following Jaffe (1986) we define proximity as the uncentered angular correlation between the vectors of the proportion of patents taken out in each technology class $\omega_{ij} = \frac{F_i F_j'}{(F_i F_i')^{\frac{1}{2}} (F_j F_j')^{\frac{1}{2}}} F_i = (F_{i1}, ..., F_{iY})$ is a 1 $\times$ Y vector where $F_{i\tau} = \frac{n_{i\tau}}{n_i}$ is firm i's number of patents in technology field τ as a share of firm i's total number of patents. To calculate $F_{i\tau}$, we use information on all patents filed between 1900 and 2011 and their 3-digit International Patent Classification (IPC), which classifies patents into 123 different technology fields. These data are available from PATSTAT. Bloom et al. (2013) show that the Jaffe measure delivers similar results to more sophisticated measures of proximity.

[28]More generally, $(E_{i,2007} = I\{z_{i,2007} \leq \tilde{z}\}$ is a binary indicator equal to one if the 2007 financial variable $z_{i,2007}$ is equal to or less than the corresponding new SME threshold for it, $\tilde{z}$.

the discontinuity-induced random fluctuations in firm i's eligibility would only affect technologically-connected firm j's R&D and innovation through R&D spillovers.

Our monadic spillover IV regression estimates the impact of the aggregate R&D spillover pool available to firm j, $spilltechR_{j,09-11}$, on firm j's innovation, $PAT_{j,09-13}$, controlling for firm j's own R&D using $E_{i,2007}$ as an instrument:

$$PAT_{j,09-13} = \alpha + \psi spilltechR_{j,09-11} + F_j(Z_{2007}) + \zeta E_{E,2007} + g(z_{j,2007}) + \pi techconnect_j + \epsilon_j$$

where $F_j(Z_{2007}) = \sum_{i,i\neq j} \omega_{ij} f(z_{i,2007})$ and Z_{2007} is a vector comprising of the 2007 assets for all firms; $f(z_{i,2007})$ and $g(z_{j,2007})$ are polynomials of firms i and j's 2007 total assets; and $techconnect_j = \sum_{i,i\neq j} \omega_{ij}$.[29] We instrument $spilltechR_{j,09-11}$ with $spilltechE_{j,2007}$. $F_j(Z_{2007})$ and $g(z_{j,2007})$ are polynomial controls for $spilltechE_{j,2007}$ and $E_{j,2007}$ respectively while $techconnect_j$ additionally controls for spillover-receiving firm j's level of "connectivity" in technology space. We estimate the above equation on the sample of firm j's with total assets in 2007 between €51m and €121m. This is a larger bandwidth than in the baseline sample as the policy-induced R&D can have spillovers on firms well beyond the policy threshold.[30] Standard errors are bootstrapped using 1,000 replications over firms.

Column (1) of Table B.18 reports the first stage for the R&D spillover term and column (2) the first stage for spillover-receiving firm j's own R&D. As expected the instrument $spilltechE_{j,2007}$ significantly predicts $spilltechR_{j,09-11}$ (column (1)) and the instrument $E_{j,2007}$ significantly predicts connected firm j's own R&D (column (2)). The instruments $spilltechE_{j,2007}$ and $E_{j,2007}$ are jointly statistically different from zero in both columns, with F-statistics of 26.9 and 6.4 respective. Interestingly, we see that in the reduced form patent model of column (3) the R&D spillover instrument, $spilltechE_{j,2007}$, has a large and significant positive effect on firm j's patents. This is consistent with

[29] Given equation (2.1) for firm i's R&D as $R_{j,09-11} = \alpha + \beta^R E_{i,2007} + f(z_{i,2007}) + \epsilon_i$, aggregating across all firm i's around the SME asset threshold and using ω_{ij} as weights gives:

$$\sum_{i,i\neq j} \omega_{ij} R_{i,09-11} = \alpha \sum_{i,i\neq j} \omega_{ij} + \beta^R \sum_{i,i\neq j} \omega_{ij} E_{i,2007} + \sum_{i,i\neq j} \omega_{ij} f(z_{i,2007}) + \sum_{i,i\neq j} \omega_{ij}\epsilon_i$$
$$\Rightarrow spilltechR_{j,09-11} = \alpha techconnect_j + \beta^R spilltechE_{j,2007} + F_j(Z_{2007}) + \nu_j$$

This equation shows that $F_j(Z_{2007})$ is the appropriate polynomial control when using $spilltechE_{j,2007}$ as instrument for $spilltechR_{j,2007}$. The key condition that $\nu_j = \sum_{i,i\neq j} \omega_{ij}\epsilon_i$ is mean independent of $spilltechE_{j,2007}$ conditional on $F_j(Z_{2007})$ follows from RD Design results. To address non-trivial serial correlation among the error term ν_j we correct the standard errors using 1,000 bootstrap replications over firms.

[30] Note that $spilltechR_{j,09-11}$ is calculated using the population of all possible firm i's, while $spilltechE_{j,2007}$ and $F_j(Z_{2007})$ are calculated using all firm i's with 2007 total assets between €51m and €121m (same as the sample on which we nomadic spillover equation), as the RD Design works best in samples of firms around the relevant threshold. Our key results are robust to using different sample bandwidths around the threshold to calculate $spilltechE_{j,2007}$ and/or to estimate the monadic spillover equation. In addition, in all reported results, we use second order polynomial controls separately on each side of the threshold for $f(z_{j,2007})$ and $g(z_{j,2007})$). In this larger sample we found that higher order terms were significant. However, using different orders of polynomial controls does not change our qualitative findings.

the hypothesis that policy-induced R&D has sizeable spillover effect on technologically-connected firms' innovation.

Turning to the IV results, column (4) suggests that there is no significant R&D spillover effect on technologically-connected firms' R&D, as already suggested by the R&D regression in column (2). By contrast, columns (5) and (6) report that the aggregate R&D spillover pool available to firm j, $spilltechR_{j,09-11}$ does have a causal impact on firm j's patenting, consistent with the patent regression in column (3). This spillover effect is robust after controlling for the policy's direct effect on firm j's R&D, either by (i) including $E_{j,2007}$ as a control in addition to the instrumented spillover term (column (5)), or (ii) including $R_{j,09-11}$ as a control and using $E_{j,2007}$ as the corresponding instrument (column (6)). The latter is a very demanding specification, and even though the corresponding spillover coefficient is no longer significant,[31] its magnitude is almost identical in both specifications.

In terms of magnitudes, the last two columns suggest that a £1m increase in R&D by a firm i with an identical technological profile will increase firm j's patenting by 0.014, which is 3.4% of the direct effect of an equivalent R&D increase by the firm *i*tself (=0.014/0.412). Combining this with the mean level of connectivity among firms in the sample gives us the total spillover effect of 0.616 (= 0.014 x 44). In other words, the total spillovers of an £1m increase in R&D on all technology-connected firms' patenting is about 1.5 x (= 0.616/0.412) the direct effect on own patenting.[32]

This presence of positive R&D spillovers on innovations is robust to a wide range of robustness tests. The reduced-form spillover coefficient capturing effect of $spilltechE_{j,2007}$ on firm j's patents (column (3)'s specification) is robust to (i) limiting firm j sample to only patenting firms, (ii) using EPO, UK, and US patent outcomes, (iii) employing the more sophisticated Mahalanobis generalization of the Jaffe proximity measure to allow for between field overlap (see Bloom et al., 2013), (iv) reconstructing the standard Jaffe measure of technological proximity using only information on patents filed up to 2008, and (v) using smaller or large sample to calculate the instrument $spilltechE_{j,2007}$ or to estimate the monadic spillover equation.

Besides spillovers in technology space, there may be some negative R&D spillovers through business stealing effects among firms in similar product markets. To address this concern, we follow Bloom et al. (2013) and construct $spillsicR_{j,09-11} = \sum_{i,i\neq j} \phi_{ij} R_{i,09-11}$ that captures the aggregate R&D spillovers pool in product market

[31]If we use robust standard errors instead of bootstrapped standard errors, the estimated coefficient (standard error) for $spilltechR_{j,09-11}$ from column (6)'s specification is 0.014 (0.007), statistically significant at 5% level.

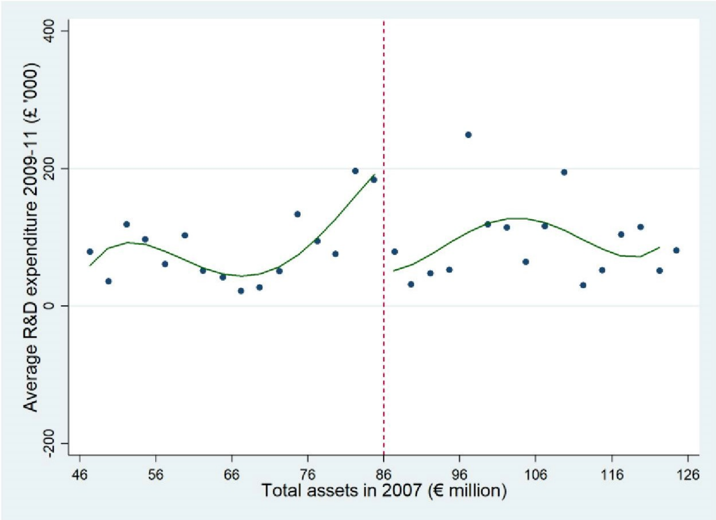
[32]Consider a firm i that increases its R&D by £1m. The spillover of this R&D increase on a firm j's patenting, as estimated by the monadic spillover equation, is $\psi\omega_{ij}$. Summing this spillover over all spillover-receiving firms j' patenting gives total spillovers of $\psi \sum_{i,i\neq j} \omega_{ij} = \psi techconnect_i$, which is the product of the spillover coefficient and firm i's level of connectivity. The estimated total spillover effect for an average firm i is then $\hat{\psi}\overline{techconnect_i} = 0.014 \times 44 = 0.616$.

space, where ϕ_{ij} is a measure of product market distance between firms i and j.[33] We also construct $spillsicE_{j,2007} = \sum_{i,i \neq j} \phi_{ij} E_{i,2007}$ as instrument for $spillsicR_{j,09-11}$. We found no significant effects of $spillsicR_{j,09-11}$ on either firm j's R&D or firm j's patents.

In summary, these findings provide evidence that policy-induced R&D have sizable positive impacts on not only R&D performing firms but also other firms in similar technology areas, as measured by patents. This further supports the use of R&D subsidies in the UK context.

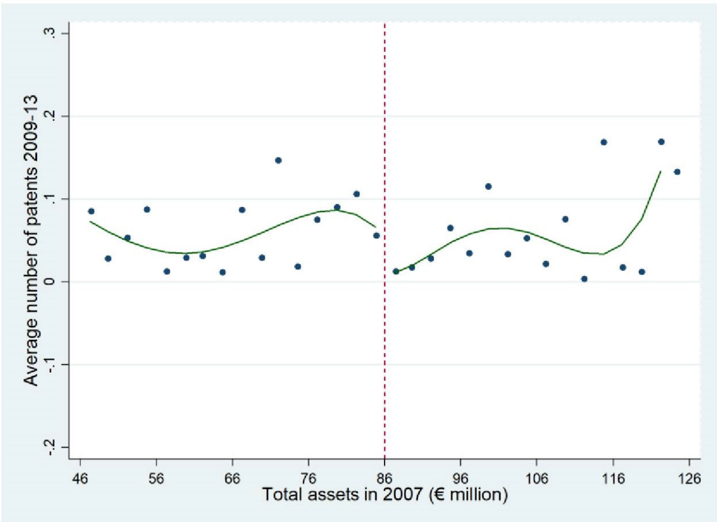
[33] $\phi_{ij} = 1$ if firm i operates in the same industry as firm j and $\phi_{ij} = 0$ otherwise. To calculate ϕ_{ij}, we use firms' primary industry codes at 3-digit Standard Industry Classification (SIC). These data are available from FAME.

Figure B.1: Discontinuity in average R&D expenditure over 2009-2011



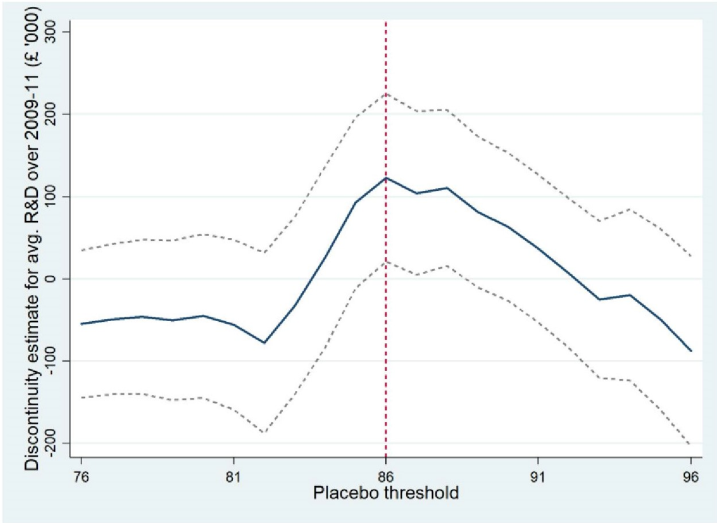
Note: The figure corresponds to CCT RD Design R&D regression. The dependent variable is average R&D expenditure over 2009-2011. The running variable is total assets in 2007 with a threshold of €86m. Controls for fourth order polynomials of the running variable separately on each side of the threshold are included. The CCT discontinuity estimate at the €86m threshold is 160.0 with a standard error of 64.2. Bin size for the scatter plot is €2.5m.

Figure B.2: Discontinuities in average number of patents over 2009-2013



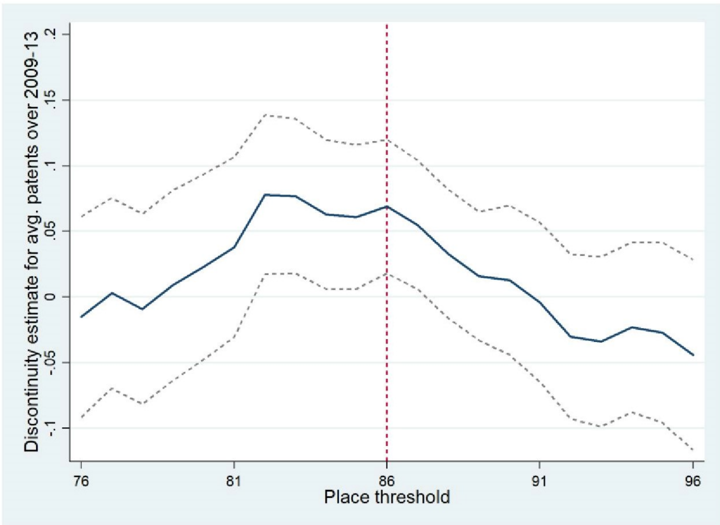
Note: The figure corresponds to CCT RD Design patent regression. The dependent variable is average number of patents over 2009-2013. The running variable is total assets in 2007 with a threshold of €86m. Controls for fourth order polynomials of the running variable separately on each side of the threshold are included. The CCT discontinuity estimate at the €86m threshold is 0.065 with a standard error of 0.023. Bin size for the scatter plot is €2.5m.

Figure B.3: Discontinuities in 2009-11 R&D at "pseudo" SME asset thresholds



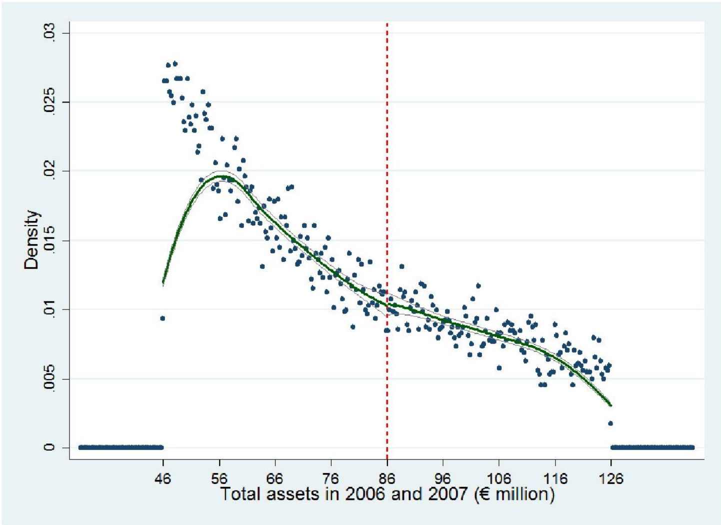
Note: This figure presents the R&D discontinuity estimates as a function of placebo threshold. The coefficient at each threshold is estimated using the baseline R&D regression based on equation (1) (OLS RD Design with average R&D expenditure over 2009-2011 as the dependent variable). The running variable is total assets in 2007. Baseline sample includes firms with total assets in 2007 €25m above and below the corresponding placebo threshold. Controls for first order polynomials of running variable separately for each side of the placebo threshold are included. The dashed lines indicate the 95% confidence interval for the discontinuity estimates.

Figure B.4: Discontinuities in 2009-13 patents at "pseudo" SME asset thresholds



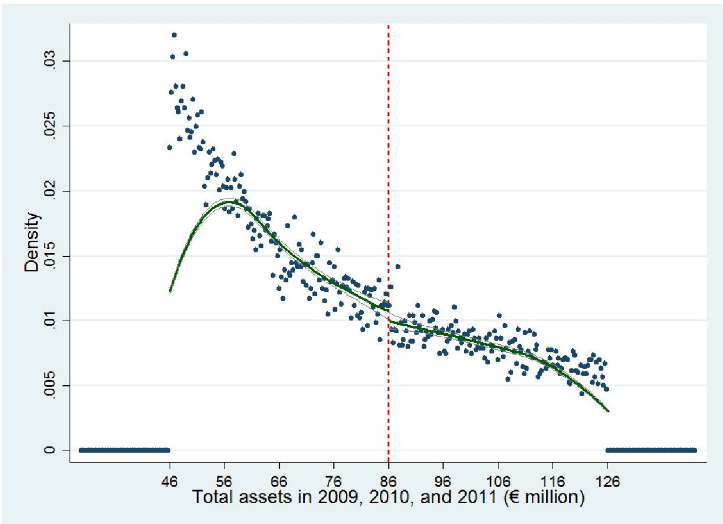
Note: This figure presents the patent discontinuity estimates as a function of placebo threshold. The coefficient at each threshold is estimated using the baseline patent regression based on equation (3) (OLS RD Design with average number of patents over 2009-2013 as the dependent variable). The running variable is total assets in 2007. Baseline sample includes firms with total assets in 2007 €25m above and below the corresponding placebo threshold. Controls for first order polynomials of running variable separately for each side of the placebo threshold are included. The dashed lines indicate the 95% confidence interval for the discontinuity estimates.

Figure B.5: McCrary test at the SME asset threshold before the policy change



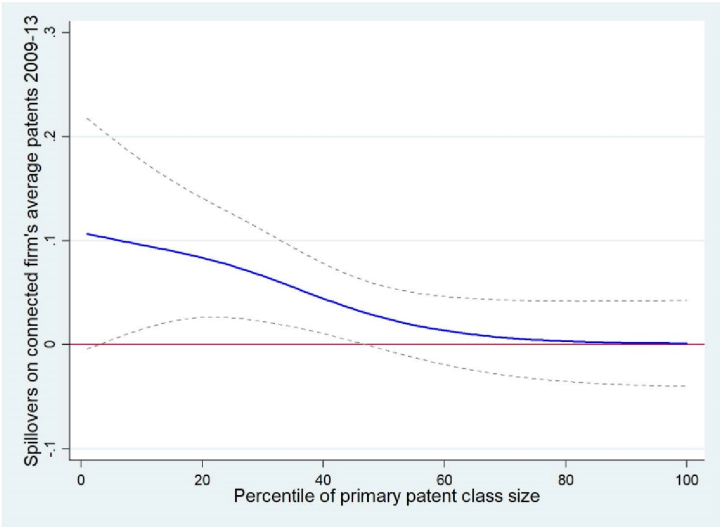
Note: McCrary test for discontinuity in distribution density of total assets at the SME asset threshold of €86m before the policy change, pooling together total assets in 2006 and 2007. Sample includes firms with total assets in [€46m, €126m] in each of the year. The discontinuity estimate (log difference in density height at the SME threshold) is 0.013, with standard error of 0.056.

Figure B.6: McCrary test at the SME asset threshold after the policy change



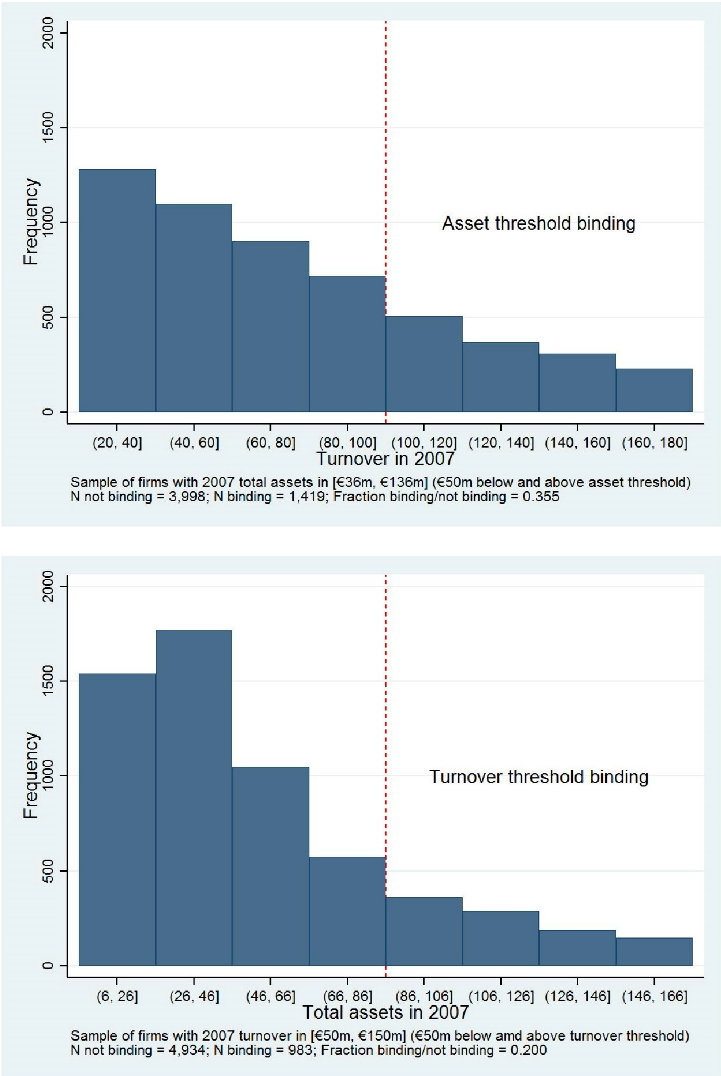
Note: McCrary test for discontinuity in distribution density of total assets at the SME asset threshold of €86m after the policy change, pooling together total assets in 2009, 2010, and 2011. Sample includes firms with total assets in [€46m, €126m] in each of the year. The discontinuity estimate (log difference in density height at the SME threshold) is -0.072, with standard error of 0.045.

Figure B.7: Spillovers on "loosely" connected firm's patents by patent class size



Note: This figure presents semi-parametric estimates of the spillover coefficient on "loosely"-connected firms' patents as a function of the technology class size percentile (the X-axis variable). Two firms are "loosely" connected technologically if they patent primarily in the same 3-digit technology class. The semiparametric estimation is based on equation (4), using a Gaussian kernel function of the X-axis variable and a bandwidth of 20% of the range (see Appendix C.1 for details). The 40th percentile of technology class size is 200. The dashed lines indicate the 90% confidence interval for the spillover coefficients.

Figure B.8: Number of firms with binding/non-binding asset and revenue thresholds



Note: Asset threshold is not binding for firms with 2007 sales in (€20m, €100m] and binding for firms with 2007 sales in (€100m, €180m]. Sales threshold is not binding for firms with 2007 total assets in (€6m, €86m] and binding for firms with 2007 total assets in (€86m, €166m].

Table B.1: Design of UK R&D Tax Relief Scheme, 2000-2012

Effective from		SME ceilings: Employment	SME ceilings: Total assets	SME ceilings: Turnover	Enhancement rate: SME	Enhancement rate: Large company	Payable credit rate: SME	Payable credit rate: Large company	Effective for
2000	April	249	€27m	€40m	50%	0%	16%	0%	Expenditure that incurred on or after April 1st, 2000
2002	April	"	"	"	"	25%	"	"	Expenditure that incurred on or after April 1st, 2002
2005	January	"	€43m	€50m	"	"	"	"	Accounting period that ended on or after January 1st, 2005
2008	April August	499	€86m	€100m	75%	30%	14%	"	Large companies: expenditure that incurred on or after April 1st, 2008 SMEs: expenditure that incurred on or after August 1st, 2008
2011	April	"	"	"	100%	"	12.5%	"	Expenditure that incurred on or after April 1st, 2011
2012	April	"	"	"	125%	"	"	"	Expenditure that incurred on or after April 1st, 2012

Note: To be considered an SME, a company must not exceed the employment ceiling and either the total asset ceiling or the sales ceiling ("ceiling tests"). The measurements and account aggregation rules for employment, total assets, and sales are set according to 1996/280/EC (up to December 31st, 2004) and 2003/361/EC (from January 1st, 2005). A company loses (acquires) its SME status if it fails (passes) the ceiling tests over two consecutive accounting periods (two-year rule). An SME working as subcontractor for a large company can only claim under the Large Company Scheme. From April 2000 to March 2012, there was a minimum requirement of £10,000 in qualifying R&D expenditure for both SMEs and large companies.

Table B.2: Tax-adjusted user cost of R&D capital over time

Tax relief scheme	(1) SME: Deduction	(2) SME: Payable credit	(3) SME: Average	(4) Large company: Deduction	(5) Large company: Payable credit	(6) Large company: Average	(7) **Arc % difference user cost**	(8) **Log difference user cost**
2006	0.157	0.152	0.154	0.179	0.200	0.190	-0.209	-0.210
2007	0.157	0.152	0.154	0.179	0.200	0.190	-0.209	-0.210
2008	0.147	0.151	0.149	0.177	0.200	0.190	-0.237	-0.238
2009	0.142	0.151	0.147	0.177	0.200	0.190	-0.254	-0.255
2010	0.142	0.151	0.147	0.177	0.200	0.190	-0.254	-0.255
2011	0.130	0.150	0.141	0.179	0.200	0.191	-0.300	-0.302
2006-2008	0.154	0.152	0.153	0.178	0.200	0.190	-0.218	-0.219
2009-2011	**0.138**	**0.151**	**0.145**	**0.177**	**0.200**	**0.190**	**-0.269**	**-0.271**

Note: Tax-adjusted user cost of R&D capital is calculated using formulae as described in sub-section 7.2. Corporate tax rate is 30% in 2006-2007, 28% in 2008-2010, and 26% in 2011. Enhancement rate is 50% for SMEs and 25% for large companies in 2006-2008, 75% for SMEs and 30% for large companies in 2008-2010, 100% for SMEs and 30% for large companies in 2011. Payable credit rate is 16% in 2006-2008, 14% in 2008-2010, and 12.5% in 2011. Share of the payable credit case is 55%. Real interest rate is 5%. Depreciation rate is 15%.

Table B.3: Robustness checks for R&D regressions

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**R&D expenditure, 2009-11 average *(£ '000)***									
Specification	Higher order polynomial controls		Alternative kernel weight		Alternative bandwidth around the asset threshold					
Below-asset-threshold indicator (in 2007)	189.9**	186.2*	144.0***	150.0**	182.8**	143.4**	204.2***	186.0***	121.0**	95.7**
	(84.7)	(108. 3)	(55.5)	(58.9)	(71.3)	(56.3)	(72.5)	(67.5)	(52.5)	(47.3)
Polynomial controls	2nd	3rd	1st	1st	1st	1st	2nd	2nd	1st	1st
Kernel weight			Epa	Tri					Tri	Tri
Sample assets (€m)	61-111	61-111	61-111	61-111	71-101	66-106	56-116	51-121	56-116	51-121
Firms	5,888	5,888	5,888	5,888	3,394	4,615	7,255	8,818	7,255	8,818

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	**R&D expenditure, 2009-11 average *(£ '000)***								
Specification	Industry & location fixed effects			Alternative LDV	Alternative winsorization parameter			Poisson	Negative binomial
Below-asset-threshold indicator (in 2007)	106.9*	121.7**	103.6**	60.8*	156.9**	87.3**	43.5*	1.31***	1.22**
	(57.2)	(52.0)	(52.2)	(33.9)	(64.6)	(38.6)	(25.0)	(0.49)	(0.49)
Fixed effects	Industry	Location	Ind. x Loc.						
Year of LDV				2007					
Winsorized window	2.5%	2.5%	2.5%	2.5%	1.0%	5.0%	No outliers	2.5%	2.5%
Firms	4,504	5,868	4,498	5,888	5,888	5,888	5,872	5,888	5,888

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. Robust standard errors are in brackets. **Panel A:** Columns (1) and (2) control for second or third order polynomials of running variable. The coefficients on the second and third order assets terms are not statistically significant. Columns (3) and (4) use Epanechnikov or triangular kernel weights. Columns (5) and (6) use samples with smaller bandwidths around the threshold, also controlling for first order polynomial of the running variable. Columns (7)-(10) use samples with larger bandwidths around the threshold. Columns (7) and (8) control for second order polynomial of the running variable to improve the fit (the coefficients on the second order assets terms are statistically significant for larger bandwidths). Columns (9) and (10) control for first order polynomial of the running variable and use triangular kernel weights. **Panel B:** Columns (1)-(3) add industry (4-digit SIC), location (2-digit postcode), and industry x location (2-digit SIC x 1-digit postcode) fixed effects. Column (4) adds R&D expenditure in 2007 as lagged dependent variable control. Columns (5)-(7) use samples with different winsorization parameter or sample excluding outliers in R&D expenditure. Column (8) and (9) uses Poisson and negative binomial specifications instead of OLS, to allow for a proportionate effect on R&D (as in a semi-log specification).

Table B.4: Robustness checks for reduced-form patent regressions

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**All patent family count, 2009-13 average**									
Specification	Higher order polynomial controls		Alternative kernel weight		Alternative bandwidth around the asset threshold					
Below-asset-threshold indicator (in 2007)	0.066	0.056	0.068**	0.067**	0.068	0.061**	0.057	0.091***	0.068***	0.063***
	(0.041)	(0.044)	(0.027)	(0.027)	(0.045)	(0.030)	(0.038)	(0.031)	(0.025)	(0.024)
Polynomial controls	2nd	3rd	1st	1st	1st	1st	2nd	2nd	1st	1st
Kernel weight			Epa	Tri					Tri	Tri
Sample assets (€m)	61-111	61-111	61-111	61-111	71-101	66-106	56-116	51-121	56-116	51-121
Firms	5,888	5,888	5,888	5,888	3,394	4,615	7,255	8,818	7,255	8,818

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	**All patent family count, 2009-13 average**								
Specification	Industry & location fixed effects			Alternative LDV	Alternative winsorization Parameter			Poisson	Negative binomial
Below-asset-threshold indicator (in 2007)	0.063*	0.065***	0.061**	0.047**	0.067**	0.063***	0.070***	1.29***	1.46***
	(0.034)	(0.025)	(0.024)	(0.022)	(0.027)	(0.024)	(0.026)	(0.46)	(0.47)
Fixed effects	Industry	Location	Ind. x Loc.						
Year of LDV				2007					
Winsorized window	2.5%	2.5%	2.5%	2.5%	1.0%	5.0%	No outliers	2.5%	2.5%
Firms	4,504	5,868	4,498	5,888	5,888	5,888	5,872	5,888	5,888

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. Robust standard errors are in brackets. **Panel A:** Columns (1) and (2) control for second or third order polynomials of running variable. The coefficients on the second and third order assets terms are not statistically significant. Columns (3) and (4) use Epanechnikov or triangular kernel weights. Columns (5) and (6) use samples with smaller bandwidths around the threshold, also controlling for first order polynomial of the running variable. Columns (7)-(10) use samples with larger bandwidths around the threshold. Columns (7) and (8) control for second order polynomial of the running variable to improve the fit (the coefficients on the second order assets terms are statistically significant for larger bandwidths). Columns (9) and (10) control for first order polynomial of the running variable and use triangular kernel weights. **Panel B:** Columns (1)-(3) add industry (4-digit SIC), location (2-digit postcode), and industry x location (2-digit SIC x 1-digit postcode) fixed effects. Column (4) adds all patent family count in 2007 as lagged dependent variable control. Columns (5)-(7) use samples with different winsorization parameter or sample excluding outliers in all patent family count. Column (8) and (9) uses Poisson and negative binomial specifications instead of OLS, to allow for a proportionate effect on patents (as in a semi-log specification).

Table B.5: Robustness checks for effects of R&D on patents (IV regressions)

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**All patent family count, 2009-13 average**									
Specification	Higher order polynomial controls		Alternative kernel weight		Alternative bandwidth around the asset threshold					
R&D expenditure *(£m)*, 2009-11 avg.	0.345	0.301	0.475**	0.449**	0.370	0.428*	0.278	0.489**	0.558**	0.655*
	(0.227)	(0.248)	(0.232)	(0.221)	(0.242)	(0.236)	(0.191)	(0.213)	(0.280)	(0.363)
Polynomial controls	2nd	3rd	1st	1st	1st	1st	2nd	2nd	1st	1st
Kernel weight			Epa	Tri					Tri	Tri
Sample assets (€m)	61-111	61-111	61-111	61-111	71-101	66-106	56-116	51-121	56-116	51-121
Firms	5,888	5,888	5,888	5,888	3,394	4,615	7,255	8,818	7,255	8,818

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	**All patent family count, 2009-13 average**							
Specification	Industry & location fixed effects			LDV control		Alternative winsorization parameter		
R&D expenditure *(£m)*, 2009-11 avg.	0.587	0.534*	0.589	0.434*	0.421*	0.428*	0.721**	1.597*
	(0.435)	(0.304)	(0.411)	(0.243)	(0.251)	(0.224)	(0.355)	(0.939)
Fixed effects	Industry	Location	Ind. x Loc.					
Year of LDV				2006-08 average	2007			
Winsorized window	2.5%	2.5%	2.5%	2.5%	2.5%	1.0%	5.0%	No outliers
Firms	4,504	5,868	4,498	5,888	5,888	5,888	5,888	5,872

Note: *** significant at 1% level, ** 5% level, * 10% level. IV estimates based on the (fuzzy) RD Design. Instrumental variable is the indicator whether total assets in 2007 is below €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of the running variable (total assets in 2007) separately for each side of the threshold are included. Robust standard errors are in brackets. **Panel A:** Columns (1) and (2) control for second or third order polynomials of running variable. Columns (3) and (4) use Epanechnikov or triangular kernel weights. The coefficients on the second and third order assets terms are not statistically significant. Columns (5) and (6) use samples with smaller bandwidths around the threshold, also controlling for first order polynomial of the running variable. Columns (7)-(10) use samples with larger bandwidths around the threshold. Columns (7) and (8) control for second order polynomial of the running variable to improve the fit (the coefficients on the second order assets terms are statistically significant for larger bandwidths). Columns (9) and (10) control for first order polynomial of the running variable and use triangular kernel weights. **Panel B:** Columns (1)-(3) add industry (4-digit SIC), location (2-digit postcode), and industry x location (2-digit SIC x 1-digit postcode) fixed effects. Columns (4) and (5) add average all patent family count over 2006-2008 or all patent family count in 2007 as lagged dependent variable control. Columns (6)-(8) use samples with different winsorization parameter or sample excluding outliers in all patent family count.

Table B.6: Additional results on effects of R&D tax relief on quality-adjusted patents

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable (2009-13 average)	**Patent count weighted by citations**			**All patent family count weighted by quality index**				**All patent family count in top quality quartile, by**		
	EPO patents	UK patents	US patents	Scaled citation	Scope	Gene-rality	Origi-nality	Scope	Gene-rality	Origi-nality
Below-asset-threshold indicator (in 2007)	0.013*	0.044*	0.056*	1.729*	0.132**	0.010**	0.030***	0.038***	0.051***	0.036***
	(0.008)	(0.026)	(0.034)	(0.954)	(0.052)	(0.005)	(0.012)	(0.013)	(0.019)	(0.014)
Dependent variable mean over 2006-08	*0.025*	*0.114*	*0.125*	*2.881*	*0.130*	*0.017*	*0.027*	*0.027*	*0.037*	*0.027*
Discontinuity estimate to baseline mean ratio	*0.53*	*0.38*	*0.45*	*0.60*	*1.02*	*0.59*	*1.12*	*1.40*	*1.38*	*1.35*
Firms	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable (2009-13 average)	**All patent family count**		**EPO patent count**					
	BTP patents	Non-BTP patents	Chem. patents	Non-chem. patents	BTP patents	Non-BTP patents	ICT patents	Non-ICT patents
Below-asset-threshold indicator (in 2007)	0.0083**	0.0573**	0.0125**	0.0206*	0.0075**	0.0262*	0.0036*	0.0262**
	(0.0034)	(0.0242)	(0.0059)	(0.0123)	(0.0032)	(0.0146)	(0.0019)	(0.0130)
Dependent variable mean over 2006-08	*0.0030*	*0.0573*	*0.0068*	*0.0211*	*0.0012*	*0.0276*	*0.0015*	*0.0270*
Discontinuity estimate to baseline mean ratio	*2.75*	*1.00*	*1.84*	*0.98*	*6.36*	*0.95*	*2.40*	*0.97*
Firms	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of the running variable separately for each side of the threshold are included. Robust standard errors are in brackets. **Panel A:** Columns (1)-(3) weight EPO patent count, UK patent count, or US patent count by citations. Columns (4)-(7) weight all patent family count by scaled citation (column 4), patent scope (column 5), patent generality index (column 6), or patent originality index (column 7). Scaled citation measures a patent's citations relative to the average citations of patents in the same patent sector x filing office x filing year cell. Patent scope counts the number of 4-digit patent classes in which a patent is classified. Generality index measures the patent-class diversity of a patent's forward citations. Originality index measures the patent-class diversity of a patent's backward citations. Columns (8)-(10) count all patent families in the top 25% in quality of their patent field x filing year cohorts, with patent quality measured by patent scope (column 8), generality index (column 9), originality index (column 10). **Panel B:** Columns (1) and (2) split all patent counts into biotechnology and pharmaceutical (BTP) patents and non-BTP patents. Column (3)-(8) split EPO patent counts into chemistry/pharmaceutical and non-chemistry/pharmaceutical patents (columns 3 and 4), BTP and non-BTP patents (columns 5 and 6), and ICT and non-ICT patents (column 7 and 8). Chemistry/pharmaceutical patents include all patents classified into patent sector (3) Chemistry. BTP patents include all patents classified into either patent field (11) Analysis of biological materials, (15) Biotechnology, or (16) Pharmaceuticals (i.e., a subset of chemistry/pharmaceutical patents). ICT patents include all patents classified into either patent field (4) Digital communication, (6) Computer technology, or (7) IT methods for management.

Table B.7: Discontinuities in the probabilities of doing any R&D or filing any patents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	**Indicator: R&D exp. > 0**			**Indicator: All patent family count > 0**				
Year	2009	2010	2011	2009	2010	2011	2012	2013
Below-asset-threshold indicator	0.008	0.006	0.013	0.011*	0.008	0.014*	0.013**	0.018***
(in 2007)	(0.011)	(0.012)	(0.011)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)
Dependent variable mean	*0.036*	*0.041*	*0.045*	*0.017*	*0.017*	*0.017*	*0.015*	*0.016*
Firms	5,888	5,888	5,888	5,888	5,888	5,888	5,888	5,888

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. Robust standard errors are in brackets. Dependent variables are indicators of whether a firm has R&D expenditure or files patents during the corresponding year.

Table B.8: Heterogeneous effects of R&D tax relief by past R&D and patents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**R&D expenditure (£ *'000*), 2009-11 avg.**		**All patent family count, 2009-13 avg.**		**EPO patent count, 2009-13 avg.**		**UK patent count, 2009-13 avg.**		**US patent count, 2009-13 avg.**	
Subsample	Past > 0	Past = 0	Past > 0	Past = 0	Past > 0	Past = 0	Past > 0	Past = 0	Past > 0	Past = 0
Below-asset-threshold	1,708*	6.3	1.50**	0.002	1.40**	-0.000	1.80**	0.007	1.89***	-0.002
indicator (in 2007)	(885)	(9.6)	(0.68)	(0.005)	(0.63)	(0.002)	(0.91)	(0.005)	(0.66)	(0.002)
Dependent variable mean over 2006-08	*1,682*	*0.0*	*2.18*	*0.00*	*1.51*	*0.00*	*2.96*	*0.00*	*1.42*	*0.00*
Difference	1,702*		1.50**		1.40**		1.79**		1.89***	
	(879)		(0.67)		(0.62)		(0.90)		(0.65)	
Firms	259	5,629	172	5,716	117	5,771	152	5,736	106	5,782

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. Robust standard errors are in brackets. Past period is the pre-policy period of 2006-2008.

Table B.9: Heterogeneous effects of R&D tax relief by industry patenting intensity

Panel A. OLS regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**R&D expenditure (£ '000), 2009-11 avg.**		**All patent family count, 2009-13 avg.**		**EPO patent count, 2009-13 avg.**		**UK patent count, 2009-13 avg.**		**US patent count, 2009-13 avg.**	
Subsample	High patent	Low patent	High patent	Low patent	High patent	Low patent	High patent	Low patent	High patent	Low patent
Below-asset-threshold indicator (in 2007)	167.4*	107.8	0.160**	0.017	0.078**	0.014	0.184**	0.017	0.084**	0.009
	(95.2)	(68.3)	(0.065)	(0.011)	(0.039)	(0.012)	(0.073)	(0.014)	(0.038)	(0.007)
Dependent variable mean over 2006-08	*124.7*	*25.0*	*0.118*	*0.020*	*0.058*	*0.007*	*0.140*	*0.024*	*0.047*	*0.006*
Difference	59.5		0.142**		0.064		0.167**		0.075*	
	(117.2)		(0.066)		(0.041)		(0.074)		(0.039)	
Firms	2,272	2,232	2,272	2,232	2,272	2,232	2,272	2,232	2,272	2,232

Panel B. IV regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable (2009-13 average)	**All patent family count**		**EPO patent count**		**UK patent count**		**US patent count**	
Subsample	High patent	Low patent	High patent	Low patent	High patent	Low patent	High patent	Low patent
R&D expenditure *(£m)*, 2009-11 average	0.954	0.161	0.463	0.128	1.101	0.162	0.500	0.080
	(0.607)	(0.103)	(0.312)	(0.081)	(0.687)	(0.158)	(0.325)	(0.051)
Firms	2,272	2,232	2,272	2,232	2,272	2,232	2,272	2,232

Note: *** significant at 1% level, ** 5% level, * 10% level. Robust standard errors are in brackets. Industry patenting intensity is calculated as the share of firms in the 4-digit industry having filed any patent before 2007. High (low) patenting subsample includes firms in industries above (below) median in patenting intensity. Examples of high-patenting industries include electric domestic appliances, basic pharmaceutical products, medical and surgical equipment, organic and inorganic basic chemicals, optical and photographic equipment, etc. **Panel A:** OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. **Panel B:** IV estimates based on the (fuzzy) RD Design. Instrumental variable is the indicator whether total assets in 2007 is below €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomial of RDD running variable (total assets in 2007) separately for each side of the threshold are included.

Table B.10: Heterogeneous effects of R&D tax relief by Biotech-pharma (BTP)

Panel A. Biotech/pharmaceutical (BTP) vs. non-BTP industries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**R&D expenditure (£ '000), 2009-11 avg.**		**All patent family count, 2009-13 avg.**		**EPO patent count, 2009-13 avg.**		**UK patent count, 2009-13 avg.**		**US patent count, 2009-13 avg.**	
Subsample	BTP ind.	Non-BTP ind.	BTP ind.	Non-BTP ind.	BTP ind.	Non-BTP ind.	BTP ind.	Non-BTP ind.	BTP ind.	Non-BTP ind.
Below-asset-threshold indicator (in 2007)	177.0	100.2*	0.116**	0.050*	0.073*	0.021	0.109*	0.064*	0.0600*	0.033*
	(109.3)	(58.0)	(0.057)	(0.029)	(0.039)	(0.016)	(0.062)	(0.036)	(0.0313)	(0.0192)
Dependent variable mean over 2006-08	*105.1*	*61.3*	*0.099*	*0.049*	*0.054*	*0.020*	*0.111*	*0.062*	*0.039*	*0.020*
Discontinuity estimate to baseline mean ratio	*1.68*	*1.64*	*1.17*	*1.01*	*1.34*	*1.05*	*0.98*	*1.04*	*1.53*	*1.65*
Difference	76.8		0.066		0.052		0.045		0.027	
	(123.7)		(0.064)		(0.043)		(0.072)		(0.037)	
Firms	1,709	4,179	1,709	4,179	1,709	4,179	1,709	4,179	1,709	4,179

Panel B. Information and communication technology (ICT) vs. non-ICT industries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**R&D expenditure (£ '000), 2009-11 avg.**		**All patent family count, 2009-13 avg.**		**EPO patent count, 2009-13 avg.**		**UK patent count, 2009-13 avg.**		**US patent count, 2009-13 avg.**	
Subsample	ICT ind.	Non-ICT ind.	ICT ind.	Non-ICT ind.	ICT ind.	Non-ICT ind.	ICT ind.	Non-ICT ind.	ICT ind.	Non-ICT ind.
Below-asset-threshold indicator (in 2007)	201.6**	82.3	0.078*	0.0645*	0.062	0.023	0.091*	0.070*	0.060	0.031**
	(101.8)	(59.1)	(0.047)	(0.032)	(0.039)	(0.014)	(0.053)	(0.038)	(0.038)	(0.015)
Dependent variable mean over 2006-08	*101.2*	*60.3*	*0.065*	*0.063*	*0.032*	*0.029*	*0.075*	*0.077*	*0.027*	*0.025*
Discontinuity estimate to baseline mean ratio	*1.99*	*1.36*	*1.20*	*1.02*	*1.91*	*0.80*	*1.21*	*0.91*	*2.24*	*1.24*
Difference	119.3		0.013		0.039		0.021		0.029	
	(117.7)		(0.057)		(0.041)		(0.066)		(0.041)	
Firms	1,969	3,919	1,969	3,919	1,969	3,919	1,969	3,919	1,969	3,919

Note: *** significant at 1% level, ** 5% level, * 10% level. Robust standard errors are in brackets. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. **Panel A:** Biotechnology and pharmaceutical (BTP) patents are those classified into either patent field (11) Analysis of biological materials, (15) Biotechnology, or (16) Pharmaceuticals. BTP-intensive industries (columns 1, 3, 5, 7, and 9) are top 20 3-digit industries in total number of BTP patent applications over 2006-2011. **Panel B:** Information and communication technology (ICT) patents are those classified into either patent field (4) Digital communication, (6) Computer technology, or (7) IT methods for management. ICT-intensive industries (columns 1, 3, 5, 7, and 9) are top 20 3-digit industries in total number of ICT patent applications over 2006-2011.

Table B.11: Heterogeneous effects of R&D tax relief by firms' past capital investments

Panel A. OLS regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	**R&D expenditure (£ '000), 2009-11 avg.**		**All patent family count, 2009-13 avg.**		**EPO patent count, 2009-13 avg.**		**UK patent count, 2009-13 avg.**		**US patent count, 2009-13 avg.**	
Subsample	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0
Below-asset-threshold indicator (in 2007)	305.5***	-36. 7	0.148***	-0.000	0.079**	-0.002	0.166**	-0.000	0.088**	-0.000
	(106.4)	(30.0)	(0.055)	(0.013)	(0.034)	(0.007)	(0.065)	(0.015)	(0.034)	(0.007)
Dependent variable mean over 2006-08	*159.6*	*4.4*	*0.123*	*0.016*	*0.058*	*0.007*	*0.147*	*0.019*	*0.048*	*0.007*
Difference	342.2***		0.148***		0.080**		0.166**		0.088**	
	(110.6)		(0.056)		(0.035)		(0.067)		(0.035)	
Firms	2,640	3,248	2,640	3,248	2,640	3,248	2,640	3,248	2,640	3,248

Panel B. IV regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable (2009-13 average)	**All patent family count**		**EPO patent count**		**UK patent count**		**US patent count**	
Subsample	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0	Past inv. > 0	Past inv. = 0
R&D expenditure *(£m)*, 2009-11 average	0.483**	0.004	0.257**	0.043	0.542**	0.002	0.288**	0.001
	(0.217)	(0.351)	(0.121)	(0.195)	(0.256)	(0.394)	(0.130)	(0.189)
Firms	2,640	3,248	2,640	3,248	2,640	3,248	2,640	3,248

Note: *** significant at 1% level, ** 5% level, * 10% level. Robust standard errors are in brackets. Past capital investments is calculated as average machinery and plant investments over 2005-2007 reported in CT600 (as coverage of capital expenditure in FAME is limited). **Panel A:** OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of running variable separately for each side of the threshold are included. **Panel B:** IV estimates based on the RD Design. Instrumental variable is the indicator whether total assets in 2007 is below €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomial of RDD running variable (total assets in 2007) separately for each side of the threshold are included.

Table B.12: Effects of R&D tax relief on other expense categories

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample	Full baseline sample					R&D performing firms				
Dependent variable (2009-11 average, *£ '000*)	**Admin exp.**	**Admin exp., excl. R&D**	**Total exp., excl. R&D**	**Capex imputed from PPE**	**Qual. M&P exp.**	**Admin exp.**	**Admin exp., excl. R&D**	**Total exp., excl. R&D**	**Capex imputed from PPE**	**Qual. M&P exp.**
Below-asset-threshold indicator (in 2007)	480	287	-1,301	20	32	1,553	-344	-5,254	-311	254
	(1,179)	(1,171)	(3,558)	(230)	(40)	(4,197)	(4,138)	(11,947)	(510)	(226)
Dependent variable mean over 2006-08	*14,806*	*14,715*	*42,875*	*3,464*	*505*	*23,490*	*22,340*	*71,470*	*2,459*	*1,743*
Firms	4,441	4,441	4,569	3,061	5,575	323	323	326	318	329

Notes: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Controls for first order polynomials of running variable separately for each side of the threshold are included. Robust standard errors are in brackets. Columns (1)-(5) employ the full baseline sample for firms with total assets in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Columns (6)-(10) use the subsample of R&D performing firms during 2009-2011 that are in the baseline sample. The dependent variables are average over the post-policy years for which data are not missing. Columns (1) and (6) look at total administrative expenses reported in FAME (columns (1) and (6)). Columns (2) and (7) look at total administrative expenses minuses qualifying R&D expenditure. Columns (3) and (8) look at total expenses reported in FAME minuses qualifying R&D expenditure. Column (4) and (9) look at capital expenditure imputed from net change in balance sheet's property, plant, and equipment reported in FAME. Column (5) and (10) look at qualifying machinery and plant investments reported in CT600 (for capital allowance tax relief purpose).

Table B.13: Effects of R&D tax relief on other measures of firms performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Before (pre-policy)			After (post-policy)					Before	5yr After	5yr Diff.
Year	2006	2007	2008	2009	2010	2011	2012	2013	2006-08 average	2009-11 average	5yr After - Before
Panel A. Dependent variable: Ln(Sales)											
Below-asset-threshold indicator (in 2007)	-0.187	0.029	-0.102	0.212	0.404**	0.307	0.198	0.188	-0.023	0.170	0.193
	(0.170)	(0.167)	(0.162)	(0.180)	(0.187)	(0.192)	(0.204)	(0.217)	(0.157)	(0.181)	(0.123)
Firms	3,292	3,439	3,394	3,312	3,296	3,260	3,207	3,153	3,451	3,451	3,451
Panel B. Dependent variable: Ln(Employment)											
Below-asset-threshold indicator (in 2007)	-0.012	0.102	0.079	0.104	0.258*	0.283*	0.289*	0.364**	0.0215	0.240*	0.219**
	(0.126)	(0.123)	(0.131)	(0.140)	(0.148)	(0.153)	(0.156)	(0.160)	(0.125)	(0.143)	(0.095)
Firms	2,468	2,548	2,430	2,443	2,553	2,470	2,370	2,281	2,403	2,403	2,403
Panel C. Dependent variable: Ln(Capital)											
Below-asset-threshold indicator (in 2007)	-0.013	-0.032	-0.007	-0.016	-0.004	0.015	0.070	0.125	-0.065	0.010	0.075
	(0.120)	(0.109)	(0.113)	(0.122)	(0.131)	(0.135)	(0.142)	(0.146)	(0.108)	(0.125)	(0.084)
Firms	3,724	3,959	3,793	3,609	3,457	3,322	3,205	3,074	3,665	3,665	3,665
Panel D. Dependent variable: Total factor productivity											
Below-asset-threshold indicator (in 2007)	-0.069	0.037	0.020	0.178	0.265	0.127	0.146	0.184	0.070	0.210	0.140
	(0.171)	(0.162)	(0.152)	(0.166)	(0.173)	(0.178)	(0.191)	(0.201)	(0.157)	(0.163)	(0.113)
Firms	1,590	1,629	1,575	1,527	1,508	1,487	1,418	1,367	1,605	1,605	1,605

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e., between €61m and €111m). Controls for first order polynomials of the running variable separately for each side of the threshold and 2-digit industry dummies are included. Robust standard errors are in brackets. **Panel A** uses sales from CT600. **Panel B** uses employment (from FAME). **Panel C** uses fixed assets (from FAME). **Panel D** uses total factor productivity calculated as $\ln(value\ added) - \propto_k \ln(capital) - \propto_l \ln(wages)$, in which value added sales minus imputed materials and $\propto_k$ and $\propto_l$ are estimated using Olley-Pakes production function estimation separately for each 2-digit industry across all firms in the FAME dataset over the 2000-2005 period. Columns (9)-(10) condition on the "balanced" sample where we observe the outcome variable in at least one year of the pre-policy sample and one year of the post-policy sample (i.e., it is a subsample of the observations in columns (1)-(8)).

Table B.14: Estimating impacts of R&D tax relief using other SME criteria

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SME criterion	Total assets		Sales				Employment	
Dependent variable	**R&D exp. *(£ '000),* 09-11 avg.**	**All patent count, 09-13 avg.**	**R&D exp. *(£ '000),* 09-11 avg.**	**All patent count, 09-13 avg.**	**R&D exp. *(£ '000),* 09-11 avg.**	**All patent count, 09-13 avg.**	**R&D exp. *(£ '000),* 09-11 avg.**	**All patent count, 09-13 avg.**
Below-SME-threshold indicator (in 2007)	123.3**	0.069***	138.6**	0.027	152.1	0.083	86.4	0.138**
	(52.1)	(0.026)	(64.2)	(0.044)	(123.2)	(0.065)	(104.6)	(0.056)
Dependent variable mean over 2006-08	*74.0*	*0.064*	*119.5*	*0.087*	*194.3*	*0.122*	*209.4*	*0.148*
Discontinuity estimate to baseline mean ratio	*1.67*	*1.09*	*1.16*	*0.31*	*0.78*	*0.68*	*0.41*	*0.93*
Sample	Total assets in [€61m, €111m]		Sales in [€50m, €150m]		Sales in [€50m, €150m] & total assets > €86m		Employment in [300, 700]	
Firms	5,888	5,888	7,101	7,101	2,085	2,085	4,526	4,526

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)
Specification	First stage	Reduced form	IV	First stage	Reduced form	IV
Dependent variable	**R&D exp. *(£ '000),* 09-11 avg.**	**All patent count, 09-13 avg.**	**All patent count, 09-13 avg.**	**R&D exp. *(£ '000),* 09-11 avg.**	**All patent count, 09-13 avg.**	**All patent count, 09-13 avg.**
Bellow-asset-threshold indicator (in 2007)	107.9*	0.129***		68.2*	0.072***	
	(57.6)	(0.045)		(37.3)	(0.026)	
Below-sales-threshold indicator (in 2007)	131.4**	0.024		71.6*	-0.013	
	(63.8)	(0.044)		(40.0)	(0.023)	
R&D expenditure (*£m*), 2009-11 average			0.696**			0.366
			(0.334)			(0.307)
Dependent variable mean over 2006-08	*119.5*	*0.087*	*0.087*	*105.0*	*0.080*	0.080
Joint F-statistics (p-value)	3.26 (0.04)	4.73 (0.01)		2.30 (0.10)	5.70 (0.00)	
Sample	Sales in [€50m, €150m]			Total assets in [€61m, €111m] or sales in [€50m, €150m]		
Firms	7,091	7,091	7,091	9,751	9,751	9,751

Note: *** significant at 1% level, ** 5% level, * 10% level. Robust standard errors are in brackets. **Panel A:** OLS estimates based on the RD Design. The running variable in columns (1) and (2) is total assets in 2007 with threshold of €86m. The running variable in columns (3) and (6) is sales in 2007 with threshold of €100m. The running variable in columns (7) and (8) is employment in 2007 with threshold of 499. Controls for first order polynomials of running variable separately for each side of the threshold are included. **Panel B:** OLS estimates based on the RD Design for R&D and patent regressions (columns 1-2 and 4-5). IV estimates based on the (fuzzy) RD Design where the instrumental variable is the indicator whether total assets in 2007 is below €86m (columns 3 and 6). The running variables are total assets in 2007 with threshold of €86m and sales in 2007 with threshold of €100m. Instrumental variable in columns (3) and (6) are the indicator whether total assets in 2007 is below €86m and the indicator whether sales in 2007 is below €100m. Controls for first order polynomials of the running variables (total assets in 2007 and sales in 2007) separately for each side of the respective threshold are included. Reported joint F-statistics for are for below-asset-threshold indicator and below-sales-threshold indicator. P-values of Anderson-Rubin weak-instrument-robust inference tests in columns (3) and (6) are 0.009 and 0.003 respectively.

Table B.15: Tax-price elasticities of R&D and patents using different approaches

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	SME status	**R&D expenditure (£ '000)**				**All patent family count**				**R&D user cost**	**Elasticity**	
Approach	Fuzziness estimate	Discon-tinuity estimate	Adjusted discon-tinuity estimate	Pre-policy baseline mean	R&D difference	Discon-tinuity estimate	Adjusted discon-tinuity estimate	Pre-policy baseline mean	Patent difference	Tax-adjusted user cost difference	R&D (wrt. R&D cost)	Patent (wrt. R&D cost)
(1) Baseline	**0.353**	**60.4**	**171.2**	**74.0**	**1.073**	**0.042**	**0.119**	**0.064**	**0.964**	**0.269**	**3.989**	**3.583**
(2) Log difference elasticity	0.353	60.4	171.2	74.0	**1.198**	0.042	0.119	0.064	**1.051**	**0.271**	4.422	3.878
(3) SME status over 2009-11	**0.248**	60.4	243.7	74.0	1.245	0.042	0.169	0.064	1.139	0.269	4.626	4.236
(4) SME status over 2008-09	**0.464**	60.4	130.3	74.0	0.936	0.042	0.090	0.064	0.829	0.269	3.481	3.081
(5) LDV discontinuity estimate	0.353	**63.4**	179.5	74.0	1.096	**0.049**	0.140	0.064	1.046	0.269	4.076	3.889
(6) Pre-policy mean over 2006-07	0.353	60.4	171.2	**77.6**	1.049	0.042	0.119	**0.065**	0.953	0.269	3.899	3.544
(7) 2007 assets in [€51m, €121m]	**0.345**	**51.8**	150.2	**69.8**	1.037	**0.038**	0.109	**0.058**	0.968	0.269	3.855	3.599
(8) Small profits corporate tax rate	0.353	60.4	171.2	74.0	1.073	0.042	0.119	0.064	0.964	**0.228**	4.706	4.227

Note: Baseline approach (i.e., arc elasticity) in row (1) is explained in detail in sub-section 7.2 and the note to Table A16 Panel A. Log-difference-elasticity approach in row (2) is explained in detail in the note to Table A16 Panel B. Rows (3)-(8) employ the baseline arc-elasticity approach as in row (1) and different alternative input estimates. Rows (3) and (4) use alternative estimates for how "sharp" the below-asset-threshold indicator is as an instrument for SME status, based on SME status over 2009-2011 (row 2) and SME status over 2008-2009 (row 3). These estimates are reported in Table 2 columns (4) and (6) respectively. Row (5) uses the discontinuity estimates with lagged dependent variable control from Table 3 column (10) (for R&D) and column (7) of Table 5 Panel B (for patents). Row (6) uses average R&D and patents over 2006-2007 as the pre-policy baseline means. Row (7) use larger baseline sample of firms with 2007 total assets between €51m and €121m and triangular weights. Relevant estimates are reported in Tables A3-5. Row (8) applies the small profits corporate tax rate in calculations of tax-adjusted user costs (see Appendix A.5). Differences between each approach and the baseline case in row (1) are highlighted in bold (for input estimates only).

Table B.16: Bootstrapping elasticity estimates

Panel A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	SME status	**R&D expenditure (£ '000)**				**All patent family count**				**Elasticity**	
	First-stage estimate	3yr After - Before estimate	Adjusted 3yr After - Before estimate	Pre-policy baseline mean	Arc % R&D difference	5yr After - Before estimate	Adjusted 5yr After - Before estimate	Pre-policy baseline mean	Arc % patent difference	R&D (wrt. R&D user cost)	Patent (wrt. R&D user cost)
Baseline sample estimates	0.353	60.4	171.2	74.0	1.073	0.042	0.119	0.064	0.964	**3.989**	**3.583**
Bootstrapped distribution											
5th percentile	0.206	8.1	24.6	58.4	0.292	0.008	0.019	0.049	0.301	**1.085**	**1.119**
10th percentile	0.236	19.8	50.9	61.5	0.529	0.016	0.042	0.052	0.502	1.966	1.866
25th percentile	0.293	39.3	108.4	67.4	0.837	0.027	0.074	0.057	0.738	3.113	2.743
50th percentile	0.357	60.4	169.3	73.8	1.079	0.042	0.118	0.064	0.963	4.010	3.580
75th percentile	0.414	82.2	247.1	80.1	1.248	0.056	0.170	0.070	1.145	4.640	4.258
90th percentile	0.468	103.7	337.1	86.1	1.380	0.072	0.232	0.076	1.292	5.130	4.801
95th percentile	0.501	119.1	404.3	90.6	1.462	0.081	0.282	0.079	1.385	5.436	5.148

Note: Panel A reports baseline estimators used to calculate "arc-percentage-difference" R&D and patent elasticities, together with their empirical distributions (see sub-section 7.2 for details). The estimators' empirical distributions are derived from 1,000 bootstrap replications. In each replication, we draw with replacement 361 observations from the subsample of 361 post-policy R&D performing firms, and 5,527 (=5,888-361) observations from the remaining subsample of 5,527 firms. Column (1) reports the discontinuity estimate in Table 3 column (5) and its empirical distribution. Column (2) corresponds to Table 4 column (9); column (4) – R&D pre-policy baseline mean; column (6) – Table 6 Panel B column (6); column (8) – patent pre-policy baseline mean. Column (3) reports policy-induced R&D, estimated as $\frac{col.(2)}{col.(1)}$. Column (5) reports policy-induced percentage difference in R&D, $\frac{R_{SME}-R_{LCO}}{(R_{SME}+R_{LCO})/2}$, estimated as $\frac{col.(3)}{col.(3)/2+col.(4)}$. Column (7) reports policy-induced patents, estimated as $\frac{col.(5)}{col.(1)}$. Column (9) reports policy-induced percentage difference in patent, $\frac{PAT_{SME}-PAT_{LCO}}{(PAT_{SME}+PAT_{LCO})/2}$, estimated as $\frac{col.(7)}{col.(7)/2+col.(8)}$. Column (10) reports R&D elasticity with respect to its tax-adjusted user cost, $\frac{\%\text{ difference in } R}{\%\text{ difference in } \rho}$, estimated as $\frac{col.(5)}{0.269}$ (percentage difference in user cost is 0.269, see Table A2 column (7)). Column (11) reports patent elasticity with respected to R&D tax-adjusted user cost, $\frac{\%\text{ difference in } PAT}{\%\text{ difference in } \rho}$, estimated as $\frac{col.(9)}{0.269}$.

Panel B.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	SME status	**R&D expenditure**				**All patent family count**				**Elasticity**	
	First-stage estimate	3yr After - Before estimate	Adjusted 3yr After - Before estimate	Pre-policy baseline mean	Log R&D difference	5yr After - Before estimate	Adjusted 5yr After - Before estimate	Pre-policy baseline mean	Log patent difference	R&D (wrt. R&D user cost)	Patent (wrt. R&D user cost)
Baseline sample estimates	0.353	60.4	171.2	74.0	1.198	0.042	0.119	0.064	1.051	**4.422**	**3.878**
Bootstrapped distribution											
5th percentile	0.206	8.1	24.6	58.4	0.302	0.008	0.019	0.049	0.303	**1.113**	**1.119**
10th percentile	0.236	19.8	50.9	61.5	0.544	0.016	0.042	0.052	0.513	2.006	1.893
25th percentile	0.293	39.3	108.4	67.4	0.893	0.027	0.074	0.057	0.774	3.296	2.857
50th percentile	0.357	60.4	169.3	73.8	1.207	0.042	0.118	0.064	1.050	4.454	3.874
75th percentile	0.414	82.2	247.1	80.1	1.464	0.056	0.170	0.070	1.303	5.404	4.808
90th percentile	0.468	103.7	337.1	86.1	1.696	0.072	0.232	0.076	1.536	6.260	5.668
95th percentile	0.501	119.1	404.3	90.6	1.864	0.081	0.282	0.079	1.705	6.876	6.293

Note: Panel B reports baseline estimators used to calculate "log-difference" R&D and patent elasticities, together with their empirical distributions. The estimators' empirical distributions are derived from 1,000 bootstrap replications. In each replication, we draw with replacement 361 observations from the subsample of 361 post-policy R&D performing firms, and 5,527 (=5,888-361) observations from the remaining subsample of 5,527 firms. Column (1) reports the discontinuity estimate in Table 3 column (5) and its empirical distribution. Column (2) corresponds to Table 4 column (9); column (4) – R&D pre-policy baseline mean; column (6) – Table 6 Panel B column (6); column (8) – patent pre-policy baseline mean. Column (3) reports policy-induced R&D, estimated as $\frac{col.(2)}{col.(1)}$. Column (5) reports policy-induced log difference in R&D, $\ln\frac{R_{SME}}{R_{LCO}}$, estimated as $\ln\frac{col.(3)+col.(4)}{col.(4)}$. Column (7) reports policy-induced patents, estimated as $\frac{col.(5)}{col.(1)}$. Column (9) reports policy-induced log difference in patent, $\ln\frac{PAT_{SME}}{PAT_{LCO}}$, estimated as $\ln\frac{col.(7)+col.(8)}{col.(8)}$. Column (10) reports R&D elasticity with respect to its tax-adjusted user cost, $\frac{\ln(R_{SME}/R_{LCO})}{\ln(\rho_{SME}/\rho_{LCO})}$, estimated as $\frac{col.(5)}{0.271}$ (log difference in user cost is 0.271, see Table A2 column (8)). Column (11) reports patent elasticity with respected to R&D tax-adjusted user cost, $\frac{\ln(PAT_{SME}/PAT_{LCO})}{\ln(\rho_{SME}/\rho_{LCO})}$, estimated as $\frac{col.(9)}{0.271}$.

Table B.17: Value for money analysis of R&D Tax Relief Scheme

Year	2006	2007	2008	2009	2010	2011	2006-11 average
Panel A. Policy parameters							
SME enhancement rate e_{SME}	50%	50%	67%	75%	75%	100%	
SME payable credit rate c_{SME}	16%	16%	15%	14%	14%	12.5%	
SME effective corporate tax rate τ_{SME}	19%	19%	21%	21%	21%	20%	
LCO enhancement rate e_{LCO}	25%	25%	30%	30%	30%	30%	
LCO effective corporate tax rate τ_{LCO}	30%	30%	28%	28%	28%	26%	
Panel B. SME tax deduction case							
Tax-adjusted user cost of R&D ρ	0.177	0.177	0.165	0.160	0.160	0.150	
Value for money ratio μ	4.19	4.19	3.99	3.89	3.89	3.63	3.87
Exchequer costs Δ_{EC} *(£m)*	50	60	80	130	160	210	115
Additional R&D Δ_R *(£m)*	210	251	319	506	622	762	445
Panel C. SME payable tax credit case							
Tax-adjusted user cost of R&D ρ	0.152	0.152	0.151	0.151	0.151	0.150	
Value for money ratio μ	2.94	2.94	2.92	2.92	2.92	2.90	2.92
Exchequer costs Δ_{EC} *(£m)*	150	180	190	190	190	220	187
Additional R&D Δ_R *(£m)*	440	528	555	555	555	639	545
Panel D. Large company deduction case							
Tax-adjusted user cost of R&D ρ	0.179	0.179	0.177	0.177	0.177	0.179	
Value for money ratio μ	1.54	1.54	1.50	1.50	1.50	1.46	1.50
Exchequer costs Δ_{EC} *(£m)*	480	550	730	670	750	780	660
Additional R&D Δ_R *(£m)*	741	849	1,095	1,005	1,125	1,139	992
Panel E: Aggregates							
Total Exchequer costs Δ_{EC} *(£m)*	680	790	1,000	990	1,100	1,210	962
Total additional R&D Δ_R *(£m)*	1,391	1,629	1,969	2,065	2,302	2,540	1,982
Value for money ratio $\mu = \Delta_R/\Delta_{EC}$	2.04	2.06	1.97	2.09	2.09	2.10	2.06
Total qualifying R&D *(£m)*	7,670	8,880	10,800	9,730	10,870	11,840	9,965
Fall of aggregate R&D without policy	18%	18%	18%	21%	21%	21%	20%

Note: Tax-adjusted user cost of R&D and value for money ratio are calculated using the formulae as described in Appendix A.6 using the above policy parameters. In addition, real interest rate is 5% and depreciation rate is 15%. Tax-adjusted user cost of R&D without any tax relief is calculated to be 0.200. Tax-price elasticity of R&D among SMEs is -3.99 as estimated in sub-section 7.2. Tax-price elasticity of R&D among large companies is -1.09 (i.e., the lower-bound elasticity estimate). Exchequer costs (Panels B-D) and total qualifying R&D (Panel E) come from HMRC national statistics. In Panels B-D, additional R&D is calculated as value for money ratios times Exchequer costs (i.e., $\Delta_R = \mu \times \Delta_{EC}$). In Panel E, total Exchequer costs and total additional R&D are the sums of the corresponding amounts in Panels B-D; value for money ratio is total Exchequer costs over total additional R&D; fall in aggregate R&D without policy if total additional R&D over total qualifying R&D.

Table B.18: Heterogeneous effects of R&D tax relief by external finance dependence

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	**R&D expenditure *(£ '000)* 2009-11 average**			**All patent family count 2009-13 average**		
	Full	High external finance dependence	Low external finance dependence	Full	High external finance dependence	Low external finance dependence
Below-asset-threshold indicator (in 2007)	171.4**	203.6*	70.3	0.100**	0.136**	0.033*
	(72.6)	(105.3)	(55.8)	(0.041)	(0.063)	(0.018)
Below-asset-threshold indicator # RZ index	8.2			0.004		
	(6.2)			(0.003)		
Difference		113.3 (119.1)			0.103 (0.066)	
Dependent variable mean over 2006-08	*75.2*	*111.6*	*40.0*	*0.069*	*0.095*	*0.045*
Discontinuity estimate to baseline mean ratio		*1.82*	*1.76*		*1.43*	*0.73*
Firms	4,503	2,217	2,286	4,503	2,217	2,286

Note: *** significant at 1% level, ** 5% level, * 10% level. OLS estimates based on the RD Design. The running variable is total assets in 2007 with a threshold of €86m. Baseline sample includes firms with total assets in 2007 within €25m below and above the cut-off (i.e. between €61m and €111m). Controls for first order polynomials of the running variable separately for each side of the threshold are included. Robust standard errors are in brackets. Rajan-Zingales (1998) index for industry external finance dependence (i.e., industry-level across-firm average of $\frac{capex-cashflow}{capex}$) is calculated at 3-digit industry level using UK firm data over 2000-2005 (Rajan and Zingales, 1998). Firms in industries with high Rajan-Zingales index are more likely to be financially constrained. High (low) external finance dependence subsample include firms with above (below) median industry Rajan-Zingales index. All right-hand-side variables are fully interacted with industry Rajan-Zingales index in columns (1) and (4).

Table B.19: R&D technology spillovers on R&D and patents

	(1)	(2)	(3)	(4)	(5)	(6)
Specification	First stage, OLS		Reduced form, OLS	IV		
Dependent variable	***spilltechRD* *(£ million)* 2009-11 avg.**	**R&D exp. *(£ million)*, 2009-11 avg.**	**All patent fam. count, 2009-13 avg.**	**R&D exp., *(£ million)*, 2009-11 avg.**	**All patent fam. count, 2009-13 avg.**	**All patent fam. count, 2009-13 avg.**
spilltechE (sum tech. proximity x indicator)	11.18***	0.053	0.174**			
	(2.20)	(0.089)	(0.074)			
Below-asset-threshold indicator (in 2007)	0.40	0.156***	0.070**	0.154**	0.063*	
	(1.28)	(0.060)	(0.029)	(0.060)	(0.037)	
spilltechR (sum tech. proximity x £m)				0.005	0.016*	0.014
				(0.008)	(0.008)	(0.011)
R&D expenditure (*£m)*, 2009-11 average						0.412
						(1.959)
Dependent variable mean over 2006-08	*25.02*	*0.070*	*0.061*	*0.070*	*0.061*	*0.061*
Firms	8,818	8,818	8,818	8,818	8,818	8,818

Note: *** Significant at 1% level, ** 5% level, * 10% level. Sample of firms with total assets in 2007 between €51m and €121m. Standard errors in brackets are corrected using 1,000 bootstrap replications over firms. Controls include second order polynomials of total assets in 2007, separately for each side of the asset threshold of €86m; $F_j(Z_{2007}) = \sum_{i,i\neq j} \omega_{ij} f(z_{i,2007})$ where $f(z_{i,2007})$'s are second order polynomials of spillover-generating firm i's total assets in 2007, also separately for each side of the asset threshold (see Appendix C.2); and $techconnect_j = \sum_{i,i\neq j} \omega_{ij}$ – a measure for spillover-generating firm j's level of connectivity in technology space. In column (5), adjusted first-stage F-statistic is 26.9; and the p-value of Anderson-Rubin weak-instrument-robust inference test is 0.018, indicating that the IV estimates are statistically different from zero even in the possible case of weak IV. In column (6), the instrument variable for $spilltechR$ is $spilltechE$ and instrument variable for R&D expenditure is below-asset-threshold indicator.

Table B.20: Descriptive statistics

Panel A. Full CT600 dataset

	Unit	2006	2007	2008	2009	2010	2011	2006-2011
No. firms	*Firm*	1,406,696	1,487,173	1,484,311	1,504,927	1,564,871	1,646,641	2,495,944
No. firms claiming R&D relief	*Firm*	6,431	7,429	8,334	9,144	10,150	12,003	20,730
SME Scheme								
No. firms claiming	*Firm*	5,153	5,855	6,570	7,354	8,238	9,921	20,205
Avg. qual. R&D expenditure	*£ (nom)*	257,752	268,904	266,730	244,854	263,811	258,541	1,569,728
Avg. est. Exchequer costs	*£ (nom)*	39,433	42,150	41,018	44,099	43,138	43,451	169,643
Large Company Scheme								
No. firms claiming	*Firm*	1,290	1,592	1,776	1,795	1,923	2,092	4,048
Avg. qual. R&D expenditure	*£ (nom)*	4,926,939	4,616,811	5,120,979	4,435,308	4,508,202	4,357,442	12,580,710
Avg. est. Exchequer costs	*£ (nom)*	371,097	346,616	412,088	376,405	382,284	357,870	1,030,878
SME subcontractors								
No. firms claiming	*Firm*	399	443	522	610	720	715	2,100
Avg. qual. R&D expenditure	*£ (nom)*	630,098	465,590	406,302	504,624	658,942	928,208	1,007,468
Avg. est. Exchequer costs	*£ (nom)*	47,406	48,014	43,043	42,618	46,771	56,809	315,560
Patenting								
No. firms having patents	*Firm*	3,093	3,085	2,965	2,806	2,682	2,662	9,420
Avg. number of patents	*Patent*	2.68	2.77	2.72	2.63	2.66	2.64	4.93
No. firms having EPO patents	*Firm*	1,453	1,448	1,376	1,409	1,358	1,125	4.770
Avg. number of EPO patents	*Patent*	0.95	0.90	0.82	0.83	0.47	0.17	4.95
No. firms having UK patents	*Firm*	3,262	3,316	3,228	3,083	2,989	2,965	8,986
Avg. number of UK patents	*Patent*	3.00	3.08	3.00	2.83	2.78	2.82	6.13

Panel B. Full FAME dataset

	Unit	2006	2007	2008	2009	2010	2011	2006-2011
No. firms	*Firm*	1,780,531	1,858,209	1,870,089	1,898,721	1,973,722	2,073,930	3,140,060
Variable coverage								
No. firms with total assets	*Firm*	1,732,169	1,807,743	1,818,448	1,843,896	1,914,848	2,015,058	3,012,397
Total assets coverage	*%*	97.3%	97.3%	97.2%	97.1%	97.0%	97.2%	95.9%
No. firms with sales	*Firm*	352,680	319,726	275,938	274,768	263,394	227,463	626,025
Sales coverage	*%*	19.8%	17.2%	14.8%	14.5%	13.3%	11.0%	19.9%
No. firms with employment	*Firm*	95,615	93,855	91,375	94,332	98,426	97,814	164,849
Employment coverage	*%*	5.4%	5.1%	4.9%	5.0%	5.0%	4.7%	5.2%

Panel C. CT600 and FAME matching

	Unit	2006	2007	2008	2009	2010	2011	2006-2011
No. CT600 firms that appear in FAME over 2006-11	*Firm*	1,353,844	1,427,132	1,442,619	1,468,000	1,529,317	1,598,012	2,358,948
As % CT600 firms	*%*	96.2%	96.0%	97.2%	97.5%	97.7%	97.0%	94.5%
Out of which								
No. firms claiming tax relief	*Firm*	6,411	7,409	8,298	9,105	10,108	11,937	20,627
As % CT600 R&D firms	*%*	99.7%	99.7%	99.6%	99.6%	99.6%	99.5%	99.5%
No. firms having patents	*Firm*	3,078	3,065	2,951	2,789	2,665	2,634	9,376
As % CT600 patenting firms	*%*	99.5%	99.4%	99.5%	99.4%	99.4%	98.9%	99.5%

Note: Average qualifying R&D expenditure and estimated Exchequer costs are computed for firms with R&D tax relief claims in the corresponding year or period. Average patents, EPO patents, and UK patents are computed for firms with corresponding patent applications in corresponding year or period.

Appendix C

Appendices to Chapter 3: Privatization and Productivity in Upstream Industries

Figure C.1: Industry-level TFP decomposition (2001-2008)

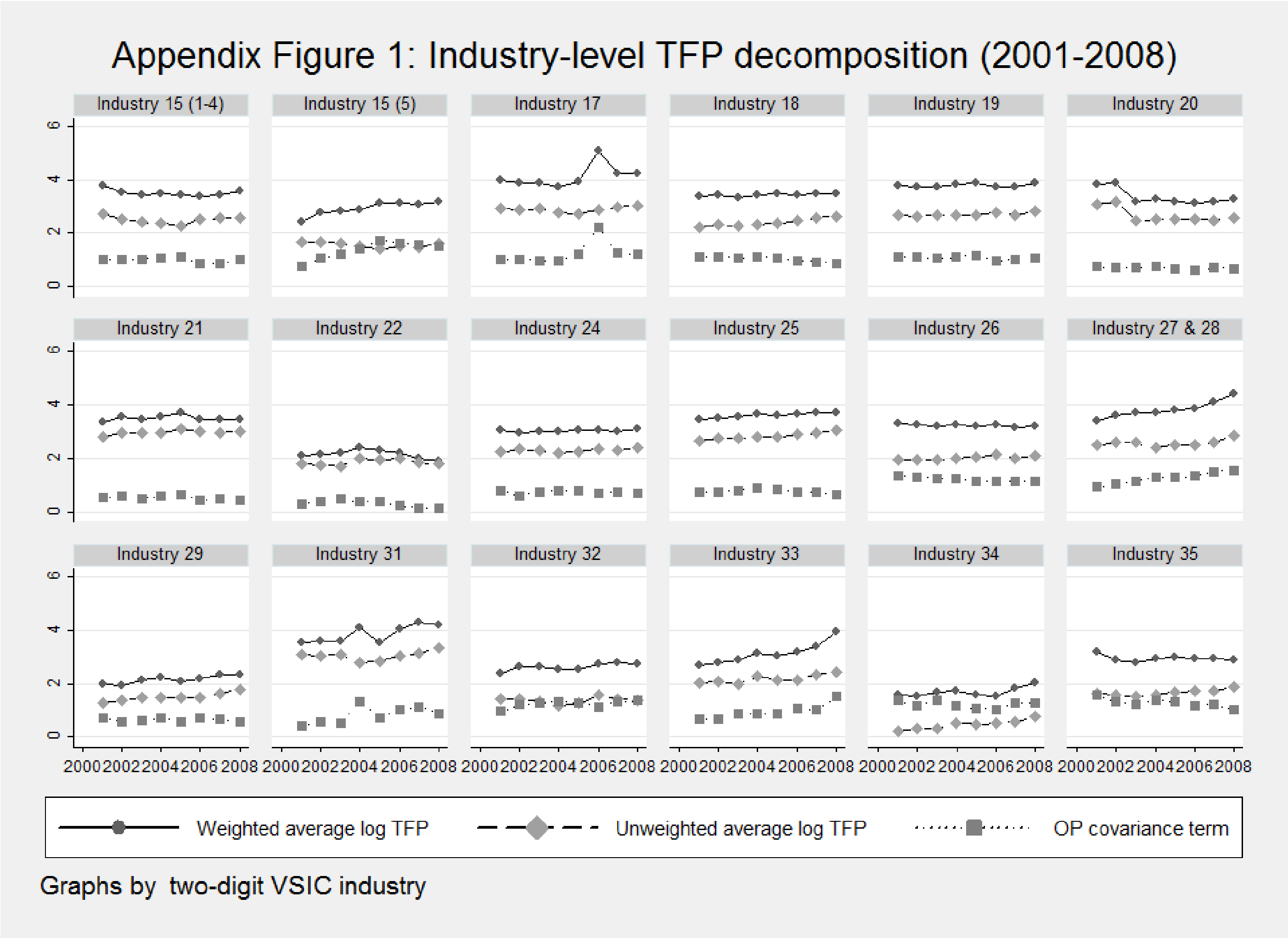


Table C.1: Firm ownership classification and private firm definitions

Panel A: Firm ownership classification from Vietnam Enterprise Census questionnaire

Firm type	Description
1	Enterprise owned by Central Government
2	Enterprise owned by Local Government
3	Limited liability company owned by Central Government
4	Limited liability company owned by Local Government
5	Joint-stock limited liability company with more than 50% state capital
6	Collective
7	Private enterprise
8	Partnership
9	Private limited liability with at most 50% state capital *(Firms are also asked to indicate if the state has controlling share in the firm.)*
10	Joint-stock company without state capital
11	Joint-stock company with at most 50% state capital
12	Enterprise with100% foreign capital
13	Joint venture between state and foreign enterprises
14	Joint venture between non-state and foreign enterprises

Panel B: Private firm definitions used in paper

Private firm type	Definition
Private firm	At least 50% domestic private ownership or non-zero foreign ownership *(i.e. firm ownership types 6-14)*
Non-zero private-share firm	Non-zero domestic private ownership or non-zero foreign ownership *(i.e. firm ownership types 5-14)*
Non-state-controlled firm	Not controlled by the state *(i.e. firm ownership types 6-14, excluding type-9 firms in which the state has controlling share and type-13 firms in which the state has more than 50% share)*
Purely-domestic private firm	At least 50% domestic private ownership and no foreign ownership *(i.e. firm ownership types 6-11)*
Foreign-affiliated firm	Non-zero foreign ownership *(i.e. firm ownership types 12-14)*
Domestic-controlled private firm	Controlled by a domestic private entity *(i.e. firm ownership types 6-11, excluding type-9 firms in which the state has controlling share)*
Foreign-controlled private firm	Controlled by a foreign entity *(i.e. firm ownership types 12-14, excluding type-13 firms in which the state has more than 50% share)*

Table C.2: OLS regression at firm level with firm fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Dependent variable	*Log TFP*									*Log TFP (L)*	*Log TFP (M)*
Sample	Private firms	State-owned firms	Firms that got privatized	Private firms	Always private firms	Never private firms	Non-zero private-share firms	Non-state-controlled firms	Private firms	Private firms	Private firms
Same-industry privatization	0.309	-1.214***	-0.869***	0.234	0.247	-1.379***	0.308	0.0623	0.672***	0.341	-0.268
	(0.269)	(0.334)	(0.278)	(0.284)	(0.291)	(0.360)	(0.191)	(0.261)	(0.243)	(0.295)	(0.188)
Downstream privatization	0.306	0.0623	0.00963	0.329	0.343	0.0346	-0.0651	0.217	0.232	0.303	0.268
	(0.249)	(0.239)	(0.266)	(0.249)	(0.254)	(0.252)	(0.216)	(0.268)	(0.237)	(0.266)	(0.186)
Upstream privatization	-0.0874	-0.158	-0.302	0.0230	0.0333	0.439	0.192	-0.0611	0.0137	-0.116	0.255
	(0.414)	(0.471)	(0.391)	(0.409)	(0.416)	(0.531)	(0.379)	(0.441)	(0.408)	(0.414)	(0.380)
Private			0.0437								
			(0.0299)								
Log demand				0.310**	0.313**	0.0244	0.328**	0.306**	0.296**	0.405***	0.101
				(0.120)	(0.123)	(0.128)	(0.128)	(0.122)	(0.125)	(0.116)	(0.103)
HHI				-0.664*	-0.666*	0.607	-0.734**	-0.555	-1.096**	-0.756*	0.149
				(0.395)	(0.399)	(1.092)	(0.344)	(0.416)	(0.479)	(0.448)	(0.360)
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes	Yes
Sector fixed effects									Yes		
Province fixed effects									Yes		
Year trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	135,511	8,354	6,042	135,511	132,226	5,150	137,223	134,963	135,511	133,786	109,607
R-squared	0.745	0.851	0.758	0.745	0.744	0.853	0.746	0.745	0.187	0.786	0.609

Notes: Private firms in colums (1)-(6) and (9)-(11) are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. Private firms in column (7) are defined as those with non-zero domestic private ownership or non-zero foreign ownership. Private firms in column (8) are defined as those controlled by domestic private or foreign entities. Same-industry, downstream, and upstream privatization measures in each column are calculated for each industry x year using the corresponding defintion of private firms. Log TFP in column (10) is estimated using employment count instead of wage bill. Log TFP in column (11) is estimated using wage bill and imputed material costs. Standard error in parentheses are clustered at industry x province level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

Table C.3: Dynamic panel regressions using TFP estimated with material costs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable					*Log TFP (M)*				
Sample	Private firms	State-owned firms	Firms that got privatized	Private firms	Private firms	Always private firms	Never private firms	Non-zero private- share firms	Non-state-controlled firms
Lagged dependent variable	0.283***	0.137*	0.253***	0.269***	0.269***	0.268***	0.0467	0.262***	0.267***
	(0.0238)	(0.0710)	(0.0617)	(0.0240)	(0.0241)	(0.0239)	(0.0833)	(0.0243)	(0.0241)
Same-industry privatization	1.038***	-1.486***	-1.065**	0.154	0.183	0.165	-1.578***	-0.0464	0.114
	(0.394)	(0.442)	(0.506)	(0.218)	(0.205)	(0.222)	(0.510)	(0.166)	(0.210)
Downstream privatization	0.395***	-0.0372	0.462**	0.497***	0.719***	0.500***	-0.166	0.0580	0.443***
	(0.113)	(0.112)	(0.211)	(0.157)	(0.175)	(0.160)	(0.134)	(0.130)	(0.145)
Upstream privatization	0.308**	1.303***	1.997***	0.360**	0.390***	0.270*	0.956**	0.565***	0.401**
	(0.138)	(0.453)	(0.464)	(0.148)	(0.147)	(0.143)	(0.478)	(0.149)	(0.167)
Private			-0.0214						
			(0.0434)						
Log demand				0.0723	0.106	0.0645	-0.343***	0.0885	0.0613
				(0.0928)	(0.0946)	(0.0943)	(0.115)	(0.100)	(0.0905)
HHI				0.304	0.0859	0.294	1.466*	0.187	0.445
				(0.405)	(0.396)	(0.401)	(0.888)	(0.407)	(0.459)
Downstream HHI					-0.991***				
					(0.278)				
Upstream HHI					-0.378				
					(0.308)				
Year trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	53,896	4,597	4,214	53,896	53,896	51,516	2,763	53,896	53,555
Number of firms	16,783	1,339	869	16,783	16,783	16,046	730	16,783	16,730

Notes: Log TFP (M) is estimated using wage bill and imputed material costs. Private firms in colums (1)-(7) are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. Private firms in column (8) are defined as those with non-zero domestic private ownership or non-zero foreign ownership. Private firms in column (9) are defined as those controlled by domestic private or foreign entities. Same-industry, downstream, and upstream privatization measures in each column are calculated for each industry x year using the corresponding defintion of private firms. Standard error in parentheses are clustered at industry x province level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

Table C.4: Dynamic panel regressions at industry level

	(1)	(2)	(3)	(4)	(5)	(6)
Sample used for aggregation	State-owned firms			Private firms		
Dependent variable	*Weighted average log TFP*	*Unweighted average log TFP*	*Olley-Pakes covariance (TFP, share)*	*Weighted average log TFP*	*Unweighted average log TFP*	*Olley-Pakes covariance (TFP, share)*
Lagged dependent variable	0.216**	0.465***	0.187	-0.264	0.526***	-0.0688
	(0.0998)	(0.134)	(0.159)	(0.203)	(0.143)	(0.138)
Same-industry privatization	-0.825	0.121	-1.019	-2.320*	-0.0791	-0.937
	(0.769)	(0.326)	(0.625)	(1.197)	(0.615)	(1.083)
Downstream privatization	0.589	0.624**	-0.0922	-0.265	0.181	-0.539
	(0.715)	(0.269)	(0.527)	(0.663)	(0.552)	(0.758)
Upstream privatization	-0.142	0.0303	0.0513	0.897	0.713	0.360
	(0.545)	(0.399)	(0.537)	(1.358)	(0.513)	(0.638)
Log demand	0.243	0.181	-0.0416	-0.209	-0.328**	0.00290
	(0.205)	(0.165)	(0.205)	(0.211)	(0.154)	(0.150)
HHI	8.876**	-0.715	9.108***	-0.143	-0.701	3.137
	(3.984)	(1.102)	(3.494)	(2.486)	(2.532)	(4.358)
Year trend	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	108	108	108	108	108	108
Number of industries	18	18	18	18	18	18

Notes: Markets are defined based on three large regions (i.e. North, Central, South). Private firms are defined as those with at least 50% domestic private ownership or non-zero foreign ownership. All privatization measures are calculated for each industry x year using this definition of private firms. Aggregate TFP measures are calculated for each industry x year from log TFPs and revenue shares of firms in the relevant sample. Standard error in parentheses are clustered at industry level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

Table C.5: Spillover on incumbent vs. entrant firms (Industry-level regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Sample used for aggregation	Incumbent private firms			Entrant private firms		
Dependent variable	*Weighted average log TFP*	*Unweighted average log TFP*	*Olley-Pakes covariance (TFP, share)*	*Weighted average log TFP*	*Unweighted average log TFP*	*Olley-Pakes covariance (TFP, share)*
Lagged dependent variable	0.191**	0.387***	0.208	-0.410**	-0.179	-0.0341
	(0.0900)	(0.120)	(0.160)	(0.169)	(0.247)	(0.159)
Same-industry privatization	-0.769	-0.169	-0.724	-1.460	0.809	-3.602**
	(0.771)	(0.397)	(0.623)	(2.106)	(0.914)	(1.702)
Downstream privatization	0.547	0.416	0.00491	2.117	1.592**	-1.264
	(0.736)	(0.325)	(0.477)	(1.721)	(0.649)	(1.406)
Upstream privatization	-0.147	-0.0938	0.417	1.383	0.261	2.908*
	(0.560)	(0.393)	(0.558)	(1.381)	(0.955)	(1.628)
Log demand	0.231	0.159	-0.0156	1.625*	0.951**	0.732
	(0.199)	(0.148)	(0.199)	(0.873)	(0.387)	(0.928)
HHI	8.743**	-0.619	9.331***	-4.170	-1.247	-10.42**
	(4.229)	(1.101)	(3.235)	(4.796)	(2.305)	(4.862)
Year trend	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	108	108	108	108	108	108
Number of industries	18	18	18	18	18	18

	(7)	(8)	(9)	(10)	(11)
Sample used for aggregation	Incumbent & entrant private firms				
Dependent variable	*OP cov. (W.Avg., share)*[a]	*OP cov. (Uw.Avg., share)*[b]	*Difference in W.Avg. log TFP (Inc. – Ent.)*	*Difference in Uw.Avg. log TFP (Inc. – Ent.)*	*Difference in revenue share (Inc. – Ent.)*
Lagged dependent variable	-0.284	-0.264			
	(0.187)	(0.254)			
Same-industry privatization	-0.0544	-0.415	-0.0742	-1.118	0.0110
	(0.998)	(0.386)	(1.996)	(0.954)	(0.0850)
Downstream privatization	0.0111	-0.695***	-0.470	-1.129***	-0.0925
	(0.558)	(0.238)	(1.143)	(0.345)	(0.0732)
Upstream privatization	-1.425**	-0.304	-2.039	-0.853	-0.210
	(0.618)	(0.351)	(1.432)	(0.823)	(0.189)
Log demand	-0.567*	-0.257**	-1.051	-0.402*	-0.115**
	(0.342)	(0.124)	(0.668)	(0.229)	(0.0490)
HHI	7.163**	-0.656	11.58*	-0.670	0.346
	(3.030)	(1.699)	(7.040)	(2.273)	(0.427)
Year trend	Yes	Yes	Yes	Yes	Yes
Number of observations	108	108	108	108	108
Number of industries	18	18	18	18	18

Notes: Markets are defined based on three large regions (i.e. North, Central, South). Private firms are defined as those with at least 50% domestic private ownership or those with non-zero foreign ownership. All privatization measures are calculated for each industry x year using this definition of private firms. Aggregate TFP measures are calculated for each industry x year from log TFPs and revenue shares of firms in the relevant sample. Standard error in parentheses are clustered at industry level. ***, **, and * denotes statistical significance at 1%, 5%, and 10% respectively.

[a] OP cov. (W.Avg., share) = (Incumbents' revenue share – Entrants' revenue share) * (Incumbents' weighted average log TFP – Entrants' weighted average log TFP)

[b] OP cov. (Uw.Avg., share) = (Incumbents' revenue share – Entrants' revenue share) * (Incumbents' unweighted average log TFP – Entrants' unweighted average log TFP)