ESSAYS ON THE LABOR MARKET, HUMAN CAPITAL, AND ECONOMIC GROWTH

by

Jingnan Liu

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Economics)

at the

UNIVERSITY OF WISCONSIN–MADISON

2024

Date of final oral examination: 05/08/2024

The dissertation is approved by the following members of the Final Oral Committee:
John Kennan, Professor, Economics
Rasmus Lentz, Professor, Economics
Dean Corbae, Professor, Economics and Finance
Randall Wright, Professor, Economics and Finance

To the moment.

ACKNOWLEDGMENTS

As I approach the culmination of my Ph.D. journey, I find myself reflecting on the incredible support and guidance I have received from so many individuals. I am deeply grateful for the inspiration, encouragement, and knowledge that have been generously shared with me. Below, I extend my heartfelt thanks to some of the extraordinary individuals whose contributions have been invaluable.

First and foremost, I extend my deepest thanks to my co-advisers, John Kennan and Rasmus Lentz, who have been pivotal in shaping my research journey. Their approaches to research have profoundly influenced my own research philosophy and aspirations.

John has instilled in me the mindset of a serious economist, always encouraging me to pursue research that will stand the test of time. He motivated me to explore the questions I am passionate about and to be patient with myself in the process. Through all the ups and downs, I found calm and wisdom in his guidance. His humor, relaxed demeanor, and deep love for research have made my academic journey not only enriching but also enjoyable.

Rasmus encouraged me to pursue fundamental research and strive for true innovation. He pushed me to think big while ensuring I paid attention to the finer details. His reminders to enjoy the process and to be a happy warrior have been a source of inspiration and motivation throughout my journey. His strong empathy and understanding have created a safe space for me to be vulnerable when faced with challenges, making me feel supported in my academic pursuits.

I would also like to thank my committee members, Dean Corbae and Randall Wright, for their generous contributions of time, expertise, and invaluable feedback, all of which have been crucial in refining and enhancing this project.

I am also grateful to many other faculties who have given me substantive comments on my research. In particular, thank you to Naoki Aizawa, Simeon Alder, Carter Braxton, Chao Fu, Rishabh Kirpalani, Paolo Martellini, Jeffrey Smith, Christopher Taber, and Matthew Wiswall.

My gratitude extends to my amazing coauthors: Xiaodong Fan, Chao Fu, Martin Ganco, Chao He, Rosemary Kaiser, Andrew Smith, Christopher Taber, Haifeng Wang, Matthew Wiswall, and Shotaro Yamaguchi. I have learned immensely from each of them, benefiting from their unique insights, rigorous scholarship, and unwavering support throughout my academic journey.

I would like to express my gratitude to the following organizations for their support: The Department of Economics for providing financial support through fellowships and travel funding, the U.S. Census Bureau for access to data used in the first chapter, Payscale.com for the data utilized in the second chapter, and the Korean Educational Development Institute for the data employed in the third chapter. I also extend my thanks to the support staff of the Economics Department, with special appreciation to Julie Anderson, Becca George, and Kim Grocholski. Additionally, Florence Honore and Robert Thomas have been invaluable resources in working with restricted-use data.

The graduate students in the Economics Department are some of the best people I've ever known, and I will miss interacting with them. Thanks to Anson Zhou for being my mentor from day one, Mark Rempel for introducing me to the colorful world of economic growth, Anna Lukianova for her hundreds of hugs, and Sharada Dharmasankar for being my writing coach and cheering me up during tough times. I can't thank my classmates enough for their constant support, especially Zaure Aitkulova, Jason Choi, Long Hong, Elise Marifian, Lois Miller, and Minseon Park. They have always been there whenever I had questions, and I feel so lucky to have shared the joys and tears with them.

I am also grateful to members of the "Labor, Firms, and Macro" Reading Group. The two-hour meetings every Saturday have helped me learn frontier research, sharpen my intuition, and improve my presentation skills. A special thanks to Xincheng Qiu, who is not only a dedicated organizer but also a role model and a kind, generous friend.

I am grateful to those who have shared their stories and insights about the real world with me. These narratives have greatly enhanced my intuition and common sense, both of which are invaluable for my research. A special acknowledgment goes to Qi Chen, who has taught me how to stay connected to real-world issues and navigate their complexities. I am also thankful to my friends Jingying He, Xiaomao Li, Ling Qin, and Fengyuan Zhu for their shared knowledge and perspectives, which have enriched my understanding and informed my work.

Finally, and most importantly, I am fortunate to have incredible support from my parents. Mom and Dad, thank you for always being there for me, for making me feel like my most authentic self, and for supporting all my decisions by every means - you are truly remarkable.

ABSTRACT

The first chapter studies how endogenous worker mobility affects inter-firm knowledge diffusion, innovation, and economic growth. I propose a framework combining endogenous growth and on-the-job search. Firms grow knowledge by in-house innovation and by hiring workers from more productive firms. Knowledge is nonrival, leading to underinvestment in innovation. Non-compete contracts address this underinvestment by allowing innovating firms to enforce buyout payments when they lose workers. However, they discourage diffusion by deterring firm entry. Linking patent records to matched employer-employee administrative data at the U.S. Census Bureau, I document that inventors diffuse knowledge across firms and are compensated for knowledge diffusion. Constructing novel microdata, I find non-compete contracts are associated with increased innovation expenditure and decreased worker mobility. I calibrate my theoretical model to match the empirical results. Knowledge diffusion, through the channel of worker mobility, accounts for 4% of the aggregate growth rate and 9% of welfare. Optimal regulation of non-compete contracts balances the innovation-diffusion tradeoff.

The second chapter (joint with Martin Ganco, Haifeng Wang and Shotaro Yamaguchi) studies the strategic use of non-compete agreements. Extant work in strategic management has focused on the role of noncompete agreements (NCAs) – a form of restrictive legal lever used by firms when managing human capital – and conceptualized them as being advantageous to firms. Challenging this notion, we highlight a novel downside of using NCAs and show how their use by some firms creates differentiation opportunities for rival firms. We analyze a unique survey dataset to examine the heterogeneity in the firms' actual use of NCAs conditional on industry and state. We find that the nonuse of NCAs is more common among firms that rely more heavily on talent and are also not the industry leaders, and such firms are more likely not to use NCAs with the goal of attracting skilled employees.

The third chapter develops a structural model of pre-college educational investment in college admission tournaments. Students are heterogeneous in ability, family wealth, and preferences for colleges and can purchase tutoring services to improve their human capital and test scores. They also face borrowing constraints. The score distribution, admission thresholds, and college assignment are joint equilibrium outcomes. The model is estimated with Korean ELS: 2005 data and can be used to study Korea's tutoring market with a wide range of policy candidates, including taxing private tutoring and reducing noise in admission. A tax lowers the overall spending on tutoring. The students from middle-income families are most responsive to the price change. Reduced signal noise incentivizes the tutoring expenditure of high-ability students and improves their chances of attending prestigious colleges.

1.1 Introduction

Worker mobility diffuses knowledge across firms. Numerous inventors, technicians, and managers change jobs, carrying technical and managerial knowledge from former employers to new ones. Semiconductor company AMD, for instance, grew into an industry leader when a former Fairchild executive onboarded engineers and managers from his previous company. Such cases highlight how knowledge flows through worker flows can improve firm productivity and foster economic growth.[1]

Since Arrow (1962), worker mobility has been recognized as a channel of inter-firm knowledge diffusion.[2] However, theory and quantitative assessments of this channel have been limited. Diffusion-based growth models have predominantly abstracted from channels of knowledge diffusion (Kortum, 1997; Luttmer, 2007; Lucas Jr and Moll, 2014; Perla and Tonetti, 2014; Benhabib, Perla, and Tonetti, 2021). Empirical work remains challenging, both because worker mobility is endogenous and because measuring knowledge diffusion is difficult.[3] Understanding the process of knowledge diffusion is central to key questions in economics, including the sources of economic growth, the implications of worker mobility, and the design of labor and industrial policies. Regulation of non-compete contracts stands out as a particularly relevant policy.

This paper studies how endogenous worker mobility affects inter-firm knowledge diffusion, innovation, and economic growth. The contributions are fourfold. First, I develop a tractable growth model incorporating endogenous worker mobility as a channel of inter-firm knowledge diffusion. Second, I provide tangible measures of inter-firm knowledge diffusion, linking administrative data on patents (United States Patent and Trademark Office), firm performance (Longitudinal Business Database), and employment history and wages (Longitudinal Employer-Household Dynamics) from the U.S. Census Bureau. Third, I propose a theory of non-compete contracts and construct rich data on non-compete contracts and employment history for executives in publicly listed U.S. firms. Lastly, I quantify the importance of the worker-mobility channel of knowledge diffusion in aggregate growth and study the optimal enforceability of non-compete contracts.

The theory introduces on-the-job search to an endogenous growth model of innovation and knowledge diffusion. A novel feature is that firms can adopt knowledge by hiring workers from

[1]According to an Endeavor Insight report from 2014, of the over 130 Bay Area tech companies listed on the NASDAQ or the New York Stock Exchange, "70 percent of these firms can be traced directly back to the founders and employees of Fairchild. The 92 public companies that can be traced to Fairchild are now worth about $2.1 trillion, which is more than the annual GDP of Canada, India, or Spain."

[2]Arrow (1962) observes that "no amount of legal protection can make a throughly appropriable commodity of something so intangible as information" and adds that "mobility of personnel among firms provides a way of spreading information."

[3]As Krugman (1991) has pointed out, "knowledge flows are invisible, they leave no paper trail by which they may be measured and tracked."

more productive firms. Firms are heterogeneous in production-related knowledge. They grow knowledge by choosing innovation intensity and by hiring workers. They also choose whether to enter and when to exit the labor market. Firms meet employed workers at random in a frictional labor market. Workers move if the surplus from trade is positive – if the worker is more valuable to the destination firm than to the origin firm. Mobility diffuses knowledge from origin to destination firms. Knowledge diffusion, together with innovation, generates aggregate growth and shapes the evolution of knowledge distribution.

A worker is a vessel for knowledge transfer. Knowledge is ***nonrival***: designs, blueprints, and production processes can be used by multiple firms simultaneously. Therefore, origin firms that lose workers do not lose knowledge, whilst destination firms that hire workers gain a transfer of knowledge. Knowledge transfer generates a surplus. The ***allocation of surplus*** shapes firms' incentives to innovate and hire workers. If the entire surplus is allocated to the origin firms, there are no incentives to enter the labor market and hire workers. Knowledge diffusion vanishes. On the other hand, if all surplus is allocated to the destination firms, innovation incentives are dampened. Firms, in this scenario, fail to internalize how their innovation decisions will improve the knowledge transfer to future employers. Besides, they may free-ride on the innovation of others, which could further stifle innovation.

Non-compete contracts can impact innovation and hiring decisions by affecting the allocation of surplus. While non-compete contracts may appear restrictive – precluding employees from moving to a competitor after leaving their employer – in practice, they often come with buyout provisions. Under this arrangement, the future employer can pay a fee demanded by the employer in exchange for a release from the non-compete restriction. I encompass these real-world features in modeling non-compete contracts. A firm and a worker enter an employment contract that includes (i) a non-compete clause restricting the worker's outside employment and (ii) a buyout payment for the worker to be released from the restrictive clause. The firm and its worker, acting as a coalition, optimally charge buyout payments from future employers. On the one hand, non-compete buyouts encourage innovation because firms will be compensated for losing workers. On the other hand, the buyouts discourage diffusion, as entering the labor market and hiring workers get more costly. Non-compete contracts generate the ***innovation-diffusion trade-offs***.

The state of the economy is summarized by the distribution of knowledge across firms. A firm's innovation and hiring decisions depend on this distribution because the knowledge levels of others determine its own returns from knowledge transfer. Firms' innovation and hiring decisions, in turn, determine how the distribution evolves. Accordingly, one of the equilibrium conditions is that firms solve Hamilton-Jacobi-Bellman equations, taking knowledge distribution as given. Another is that the distribution evolves according to Kolmogorov forward equations given the decision rules of individual firms. I focus on a particular class of solutions to these equations: balanced growth paths (BGPs). Along the BGP, aggregate knowledge grows at a constant rate, and the distribution of relative knowledge levels remains stationary.

The aggregate growth rate is the sum of three components: (i) innovation rate, (ii) diffusion rate through worker mobility, and (iii) growth from firm entry and exit. I show that by introducing a small perturbation to the random search technology – firms randomly meet workers from all firms that are more productive than they are – the BGP equilibrium can be solved analytically and admits a theoretical decomposition of the aggregate growth rate. The unique stationary distribution of knowledge is a Pareto distribution, whose shape parameter is determined by the rate of hiring workers relative to the rate of innovation. Low search frictions compress the cross-sectional distribution of knowledge because low-productivity firms hire and catch up quickly. Low innovation costs spread the distribution because high-productivity firms can easily innovate and escape from the pack. The rates of hiring and innovation jointly determine the knowledge dynamics.

The decentralized equilibrium features inefficiency because of innovation and search externalities. The innovation externality stems from knowledge diffusion. When workers move, they diffuse knowledge from the origin to the destination firms, creating a surplus from diffusion. However, in the presence of search frictions, this surplus is only partially appropriated by origin firms. This creates a wedge between private returns and social returns to innovation. Innovation is underinvested. The search externality encapsulates congestion and market thickness externalities arising in the random search environments (Hosios, 1990). When entering the labor market, firms do not internalize their negative impact on other firms' probability of hiring a worker (congestion externality) and their positive impact on workers to find a job (thick market externality). Absent knowledge diffusion, efficient allocation is restored if the Hosios condition holds, namely, if destination firms' surplus share equals the matching elasticity. With knowledge diffusion, however, the equilibrium is no longer efficient under the Hosios condition, and origin firms must be compensated for innovation externality.

I apply the theory to the matched employer-employee datasets. Measuring how worker mobility diffuses knowledge across firms is challenging. Both because it requires data that tracks workers across firms and traces the associated knowledge flows, and because "knowledge flows are invisible, they leave no paper trail by which they may be measured and tracked" (Krugman, 1991). To overcome this challenge, I construct comprehensive data on worker employment history and earnings, patent records, and firm-level measures by linking the Longitudinal Employer-Household Dynamics (LEHD), Longitudinal Business Database (LBD), and United States Patent and Trademark Office (USPTO) PatentsView Database at the U.S. Census Bureau. The data offers a unique opportunity to identify worker mobility directly and observe (i) inter-firm knowledge flows, (ii) firm productivity, and (iii) worker earnings.

With the data, I offer tangible measures of inter-firm knowledge diffusion through mobile workers. Specifically, for each pair of origin and destination firms associated with worker mobility, I count how many post-mobility new patents of the destination firm cite the pre-mobility patent stock of the origin firm. Each citation is treated as one instance of the destination firm drawing upon the knowledge stock of the origin firm. I view patents as a measurable form of knowledge and citations as direct evidence of knowledge diffusion. I further restrict my analysis to mobile inventors because inventors are directly exposed to their employer's technical knowledge, and that knowledge can be measured with patents. Compared with previous studies, where inventor mobility is typically inferred from patent trajectories, the administrative data allows me to track the entire employment history of inventors and precisely identify mobility.

Using an event study approach, I examine the impact of inventor mobility on knowledge diffusion, firm productivity, and inventor compensation. I document four new findings. First, after hiring an inventor from a more productive firm, the destination firm more intensively draws upon the knowledge (cite the patent stock) of the origin firm. Second, after hiring an inventor from a more productive firm, the destination firm experiences 5% growth in annual productivity, as measured by total factor productivity of revenue (TFPR). Third, the inventor labor market is mobile, and nearly half (49%) of inventor mobility occurs from more to less productive firms. Fourth, inventors who move down the firm productivity ladder are compensated for knowledge diffusion, experiencing 4% growth in quarterly earnings. This set of evidence collectively suggests that inventors diffuse knowledge across firms and are compensated for knowledge diffusion.

I next delve into the innovation-diffusion trade-offs of non-compete contracts. For this analysis, I assemble a unique dataset by extracting non-compete contracts from U.S. Securities and Exchange Commission (SEC) filings with machine learning and natural language processing tools. I further link the contracts to matched firm-executive data constructed from Compustat and BoardEx. My sample covers 34,786 executives from 9,255 U.S. publicly traded companies, out of which 65% are bound by non-compete contracts. This micro-level non-compete data allows me to utilize within-firm variation in non-compete use and examine whether changes in non-compete use predict changes in R&D investment and worker mobility. I find that, within a firm, non-compete use is associated with increases in R&D expenditure and decreases in mobility rate. Specifically, shifting from nonuse to all use of non-compete contracts among executives is associated with a 4% rise in a firm's R&D expenditure. For a given executive, signing a non-compete contract is associated with 6 percentage points (pp) decline in mobility rate.

Linking the model to data, I quantify the importance of the worker mobility channel of diffusion and study the optimal regulatory policy of non-compete contracts. I numerically solve for a BGP equilibrium by adapting an algorithm from Achdou, Han, Lasry, Lions, and Moll (2022). I calibrate the model parameters via simulated method of moments (SMM). Key calibration targets include the average change in firm productivity after hiring inventors, firm hiring rate, inventor mobility rate, the R&D expenditure-to-sales ratio, and the TFP growth rate. The calibrated model captures the targeted and non-targeted moments well.

With the calibrated model, I perform two main exercises. The first quantitative exercise is a growth decomposition. I decompose the aggregate TFP growth rate into three additive components: (i) innovation rate, (ii) diffusion rate through worker mobility, and (iii) net growth from firm entry and exit. I find that knowledge diffusion, through the channel of worker mobility, accounts for 61.3% of the TFP growth rate. I complement the exercise where I shut down knowledge diffusion associated with worker mobility. Shutting down the channel leads to a 4% drop in the TFP growth rate and a 9.20% decline in welfare. The welfare decline would come through the declining growth and reduced entry of new firms.

The second quantitative exercise studies the optimal regulation of non-compete contracts. The optimal regulation suggests that, by allocating 40% of the surplus to destination firms and 60% of the surplus to origin firms, welfare can improve by 10.64%.

Related literature. This paper contributes to several strands of literature. First, this paper builds on and adds to the theoretical literature on endogenous growth and knowledge diffusion. Seminal papers include Kortum (1997); Luttmer (2007); Lucas Jr (2009); Lucas Jr and Moll (2014); Perla and Tonetti (2014); König, Lorenz, and Zilibotti (2016); Buera and Oberfield (2020); Hopenhayn and Shi (2020); Benhabib, Perla, and Tonetti (2021). The diffusion-based growth models have predominantly abstracted from channels of diffusion[4] This paper unpacks the "black box" of inter-firm knowledge diffusion and isolates a particular channel: worker mobility. I enrich the framework of Lucas Jr and Moll (2014), Perla and Tonetti (2014) in a tractable way to introduce endogenous worker mobility as a channel to diffuse knowledge. The extra margin not only sheds light on the role of the labor market in aggregate growth but also has novel aggregate implications as worker mobility and aggregate growth are jointly determined.

Second, this paper adds to the empirical literature evaluating the impact of inventor mobility on knowledge diffusion (Jaffe, Trajtenberg, and Henderson, 1993; Almeida and Kogut, 1999; Singh and Agrawal, 2011; Stoyanov and Zubanov, 2012; Kaiser, Kongsted, and Rønde, 2015; Braunerhjelm, Ding, and Thulin, 2020). This literature has been constrained by the limited availability of matched employer-inventor data to identify the inter-firm mobility of inventors. Besides, measuring knowledge diffusion has proved challenging. I bring new data that tracks the full employment history of inventors from the U.S. Census Bureau. I offer a tangible measure of knowledge diffusion using information on patent citations.

Third, the paper adds to the growing literature on non-compete contracts (Balasubramanian, Chang, Sakakibara, Sivadasan, and Starr, 2022; Baslandze, 2022; Gottfries and Jarosch, 2023; Jeffers, 2023; Johnson, Lipsitz, and Pei, 2023; Shi, 2023). Shi (2023) has pioneered a theoretical framework to rationalize the design of noncompete clauses in a labor search framework and studied the welfare effects of regulating these contracts. This paper complements Shi (2023) by integrating non-compete contracts into an endogenous growth model. I focus on the knowledge diffusion aspect

[4] In diffusion-based growth models, agents can increase their productivity by interacting with others, typically described as random draws from an exogenous or endogenous distribution.

and emphasize that nonrivalry generates inefficiency. I contribute to the literature by evaluating the impact of non-compete contracts on knowledge diffusion and economic growth.

The rest of the paper proceeds as follows. Section 1.2 presents the theory, predictions, and efficiency properties. Section 1.3 describes the data and empirical evidence. The section 1.4 quantifies the importance of the worker-mobility channel of diffusion in aggregate growth and characterizes the optimal regulation of non-compete contracts.

1.2 Model

Baseline Model

Environment

Time is continuous and infinite, $t \in [0, \infty)$. Agents are risk-neutral and discount the future at a common rate ρ. Two types of agents populate the economy: an endogenous measure N_t of firms, and a unit measure of workers. Firms are heterogeneous over production-related knowledge and labor market state. Each firm has knowledge Z_t, and can be matched with a worker or be vacant. Each worker has the same knowledge as his or her employer, and will always be employed but search on the job.[5] Knowledge is created through firm innovation and diffused through worker mobility.

Preference. A representative household comprises all workers and owns all firms in the economy. The household derives utility from consumption

$$\int_{t=0}^{\infty} e^{-\rho t} Y_t \, dt,$$

may borrow and lend in the financial market at interest rate r_t. The household's Euler equation gives the equilibrium interest $r_t = \rho$.

Production. Firms produce a homogeneous consumption good in a perfectly competitive market with the price normalized to one. Firms have access to a costless linear production technology. Regardless of labor market state, a firm with knowledge Z_t produces Z_t units of the good and earns Z_t revenue.

Innovation. While matched with a worker, a firm can grow its knowledge by choosing innovation intensity μ_t at cost $\kappa_{r\&d}(\mu_t)Z_t$.[6] The cost function $\kappa_{r\&d}(\cdot)$ is strictly increasing, continuously differentiable, and convex. Innovation intensity governs the speed of knowledge growth:

$$d\log(Z_t) = \mu_t \, dt.$$

[5] I leave out unemployment to focus on inter-firm knowledge diffusion when workers change jobs.

[6] Innovation cost is proportional to firm productivity, reflecting the view that innovation requires labor time at the cost of foregone production.

The worker learns new innovations while on the job and has the same knowledge as his or her employer.

Knowledge diffusion. While vacant, a firm can adopt knowledge by hiring workers from more productive firms. Labor market is frictional. Knowledge diffusion is the outcome of search, matching, and learning between firms and workers.

- **Search.** A vacant firm meets a worker at rate $\lambda(\theta_t)$. The meeting rate $\lambda(\theta_t)$ is determined by the equilibrium market tightness θ_t. Meeting is random: a vacant firm randomly draws a worker from the distribution of matched firms $\mathcal{F}_m(\cdot, t)$. Function $m(1, \theta_t)$ gives the total number of meetings between a unit measure of workers and measure θ_t of vacant firms and has constant return to scale.[7]

- **Matching.** Upon meeting, the vacant firm Bertrand competes with the worker's employer in a sequential auction as in Postel-Vinay and Robin (2002): The vacant firm makes a take-it-or-leave-it offer to the worker; the current employer makes a take-it-or-leave-it counteroffer; the worker decides. An offer specifies the expected wage value that a firm promises to a worker. The worker moves to, or is retained by, the firm that offers a higher promised value.

- **Learning.** When a firm hires a worker from a more productive firm ($Z' > Z$), the firm will catch up to the worker with probability p, or bring the worker to its current level with probability $1-p$. When a firm hires a worker from a less productive firm, the worker will always learn from the firm. Learning is instantaneous. In the end, the firm and worker knowledge will be equal to each other.

Separation. At each instant, a matched pair of firm and worker has the option to separate. Upon separation, a firm becomes vacant, and a worker leaves the labor market. Each leaving worker is replaced by a newborn worker randomly matched with a vacant firm. As a result, the outflow of separated matches is offset by the inflow of newly formed matches with newborn workers.

Exit. At each instant, any firm has the option to exit the economy. Upon exit, a firm produces $\kappa_{\text{scrap}}(t)$ units of good, which I refer to as its scrap value. Scrap value grows as the economy grows. At the reservation knowledge, $\underline{Z}_t$, a firm should be indifferent between continuing to operate and exiting.

Entry. A large pool of potential firms may enter the economy by paying an entry cost $\kappa_{\text{entry}}(t)$. A firm enters vacant and draws initial knowledge from an exogenous distribution $\mathcal{F}_e(\cdot, t)$. Let $\mathcal{V}(Z, t)$ be the value of a vacant firm Z. The measure of entrants is determined by an *ex-ante* free entry

[7] Note that market tightness is defined as the number of vacant firms divided by the number of workers. In this setting, θ_t is equal to the number of vacant firms, because the number of workers is normalized to be one.

condition:

$$\int_{\underline{Z}_t}^{\infty} \mathcal{V}(Z,t)\ d\mathcal{F}_e(Z,t) = \kappa_{\text{entry}}(t). \tag{1.1}$$

Contract

Information and commitment. Information is complete. The payoff-relevant information – firm knowledge, innovation intensity, wage, and outside offer – is perfectly observable. Firms and workers enter long-term contracts. Workers cannot commit. Firms can commit to the expected wage value they have promised to workers. Both workers and firms can costlessly leave the employment relationship for their respective outside options: Workers can quit to another firm upon outside meeting or leave the labor market at will; firms can separate from matches at any time.

Competition over worker. When a vacant firm meets the worker, the competition for the worker occurs in a sequential auction as in Postel-Vinay and Robin (2002). The vacant firm makes a take-it-or-leave-it offer that promises the worker expected wage value (present discounted value of wages). The contract responds by matching the outside offer. As both firms have participation constraints – firms can separate from matches at will – the promised wage value in an offer will not exceed the firm's willingness to pay (the marginal value the worker brings). The worker ends up in the firm with the higher willingness to pay, and receives the wage value equal to the second-highest willingness to pay.

Contract. Consider a firm contracting with a worker at time t_0. A contract specifies state-contingent innovation intensity μ_t and wage w_t, where state includes firm knowledge Z_t and promised wage value $\mathcal{W}_t$ to the worker: [8]

$$\mathcal{C} = \left\{ \mu_t(s_t)\ ,\ w_t(s_t) \right\}_{t=t_0}^{\infty} \quad , \quad \text{where} \quad s_t := \left\{ Z_t\ ,\ \mathcal{W}_t \right\}.$$

Knowledge Z_t evolves through innovation and diffusion. Promised wage value $\mathcal{W}_t$ can evolve when an outside offer arrives. Since the contracts respond to outside offers by matching counteroffers, the promised wage value is determined by the willingness to pay of the best outside offer a worker has previously received.

Firm's Problem

A firm designs a contract $\mathcal{C}$ to maximize the present discounted value of profits subject to a given wage value promised to the worker.

[8] As workers always have the sample knowledge as their employers, worker knowledge can be abstracted from state space.

Promise-keeping. The promised wage value is delivered through (i) a sequence of state-contingent wages $\{w_t(s_t)\}_{t=t_0}^{T}$, and (ii) continuation value upon separation $\mathcal{W}_T$, which equals the firm's willingness to retain the worker $\mathcal{M}(Z_T, T) - \mathcal{V}(Z_T, T)$. The firm has to honor the promise and hence faces a promise-keeping (PK) constraint:

$$E_{T,\{s_t\}_{t=t_0}^{T}}\left[\int_{t_0}^{T} e^{-rt} w_t \, dt + e^{-rT}\left(\mathcal{M}(Z_T, T) - \mathcal{V}(Z_T, T)\right)\right] \geqslant \mathcal{W}_{t_0}. \tag{PK}$$

Firm optimality. The firm earns profits from production and pays innovation costs and wages. The problem of a firm consists of choosing a sequence of innovation intensities and wages $\{\mu_t, w_t\}_{t=t_0}^{T}$ to maximize the profit value:

$$\begin{aligned} \max_{\mathcal{C}} \quad & E_{T,\{s_t\}_{t\geqslant t_0}}\left[\int_{t_0}^{T} e^{-rt}\Big(Z_t - \kappa_{r\&d}(\mu_t) Z_t - w_t\Big) dt + e^{-rT}\mathcal{V}(Z_T, T)\right] \\ \text{s.t.} \quad & E_{T,\{s_t\}_{t\geqslant t_0}}\left[\int_{t_0}^{T} e^{-rt} w_t \, dt + e^{-rT}\left(\mathcal{M}(Z_T, T) - \mathcal{V}(Z_T, T)\right)\right] \geqslant \mathcal{W}_{t_0} \end{aligned} \tag{PK}$$

Private efficiency. The structure of the economy allows us to simplify the firm's problem. With a firm's commitment to promised wage value and risk-neutral preferences, the optimal contract is ***privately efficient***: A contract maximizes the joint value of a firm-worker match. The joint value is the sum of the firm's profit value and the worker's wage value within the match.

A firm's problem can thus be solved in two stages: A first stage in which the firm chooses innovation intensities to maximize the joint value, and a second stage in which the firm sets the wages that deliver the promised value. As both firms and workers are risk-neutral, wage transfers between the firm and its worker leave the joint value unchanged. In what follows, I will focus on the problem of joint value maximization in the first stage.

Value Functions

The investment and allocative decisions – innovation, worker mobility, entry, exit – can be characterized by a set of Hamilton-Jacobi-Bellman equations.

Joint value. A matched firm grows knowledge by choosing innovation intensity. Let $\mathcal{M}(Z, t)$ be the joint value of a firm and a worker in a match. The joint value can be characterized by the Hamilton-Jacobi-Bellman (HJB) equation:

$$r_t \, \mathcal{M}(Z, t) \;\; = \max_{\mu} \; \underbrace{Z}_{\text{Production}} \; - \; \underbrace{\kappa_{r\&d}(\mu)\, Z}_{\text{Innovation Cost}} \; + \; \underbrace{\mu Z \, \partial_z \mathcal{M}(Z, t)}_{\text{Gains from Innovation}} \; + \partial_t \mathcal{M}(Z, t) \tag{1.2}$$

The annuitized value of a firm-worker match (the left-hand side) consists of the flow profit net of innovation cost, gains from innovation, and capital gains from economy-wide growth (the last term).

It is worth noting that worker mobility does not affect the joint value of this match. This is because when the worker leaves, the firm's loss of value is exactly compensated by the worker's wage value at the new employer. As a result, the joint value remains the same. The optimal innovation intensity maximizes joint value and satisfies the first-order condition:

$$\partial_\mu \kappa_{r\&d}(\mu) = \partial_z \mathcal{M}(Z,t). \tag{1.3}$$

Vacant value. A vacant firm grows knowledge by hiring workers in a frictional labor market. Let $\mathcal{V}(Z,t)$ be the value of a vacant firm. The HJB equation for a vacant firm is

$$r_t\, \mathcal{V}(Z,t) = \underbrace{Z}_{\text{Production}} + \underbrace{\lambda(\theta_t)}_{\text{Meeting Rate}} \int \underbrace{\left[\mathcal{S}(Z,Z',t)\right]^+}_{\text{Surplus from Trade}} d\mathcal{F}_m(Z',t) + \partial_t \mathcal{V}(Z,t). \tag{1.4}$$

The firm earns flow profit from production. At rate $\lambda(\theta_t)$, the firm randomly meets a worker sampled from the distribution of matches, $\mathcal{F}_m(Z',t)$. Upon meeting, the firm gains surplus $S(Z,Z',t)$, if it hires the worker successfully. The surplus from trade $S(Z,Z',t)$ follows

$$\mathcal{S}(Z,Z',t) = \begin{cases} \Big[\mathcal{M}(Z,t) - \mathcal{V}(Z,t)\Big] - \Big[\mathcal{M}(Z',t) - \mathcal{V}(Z',t)\Big] & \text{if} \quad Z \geqslant Z' \\ \Big[p\mathcal{M}(Z',t) + (1-p)\mathcal{M}(Z,t) - \mathcal{V}(Z,t)\Big] - \Big[\mathcal{M}(Z',t) - \mathcal{V}(Z',t)\Big] & \text{if} \quad Z < Z' \end{cases}$$

In this expression, the first bracket is the increase in value due to hiring, representing the marginal value of the worker. Workers are valuable because they facilitate innovation and transfer knowledge from their former employers. The second bracket is the wage value the firm pays to the worker. The wage value equals the former employer's willingness to pay because, in Bertrand's competition, a poaching firm offers a wage value that is exactly sufficient to induce a worker to move.

Mobility occurs when there is a positive surplus from trade. Workers move if the value they bring to the destination firm exceeds their value at the origin firm. As knowledge is nonrival, firms losing workers do not lose knowledge, while firms hiring from more productive firms gain knowledge. Workers can be more valuable to the less productive firms, leading them to move down the firm productivity ladder voluntarily. This *down-the-ladder mobility* contrasts the implications of most labor market sorting models but align with empirical patterns.

Knowledge distributions

The distributions of knowledge are endogenous and equilibrium objects. They impact firms' innovation and hiring decisions and evolve as a result of innovation and worker mobility. The evolution of knowledge distributions can be characterized by Kolmogorov Forward (KF) equations.

Firm-worker matches. Consider a cumulative density function (CDF), $\mathcal{F}_m(\cdot,t)$, representing the knowledge distribution among firm-worker matches. The KF equation describes the inflows and outflows for each point of the distribution and is given by:

$$\begin{aligned}
\partial_t \mathcal{F}_m(Z,t) = \quad & -\underbrace{\mu(Z,t) Z\, \partial_Z \mathcal{F}_m(Z,t)}_{\text{Innovation}} - \underbrace{[g-\mu(\underline{Z}_t,t)]\, \underline{Z}_t\, \partial_Z \mathcal{F}_m(\underline{Z}_t,t)\, [1-\mathcal{F}_v(Z,t)]}_{\text{Endogenous separation - Replacement}} \qquad (1.5)\\
\textit{Workers move up}: \quad & -\lambda(\theta_t)\theta_t \int_{\underline{Z}_t}^{Z} \int_{Z}^{\infty} \mathbb{1}\left\{\mathcal{S}(Z',Z'',t)>0\right\} d\mathcal{F}_v(Z',t) d\mathcal{F}_m(Z'',t)\\
\textit{Workers move down}: \quad & +(1-p)\lambda(\theta_t)\theta_t \int_{Z}^{\infty} \int_{\underline{Z}_t}^{Z} \mathbb{1}\left\{\mathcal{S}(Z',Z'',t)>0\right\} d\mathcal{F}_v(Z',t) d\mathcal{F}_m(Z'',t)
\end{aligned}$$

The left-hand side describes the instantaneous change in CDF evaluated at knowledge Z at time t. The first term on the right-hand side reflects the outflows that arise from in-house innovation. Since the $\partial_Z \mathcal{F}_m(Z,t)$ amount of matches at knowledge Z choose innovation intensity $\mu(Z,t)$, they grow above Z at rate $\mu(Z,t)Z$ and are subtracted from the CDF. The second term reflects the net outflows due to voluntary separation. At each instant, the $\partial_Z \mathcal{F}_m(\underline{Z}_t,t)$ amount of matches at the minimum of the support, $\underline{Z}_t$, choose to separate and hence leave the distribution. Each separating match is replaced by a newborn worker randomly matched with a vacant firm. The new match has probability $\mathcal{F}_v(Z,t)$ of having a productivity less than or equal to Z and the number of new matches is $[g-\mu(\underline{Z}_t,t)]\, \underline{Z}_t\, \partial_Z \mathcal{F}_m(\underline{Z}_t,t)$. On net, $[g-\mu(\underline{Z}_t,t)]\, \underline{Z}_t\, \partial_Z \mathcal{F}_m(\underline{Z}_t,t)\, [1-\mathcal{F}_v(Z,t)]$ is subtracted from the corresponding distribution.

The second line reflects the loss of matches when workers move to more productive firms. Workers in matches with knowledge at or below Z search on the job, meet vacant firms at rate $\lambda(\theta_t)\theta_t$. If these workers move to firms with knowledge above Z, the separated matches are subtracted from $\mathcal{F}_m(Z,t)$. The third term is the inflow of matches when workers move down the productivity ladder.

Vacant firms. Denote $\mathcal{F}_v(\cdot,t)$ the knowledge distribution among vacant firms. The KF equation is:

$$\begin{aligned}
\partial_t \mathcal{F}_v(Z,t) = \quad & \underbrace{\frac{N_e(t)}{\theta_t}\mathcal{F}_e(Z,t)}_{\text{Entry}} - \underbrace{g\underline{Z}_t\, \partial_Z \mathcal{F}_v(\underline{Z}_t,t)}_{\text{Endogenous Exit}} - \underbrace{\delta_v \mathcal{F}_v(\underline{Z}_t,t)}_{\text{Exogenous Exit}} \qquad (1.6)\\
\textit{Workers move up}: \quad & +\lambda(\theta_t) \int_{Z}^{\infty} \int_{\underline{Z}_t}^{Z} \mathbb{1}\left\{\mathcal{S}(Z',Z'',t)>0\right\} d\mathcal{F}_m(Z'',t) d\mathcal{F}_v(Z',t)\\
\textit{Workers move down}: \quad & -\lambda(\theta_t) \int_{\underline{Z}_t}^{Z} \int_{Z}^{\infty} \mathbb{1}\left\{\mathcal{S}(Z',Z'',t)>0\right\} d\mathcal{F}_m(Z'',t) d\mathcal{F}_v(Z',t)
\end{aligned}$$

The first component on the right-hand side is the inflows coming from entry. At each instant, N_{et} measure of firms enter the economy and draw initial knowledge from the entry distribution, $\mathcal{F}_e(Z,t)$. The total measure of entrants flowing in below Z is $N_{et}\mathcal{F}_e(Z,t)$. The second component reflects the loss of mass in the distribution due to voluntary exit. At each instant, firms at the minimum of the support, $\underline{Z}_t$, choose to exit the economy. The exit threshold $\underline{Z}_t$ acts as an absorbing barrier sweeping through the distribution from below. As it moves forward at the growth rate g, it collects the $\partial_Z \mathcal{F}_v(\underline{Z}_t, t)$ mass of firms at the minimum of the support and removes them from the distribution. The third component reflects exogenous exit, and since exogenous exit occurs uniformly across all firms, $\delta_v \mathcal{F}_v(\underline{Z}_t, t)$ measure of firms escape from the CDF. The second line describes the outflow that arises from hiring workers. Firms with knowledge at or below Z leave the vacant state upon hiring workers. If they hire workers from firms at or below Z, the outflow of the vacant firms is exactly offset by inflow from the separated matches, leaving $\mathcal{F}_v(Z, t)$ unchanged. If the workers are from firms above Z, the vacant firms are subtracted from $\mathcal{F}_v(Z, t)$. The second line describes the inflow that arises from separated matches.

The economy has θ_t measure of vacant firms. The law of motion for θ_t follows

$$\partial_t \theta_t = \underbrace{N_{et}}_{\text{Entry}} - \underbrace{g \underline{Z}_t \, \partial_Z \mathcal{F}_v(\underline{Z}_t, t) \, \theta_t}_{\text{Endogenous Exit}} - \underbrace{\delta_v \, \theta_t}_{\text{Exogenous Exit}} \quad .$$

Balanced Growth Path

BGP equilibrium.

In equilibrium, firms innovate and enter the labor market optimally, taking the knowledge distributions as given. The distributions evolve as firms' choices dictate. A formal definition follows.

Definition (Equilibrium). A recursive competitive equilibrium consists of: value functions $\{\mathcal{M}(Z, t), \mathcal{V}(Z, t)\}$, innovation intensity $\mu(Z, t)$, knowledge distributions $\{\mathcal{F}_m(Z, t), \mathcal{F}_v(Z, t)\}$, market tightness θ_t, and interest rate r_t such that:

1. Given $\{r_t, \theta_t, \mathcal{F}_m(Z, t), \mathcal{F}_v(Z, t)\}$, $\{\mathcal{M}(Z, t), \mathcal{V}(Z, t)\}$ solve the HJB equations (1.2) (1.4), with $\mu(Z, t)$ the associated decision rules (1.3);
2. Given $\{\mu(Z, t), \theta_t\}$, $\{\mathcal{F}_m(Z, t), \mathcal{F}_v(Z, t)\}$ evolve according to KF equations (1.5) (1.6);
3. Given $\{\mathcal{F}_v(Z, t), \mathcal{F}_m(Z, t)\}$, market tightness θ_t adjusts so that the free entry condition (1.1) holds;
4. r_t is consistent with the representative household's inter-temporal marginal rate of substitution.

Definition (BGP). A balanced growth path (BGP) equilibrium is a recursive competitive equilibrium such that the growth rate g of aggregate consumption Y_t is constant and the distributions of knowledge are stationary when rescaled, i.e.,

$$g = \frac{\dot{Y}_t}{Y_t} \quad , \quad \mathcal{F}_v(Z, t) = \mathcal{F}_v(Ze^{-gt}, 0) \quad , \quad \mathcal{F}_m(Z, t) = \mathcal{F}_m(Ze^{-gt}, 0).$$

Normalization. I study the economies in equilibrium on balanced growth paths. To compute the BGP equilibrium, it is convenient to normalize the economy and transform this system into a set of stationary equations. Let $g := \frac{\dot{Y}_t}{Y_t}$ be the growth rate on the balanced growth path. Define normalized state variable z, value functions $\{\boldsymbol{M}(z), \boldsymbol{V}(z)\}$, and distributions $\{\boldsymbol{F}_m(z), \boldsymbol{F}_v(z)\}$ as:

$$\begin{aligned} z := \log(Z) - gt \quad , \quad \boldsymbol{M}(z) &:= e^{-gt}\mathcal{M}(Z,t) \quad , \quad \boldsymbol{F}_m(z) := \mathcal{F}_m(Z,t) \\ \boldsymbol{V}(z) &:= e^{-gt}\mathcal{V}(Z,t) \quad , \quad \boldsymbol{F}_v(z) := \mathcal{F}_v(Z,t) \end{aligned}$$

The main advantage of the normalized system is that it reduces the value function to one of state variable z alone, removing the dependence on time. This mirrors the normalization of the knowledge distribution. Thus, computing a balanced growth path equilibrium using the normalized system of equations involves solving ordinary differential equations.

On the balanced growth path, the normalized continuation value functions for matches and vacant firms in equations (1.2) (1.4) simplify to

$$(r-g)\boldsymbol{M}(z) = \max_{\mu} \underbrace{e^z}_{\text{Production}} - \underbrace{\kappa_{r\&d}(\mu)\, e^z}_{\text{Innovation Cost}} + \underbrace{(\mu - g)\partial_z \boldsymbol{M}(z)}_{\text{Gains from Innovation}} \tag{1.7}$$

$$(r-g)\boldsymbol{V}(z) = \underbrace{e^z}_{\text{Production}} + \underbrace{\lambda(\theta)}_{\text{Meeting Rate}} \int \underbrace{\left[\boldsymbol{S}(z,z')\right]^+}_{\text{Surplus from Trade}} d\boldsymbol{F}_m(z') - g\, \partial_z \boldsymbol{V}(z), \tag{1.8}$$

where the normalized gains from trade $\boldsymbol{S}(z,z')$ follow

$$\boldsymbol{S}(z,z') = \begin{cases} \Big[\boldsymbol{M}(z) - \boldsymbol{V}(z)\Big] - \Big[\boldsymbol{M}(z') - \boldsymbol{V}(z')\Big] & \text{if} \quad z \leqslant z' \\ \Big[p\boldsymbol{M}(z') + (1-p)\boldsymbol{M}(z) - \boldsymbol{V}(z)\Big] - \Big[\boldsymbol{M}(z') - \boldsymbol{V}(z'))\Big] & \text{if} \quad z < z' \end{cases}$$

Given the static profit functions and perceived laws of motion for knowledge distributions, each firm-worker match chooses how intensively to innovate, μ. The first-order condition for innovation intensity gives the dynamic decision rule:

$$\partial_\mu \kappa_{r\&d}(\mu) = e^{-z}\partial_z \boldsymbol{M}(z)$$

On the BGP, the market tightness θ is constant and is equilibrated through the free entry condition:

$$\int \boldsymbol{V}(z)\ d\boldsymbol{F}_e(z) = \kappa_{\text{entry}}.$$

Firms keep entering the market until the expected entry value equals the normalized entry cost $\kappa_{\text{entry}} = e^{-gt}\kappa_{\text{entry}}(t)$. The mass of entrants $n_e = N_{et}$ is constant over time.

The normalized knowledge distributions are stationary. The inflow balances the outflow for each point in the support of the distribution. The KF equations (1.5) (1.6) take the form

$$
\begin{aligned}
0 = & -\partial_z\left[(\mu(z))\, f_m(z)\right] + g\,\partial_z f_m(z) \\
& + \lambda(\theta)\theta\, f_v(z) \int_{\underline{z}}^{z} \mathbb{1}\left\{S(z,z') > 0\right\} \mathrm{d}F_m(z') + (1-p)\lambda(\theta)\theta\, f_v(z) \int_{z}^{\infty} \mathbb{1}\left\{S(z,z') > 0\right\} \mathrm{d}F_m(z') \\
& - (1-p)\lambda(\theta)\theta\, f_m(z) \int_{\underline{z}}^{z} \mathbb{1}\left\{S(z',z) > 0\right\} \mathrm{d}F_v(z') - \lambda(\theta)\theta\, f_m(z) \int_{z}^{\infty} \mathbb{1}\left\{S(z',z) > 0\right\} \mathrm{d}F_v(z') \\
0 = & \; g\partial_z f_v(z) + \frac{n_e}{\theta} f_e(z) - \delta_v f_v(z) \\
& + \lambda(\theta) f_m(z) \int_{\underline{z}}^{\infty} \mathbb{1}\left\{S(z',z) > 0\right\} \mathrm{d}F_v(z') - \lambda(\theta) f_v(z) \int_{\underline{z}}^{\infty} \mathbb{1}\left\{S(z,z') > 0\right\} \mathrm{d}F_m(z')
\end{aligned}
$$

Besides, the measure of vacant firms θ is stable over time. The law of motion for θ is

$$
0 = \underbrace{n_e}_{\text{Entry}} - \underbrace{\theta\, g f_v(\underline{z})}_{\text{Endogenous Exit}} - \underbrace{\delta_v \theta}_{\text{Exogenous Exit}} .
$$

At each instant, the lowest productive firms choose to exit the economy. The number of firms hitting the exit boundary per instant is the product of two terms: The measure of firms at the boundary, $\theta f_v(\underline{z})$, and the relative speed g at which the firm drifts towards the boundary. The exit of unproductive firms is replaced by, on average, more productive entrants. At each instant, a measure n_e of firms enter the economy. The inflow of entrants balances the outflow of exiting firms.

Growth decomposition

The aggregate growth rate is given by the growth rate of aggregate knowledge. Define the normalized the aggregate knowledge as

$$
\mathcal{Z} = \underbrace{\int_{\underline{z}}^{\infty} e^z\, \mathrm{d}F_m(z)}_{\text{Matches}} + \underbrace{\theta \int_{\underline{z}}^{\infty} e^z\, \mathrm{d}F_v(z)}_{\text{Vacant Firms}} .
$$

The knowledge growth comes from three sources: (i) innovation, (ii) diffusion through worker mobility, and (iii) net growth from firm entry and exit. Although the three sources are not independent, the aggregate growth rate can be decomposed into three additively separable components as follows:

$$
g = \frac{1}{\mathcal{Z}} \left\{
\begin{aligned}
& \underbrace{\int_{\underline{z}}^{\infty} \mu(z) e^z\, \mathrm{d}F_m(z)}_{\text{Innovation}} + \underbrace{\lambda(\theta)\theta p \int_{\underline{z}}^{\infty} \int_{z}^{\infty} \left(e^{z'} - e^{z}\right) \mathbb{1}\left\{S(z,z') > 0\right\} \mathrm{d}F_m(z') \mathrm{d}F_v(z)}_{\text{Diffusion through worker mobility}} \\
& \underbrace{+ n_e \int_{\underline{z}}^{\infty} e^z \mathrm{d}F_e(z)}_{\text{Entry}} - \underbrace{\theta\, g f_v(\underline{z})}_{\text{Endogenous Exit}} \underbrace{- \delta_v\, \theta \int_{\underline{z}}^{\infty} e^z \mathrm{d}F_v(z)}_{\text{Exogenous Exit}}
\end{aligned}
\right\} .
$$

An analytical BGP

This subsection focuses on cases where a BGP equilibrium can be solved analytically. I make a perturbation to the labor search technology: firms randomly meet workers from all firms that are more productive than they are. With this perturbation, the BGP equilibrium can be solved analytically and admits a theoretical decomposition of the aggregate growth rate. I will return to the general model when I study the planner's problem in subsection 1.2 and quantitative analysis in section 1.4.

Assumptions. I deviate from the random search technology and assume search is *semi-random*: a vacant firm randomly meets a worker drawn from all matches that are more productive than they are. Let $\mathcal{F}_m(\cdot, t)$ be the knowledge distribution of firm-worker matches. Then, a vacant firm Z draws a worker from the conditional distribution $\mathcal{F}_m(Z'|Z' \geqslant Z, t)$.

To maintain tractability, I make the following functional form assumptions. The innovation cost function is $\kappa_{r\&d}(\mu) = \tilde{\kappa}_{r\&d}\, \mu^{\gamma}$. The initial distributions of knowledge are Pareto with the minimum of the support normalized to one:

$$\mathcal{F}_m(Z, 0) = 1 - \left(\frac{1}{Z}\right)^{\zeta_m} \quad , \quad \mathcal{F}_v(Z, 0) = 1 - \left(\frac{1}{Z}\right)^{\zeta_v}.$$

Although knowledge diffusion and innovation are endogenous, the BGP equilibrium can be characterized analytically.

Proposition 1.1. *Given Pareto initialization, there exists a unique Balanced Growth Path Equilibrium, where:*

1. *The value functions are linear in knowledge level*

$$\mathcal{M}(Z) = \tilde{\kappa}_{r\&d}\, \gamma\, \mu^{\gamma-1} Z$$

2. *The optimal innovation intensity is constant and solves*

$$\rho\, \tilde{\kappa}_{r\&d}\, \gamma\, \mu^{\gamma-1} = \tilde{\kappa}_{r\&d}\, (\gamma - 1)\mu^{\gamma} + 1$$

3. *The distributions of knowledge are Pareto with shape parameters:*

$$\zeta_v = \zeta_m - \frac{\lambda(\theta)}{\mu}$$

4. *The growth rate follows*

$$g = \mu = \frac{\lambda(\theta)}{\zeta_m - \zeta_v}$$

Growth Decomposition.

$$g = \frac{1}{\mathfrak{Z}}\left[\underbrace{\frac{\mu\,\zeta_m}{\zeta_m - 1}}_{\text{Innovation}} + \underbrace{\frac{\lambda(\theta)\,\zeta_v}{(\zeta_m - 1)(\zeta_v - 1)}\theta}_{\text{Diffusion through Worker Mobility}} + \underbrace{\frac{g\,\zeta_v}{\zeta_m - 1}\theta}_{\text{Entry-Exit}}\right],$$

where normalized aggregate output $\mathfrak{Z}$ is given by

$$\mathfrak{Z} = \frac{\zeta_m}{\zeta_m - 1} + \theta\frac{\zeta_v}{\zeta_v - 1}$$

Planner's Problem

Despite being bilaterally efficient, the decentralized equilibrium is socially inefficient. Firms do not internalize the effect their innovation and entry decisions have on the evolution of the knowledge distribution and, in turn, the distribution for future hiring firms. A constrained planner's problem is useful for characterizing inefficiency.

Definition (Constrained planner's problem). I consider a planner whose objective is to maximize social welfare, defined as the present discounted value of aggregate output net of investment and entry costs. The planner is subject to the same labor market frictions and technological constraints faced by firms in decentralized equilibrium. The key difference is that the planner internalizes the social benefits of innovation and the congestion externality of entrants. The planner chooses innovation intensities for each firm and the number of entrants, solving the HJB equation:

$$\rho\boldsymbol{\Omega}(f_m(\cdot,t),\phi_v(\cdot,t)) = \max_{\substack{\mu(\cdot,t)\\ N_e(t)}}\left\{\begin{array}{l}\displaystyle\int_{\underline{Z}_t}^{\infty}\underbrace{[1-\kappa_{r\&d}(\mu(Z,t))]\,Z}_{\text{Output - Innovation Cost}}\,f_m(Z,t) + \underbrace{\widehat{\mathcal{M}}(Z,t)}_{\text{Social Value: Match}}\quad\partial_t f_m(Z,t)dZ\\ \displaystyle -\underbrace{\kappa_{\text{entry}}(t)}_{\text{Entry Cost}}N_e(t) + \int_{\underline{Z}_t}^{\infty}\underbrace{Z}_{\text{Output}}\,\phi_v(Z,t) + \underbrace{\widehat{\mathcal{V}}(Z,t)}_{\text{Social Value: Vacant}}\quad\partial_t\phi_v(Z,t)dZ\end{array}\right\}$$

Social welfare is denoted by $\boldsymbol{\Omega}(f_m(\cdot,t),\phi_v(\cdot,t))$, where state variables $f_m(\cdot,t)$ and $\phi_v(\cdot,t)$ describe densities of vacant firms and firm-worker matches at instant t. The law of motion for vacant firms,

$\partial_t \phi_v(Z,t)$, and for matches, $\partial_t f_m(Z,t)$, follow:

$$
\begin{aligned}
\partial_t f_m(Z,t) \quad &= -\underbrace{\partial_z\left[\mu(Z,t)Z f_m(Z,t)\right]}_{\text{Innovation}} + \lambda(\theta_t)\eta(Z,t)\phi_v(Z,t)\int_{\underline{Z}_t}^{Z} \mathbb{1}\{\mathcal{S}(Z,X,t)>0\}\, d\mathcal{F}_m(X,t) \\
&+(1-p)\lambda(\theta_t)\eta(Z,t)\phi_v(Z,t)\int_{Z}^{\infty} \mathbb{1}\{\mathcal{S}(Z,X,t)>0\}\, d\mathcal{F}_m(X,t) \\
&-(1-p)\lambda(\theta_t) f_m(Z,t)\int_{\underline{Z}_t}^{Z} \eta(Z',t)\mathbb{1}\left\{\mathcal{S}(Z',Z,t)>0\right\} d\Phi_v(Z',t) \\
&-\lambda(\theta_t) f_m(Z,t)\int_{Z}^{\infty} \eta(Z',t)\mathbb{1}\left\{\mathcal{S}(Z',Z,t)>0\right\} d\Phi_v(Z',t)
\end{aligned}
$$

$$
\begin{aligned}
\partial_t \phi_v(Z,t) \quad &= \underbrace{n_e(t) f_m(Z,t)}_{\text{Entry}} + \underbrace{(1-\tau)\lambda(\theta_t) f_m(Z,t)\int_{\underline{Z}_t}^{\infty} \eta(Z',t)\mathbb{1}\left\{\mathcal{S}(Z',Z,t)>0\right\} d\Phi_v(Z',t)}_{\text{Match Separation - Business Stealing}} \\
\textit{Hire}: \quad &-\lambda(\theta_t)\eta(Z,t)\phi_v(Z,t)\int_{\underline{Z}_t}^{\infty} \mathbb{1}\{\mathcal{S}(Z,X,t)>0\}\, d\mathcal{F}_m(X,t)
\end{aligned}
$$

$$
\partial_t \phi_v(\underline{Z}_t,t) \quad = \underbrace{n_e(t) f_m(\underline{Z}_t,t)}_{\text{Entry}} - g\,\phi_v(\underline{Z}_t,t) - g\underline{Z}_t\,\partial_z \phi_v(\underline{Z}_t,t)
$$

$$
\partial_t n_v(t) \quad = \underbrace{n_e(t)}_{\text{Entry}} - \underbrace{n_v(t)\; g\underline{Z}_t\; f_v(\underline{Z}_t,t)}_{\text{Exit}} - \underbrace{n_v(t)\;\lambda(\theta_t)\tau\int_{\underline{Z}_t}^{\infty}\eta(Z,t)\int_{\underline{Z}_t}^{\infty}\mathbb{1}\left\{\mathcal{S}(Z,Z',t)>0\right\} d\mathcal{F}_m(Z',t)d\mathcal{F}_v(Z,t}_{\text{Business Stealing}}
$$

Definition (Social value). Let $\widehat{\mathcal{V}}(Z,t)$ be the social value of a vacant firm of knowledge Z at instant t, $\widehat{\mathcal{M}}(Z,t)$ be the social value of a firm-worker match:

$$
\widehat{\mathcal{V}}(Z,t) := \frac{\delta\boldsymbol{\Omega}(f_m(\cdot,t),\phi_v(\cdot,t)}{\delta\phi_v(Z,t)} \quad , \quad \widehat{\mathcal{M}}(Z,t) := \frac{\delta\boldsymbol{\Omega}(f_m(\cdot,t),\phi_v(\cdot,t))}{\delta f_m(Z,t)}.
$$

The social values are defined along the optimal trajectory of the densities, $\phi_v(\cdot,t)$ and $f_m(\cdot,t)$. Here $\dfrac{\delta}{\delta\phi_v(Z,t)}$ is the "functional derivative" of the social welfare with respect to $\phi_v(\cdot,t)$ at point Z.

Social value of vacant firms. The social value of a vacant firm Z satisfies the HJB equation:

$$
\begin{aligned}
\rho\widehat{\mathcal{V}}(Z,t) = \max_{\eta} \quad & Z - c_{\text{search}}(\eta)Z + \partial_t\widehat{\mathcal{V}}(Z,t) + \lambda(\theta_t)\,\eta\int_{\underline{Z}_t}^{\infty}\left[\widehat{\mathcal{S}}(Z,Z',t)\right]^{+} d\mathcal{F}_m(Z',t) \\
& \underbrace{+\lambda'(\theta_t)\eta(Z,t)\int_{\underline{Z}_t}^{\infty}\eta(Z',t)\int_{\underline{Z}_t}^{\infty}\left[\widehat{\mathcal{S}}(Z',X,t)\right]^{+} d\mathcal{F}_m(X,t)d\Phi_v(Z',t)}_{\text{Congestion Externality}}
\end{aligned}
$$

where social gains from trade $\widehat{\mathcal{S}}(Z, Z', t)$ follows

$$\widehat{\mathcal{S}}(Z, Z', t) = \begin{cases} \left[\widehat{\mathcal{M}}(Z,t) - \widehat{\mathcal{V}}(Z,t)\right] - \left[\widehat{\mathcal{M}}(Z',t) - \widehat{\mathcal{V}}(Z',t)\right] & \text{if} \quad Z > Z' \\ \left[p\widehat{\mathcal{M}}(Z',t) + (1-p)\widehat{\mathcal{M}}(Z,t) - \widehat{\mathcal{V}}(Z,t)\right] - \left[\widehat{\mathcal{M}}(Z',t) - \widehat{\mathcal{V}}(Z',t))\right] & \text{if} \quad Z \leqslant Z' \end{cases}$$

Social value of matches. The social value of a match Z satisfies the HJB equation:

$$\begin{aligned} \rho\widehat{\mathcal{M}}(Z,t) = \max_{\mu} \quad & Z - c_{r\&d}(\mu)Z + \mu Z\,\partial_z\widehat{\mathcal{M}}(Z,t) + \partial_t\widehat{\mathcal{M}}(Z,t) \\ & \underbrace{+\lambda(\theta_t)\int_{\underline{Z}_t}^{\infty} \eta(Z',t)\left[\widehat{\mathcal{S}}(Z',Z,t)\right]^+ d\Phi_v(Z',t)}_{\text{Externality to Searching Firms}} + \underbrace{n_e(t)\widehat{\mathcal{V}}(Z,t)}_{\text{Externality to Entrants}} \\ & \underbrace{-\lambda'(\theta_t)\theta_t\int_{\underline{Z}_t}^{\infty} \eta(Z',t)\int_{\underline{Z}_t}^{\infty}\left[\widehat{\mathcal{S}}(Z',X,t)\right]^+ d\mathcal{F}_m(X,t)d\Phi_v(Z',t)}_{\text{Congestion Externality}} \end{aligned}$$

The planner's social value has six components. The first four components resemble the joint value of a match, with private continuation value replaced by social value. The fifth component captures the externality to entrants, namely the expected value from knowledge improvement to entrants. The second line describes the externality to searching firms.

Social value of entrants. The social value of an entrant satisfies the HJB equation:

Non-compete Contract

Contract. A firm designs a contract $\mathcal{C}$ to maximize the present discounted value of profits subject to the promised utility W_{t_0} to its worker. A contract specifies whether to include a non-compete clause, innovation rates, wages, and potential buyout offers for every future sequence of states:

$$\mathcal{C} = \left\{ \mathbb{1}_{\{NCA\}_{t_0}}\ ,\ \left\{ \mu_t(s_t)\ ,\ w_t(s_t)\ ,\ B_t(s_t, Z_t')\right\}_{t=t_0}^{\infty} \right\} \quad , \quad \text{where} \quad s_t := \left\{ Z_t\ ,\ H_t\ ,\ W_t \right\}.$$

Non-compete clause $\mathbb{1}_{\{NCA\}_{t_0}} \in \{0, 1\}$ is a one-time decision made upon match formation at instant t_0. Innovation rate μ_t and wage w_t are state-contingent, where state s_t includes firm knowledge Z_t, worker knowledge H_t, and the utility promise W_t made to the worker. Since the contracts respond to outside offers by matching counteroffers, the utility promise is determined by the willingness-to-pay of the best outside offer a worker has previously received.

Non-compete clause. A non-compete clause enables rent appropriation from future employers in the market for knowledge workers. This is achieved through buyout payment, where the current employer receives a lump sum payment from the hiring firm in exchange for waiving the non-compete clause. The game tree in Figure 1.1 describes the dynamics of the contract.

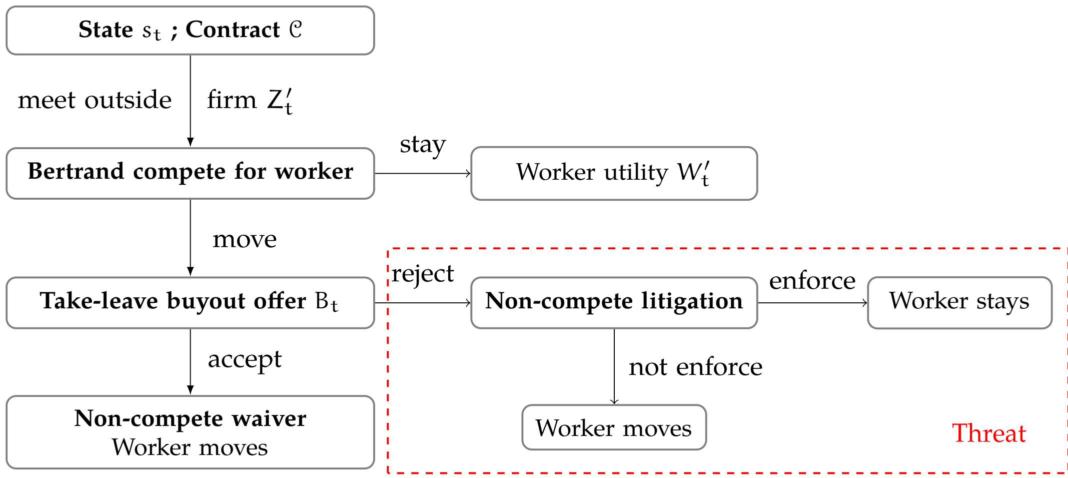


Figure 1.1: Game tree in search-and-matching stage

Consider a worker H employed at firm Z whose current contract delivers utility promise W. Suppose the worker receives an outside offer valued $\tilde{W}'$ from firm Z' and intends to move. In the presence of a non-compete clause, the incumbent firm makes a take-it-or-leave-it offer of buyout B to the outside firm, using non-compete litigation as a threat. If the outside firm accepts the offer, the worker is released from the clause. If the outside firm rejects the offer but forms a match with the worker, the incumbent firm can sue the outside firm in court. A non-compete clause is enforced with probability β, where β is exogenous and governed by state law. Enforcement leads to the separation of the new match.

The game can be solved recursively. Given the buyout offer B, the outside firm chooses whether to accept, weighing the two options:

$$\mathcal{O}(Z', H, \tilde{W}', B, t) = \max\left\{ \underbrace{\mathcal{M}(Z', H, t) - \tilde{W}' - B}_{\text{Accept buyout}} \; , \; \underbrace{\beta \mathcal{V}(Z', t) + (1-\beta)\left[\mathcal{M}(Z', H, t) - \tilde{W}'\right]}_{\text{Reject buyout}} \right\}.$$

The first element describes the value of accepting, which equals the marginal value of the worker net of the buyout payment. The second element represents the expected value of rejecting: if the non-compete clause is enforced, which occurs with probability β, the outside firm becomes vacant with value $\mathcal{V}(Z', t)$; otherwise, the firm collects the marginal value of the worker.

Accordingly, the continuation value of the incumbent firm reads as

$$\Pi(Z, H, \tilde{W}', B, t) = \begin{cases} B + \mathcal{V}(Z, t) & \text{if buyout offer is accepted} \\ \beta\left[\mathcal{M}(Z, H, t) - \tilde{W}'\right] + (1-\beta)\,\mathcal{V}(Z, t) & \text{if buyout offer is rejected} \end{cases}$$

If the buyout offer is accepted by outside firm, the incumbent firm receives the buyout payment and becomes vacant (first row). The second row captures the expected value when offer is rejected: if the non-compete clause is enforced, the new match will be terminated, and the incumbent firm will rehire the worker and promise utility $\tilde{W}'$; if the non-compete clause is not enforced, the incumbent firm will be vacant.

Going back one step in the game tree, incumbent and outside firms Bertrand compete for the worker. The incumbent firm matches the outside firm's willingness to pay up to its own willingness to pay. The worker chooses the firm with higher willingness to pay, and obtains utility promise W':

$$W' = \min\left\{ \underbrace{\mathcal{M}(Z,H,t) - \mathcal{V}(Z,t) - B}_{\text{Incumbent WTP}},\ \underbrace{\mathcal{O}(Z',\tilde{W}',B,t) - \mathcal{V}(Z',t) + \tilde{W}'}_{\text{Outside WTP}} \right\}.$$

1.3 Data and Empirical Analysis

This section empirically examines inter-firm knowledge diffusion through worker mobility, and the innovation-diffusion trade-off associated with non-compete contracts. My analysis centers on two groups of workers: inventors and executives. Inventors and executives are particularly suitable for studying knowledge diffusion because of their direct exposure to employers' technical or managerial knowledge, and the prevalence of non-compete contracts among those workers. I build two new sets of matched employer-employee data. The first set is matched employer-inventor administrative data from the U.S. Census Bureau. The main advantage of administrative data is that it enables us to identify inventor mobility directly and observe earnings. The second set is matched employer-executive data within U.S. publicly traded firms. I compile the data by scraping and analyzing employment contracts from SEC Edgar. This data offers unique, granular information on non-compete contracts among executives in publicly traded firms.

Section 1.3 describes the primary datasets employed and variable constructions. Section 1.3 presents the empirical strategy and findings.

Data

Matched Employer-Inventor Data

I construct a new dataset containing patent records, firm-level measures, worker employment history, and earnings using the United States Patent and Trademark Office (USPTO) PatentsView Database, Longitudinal Business Database (LBD), Revenue-enhanced Longitudinal Business Database (RELBD), and Longitudinal Employer-Household Dynamics (LEHD) from 1997 through 2019. The data offers a unique opportunity to observe (i) inventor mobility, (ii) inventor compensation, (iii) inter-firm knowledge flows, and (iv) firm productivity.

Patent citations. I identify inventors and measure inter-firm knowledge diffusion using USPTO data from 1976 onward. This data contains rich information for granted patents, including application and grant dates, citations to other patents, and the name and address of patent assignees (firms, institutions, or individuals that own the property rights of a patent). I use this data to identify inventors and assignees of granted patents.

I use patent citations to track inter-firm knowledge flows. Specifically, for each pair of origin and destination firms associated with a mobile worker, I count how many post-mobility new patents of the destination firm cite the pre-mobility patent stock of the origin firm. Patent citations are informative of knowledge diffusion because legal obligations mandate patent applicants to disclose any relevant "prior art" they know. So, each citation indicates one instance of the destination firm drawing upon the knowledge stock of the origin firm. Admittedly, not all citations represent knowledge diffusion, as some may be introduced to distinguish the invention from dissimilar ones or to protect the firm from legal disputes. Nevertheless, patent citations offer useful and tangible measures for tracing knowledge flows.

Firm productivity. I measure firm productivity using revenue and payroll information from the LBD and RELBD. The LBD tracks the universe of U.S. business establishments and firms with at least one paid employee from 1976 onward. It provides rich information on employment, labor costs, industry codes, business names, and location. I augment LBD with revenue information from RELBD, a subset of the LBD merged with income tax filings. I further collect industry-by-year-level labor shares from the U.S. Bureau of Labor Statistics, and calculate the total variable cost of a firm by dividing labor cost by industry-level labor share. Firms' annual productivity is measured using revenue-based Total Factor Productivity (TFPR). This productivity measure is equivalent to revenue per unit of composite input when there are constant returns to scale in production.

$$\mathsf{TFPR} = \frac{\text{Revenue}}{\text{Labor cost/Industry-level labor share}}$$

Inventor mobility. I collect inventor mobility and earnings from LEHD. The LEHD is a matched employer-employee dataset that covers over 95% of private sector workers. My access to the LEHD dataset spans from 1991 to 2021 and across 29 states, collectively representing over 60% of private sector employment in the United States.[9] For each worker, I track their employers and earnings every quarter. I assign inventor records in USPTO to workers in the LEHD, built on a crosswalk developed by Akcigit and Goldschlag (2023). Ultimately, I observe the employment histories and earnings of approximately 826,000 inventors from 1997 to 2021.

Compared with the prior literature that infers inventor mobility from patent trajectories, the inventor LEHD offers a unique opportunity to identify inventor movements between jobs and to observe earnings. Following the standard practice in the literature, I keep the primary job (job

[9]The 29 states include Alabama, Arizona, California, Colorado, Connecticut, Delaware, Idaho, Indiana, Kansas, Maryland, Maine, North Dakota, Nebraska, New Jersey, New Mexico, Nevada, New York, Ohio, Oklahoma, Oregon, Pennsylvania, South Dakota, Tennessee, Texas, Utah, Virginia, Washington, Wisconsin, and Wyoming.

with the highest earnings) if a worker is employed in multiple firms in a quarter. I define job-to-job mobility as moving to a new employer after leaving a previous job, either in the same or the subsequent quarter. Earnings are reported quarterly and normalized to 2012 dollars.

Inventor sample. Linking these datasets, I construct a matched employer-inventor data containing around 826,000 inventors between 1997 and 2019.[10] The inventor labor market is fluid, with 76.8% of inventors having changed jobs at least once during the sample period. To examine the impact of inventor mobility, I focus my analysis on the inventors who have experienced job changes. I also restrict attention to inventors aged 18 to 65 (inclusive). Table 1.1 presents some basic summary statistics for my analysis sample. The sample includes approximately 634,000 inventors employed by 325,000 firms from 1997 to 2019. On average, an inventor has been employed in 4.23 firms and earns mean quarterly earnings of $39,810. The average job tenure is 18.45 quarters. About 5.98% of inventors change jobs each quarter, and 48.69% of movements are down the firm productivity ladder. Notably, 5.98% of inventors move to a new job each quarter, with nearly half of these movers (48.69%) moving to less productive firms.

Table 1.1: Summary Statistics of Analysis Sample

Inventors		Firms	
Quarterly earnings	$39,810	Productivity	2.593
Log quarterly earnings	10.13	Log productivity	0.293
# Employers	4.23	# Employees	303.7
Tenure (quarters)	18.45	Revenue	$87,080
Quarterly mobility rate	5.98%	Labor cost	$16,290
Move to less productive firms	48.69%	# Years to hire inventors	5.39
# Patents per year	0.07	# Patents per year	0.69
Inventor age	41.95	Firm age	17.86
Observations (rounded to 000s)	634,000	Observations (rounded to 000s)	325,000

Matched Employer-Executive Data

I construct a matched firm-executive panel dataset for 1992-2021 that contains information on employment history, non-compete contracts, innovation, and productivity.

Executive employment I collect the employment histories of 112,046 executives in 18,012 publicly listed firms from 1992 to 2021. On average, an executive holds two jobs and spends 6 years on each. Overall, 38% of executives change jobs at least once throughout the sample period, and 10% of executives change jobs annually.

[10]The linking process involves assigning USPTO PatentsView patent records to firms in the Longitudinal Business Database (LBD) using the Business Dynamics Statistics of Patenting Firms (BDS-PF) crosswalk from the Census (Dreisigmeyer, Goldschlag, Krylova, Ouyang, Perlman, et al., 2018), and matching inventor records to workers in the LEHD following a methodology developed by Akcigit and Goldschlag (2023).

The data construction starts with the ExecuComp database, covering 55,074 executives from 1992 onward. Each executive was among the five top-paid employees (some up to fifteen) in firms comprising the S&P 500, MidCap 400, and SmallCap 600 indices. I compliment ExecuComp with the Capital IQ People Intelligence database, which provides a broader coverage of over 2 million executives, board members, and investment professionals in U.S. public and private firms since 1992. I further use BoardEx to obtain the complete employment histories of over 1 million individuals who served as executives, senior managers, and directors in U.S. public and private firms between 1999 and 2021. Each database has internally coherent identifiers referring to unique individuals. Individuals are linked using identifier crosswalks provided by WRDS and fuzzy name matching when crosswalks are unavailable.

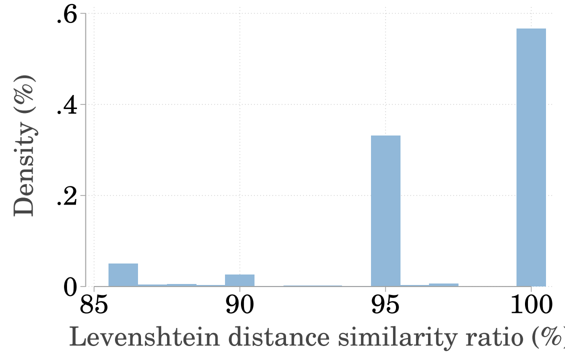
The compiled dataset contains the different firms an executive worked for, job positions at each firm, the dates those positions began and ended, and detailed compensation. The primary job is kept if an executive works for multiple firms in a fiscal year, often as a board member. A primary job is a managerial position involving day-to-day operations, such as being named executive officer, president, and founder. In the rare cases where an executive serves managerial positions simultaneously for multiple firms, the job with the highest annual compensation is identified as the primary job. Job mobility is defined as occurring in the last fiscal year when an executive is in office for the greater part of the fiscal year. Mobility due to acquisition, bankruptcy, or delisting is excluded.

Non-compete contracts I create a unique dataset of non-compete contracts with broad coverage over 34,786 executives in 9,255 publicly listed U.S. firms for 2000-2021. The dataset is constructed from the disclosed employment contracts in the SEC's EDGAR system. For each employment contract, I observe whether a non-compete clause is included, the duration of non-competes, the executive name, and the firm identifier. Overall, 65% of executives have signed non-competes.

The starting point is to gather all mandatory disclosures filed with the SEC using web crawling algorithms. Employment contracts of named executive officers and directors are among the mandatory disclosures and are generally attached under the "exhibit 10" designation to 10-K, 10-Q, and 8-K filings. Relevant employment contracts include employment agreements, letters of employment, non-compete agreements, retention agreements, separation and severance agreements, and contract amendments.

To identify employment contracts from other mandatory disclosures under the "exhibit 10" designation, I train a text classification algorithm. First, I label 29% of the 818,719 collected disclosures into employment or other contracts. The labeled sample is randomly split into a training set (1/3) and a test set (2/3). With the training set, I extract text-based features using the bag-of-words model and train a random forest classification algorithm. The classification algorithm contains top keywords and corresponding weights as described in Figure 1.2a, and yields an accuracy rate of 98.3% on the test set. I leverage the algorithm to categorize the unlabeled disclosures. In total, 128,252 employment contracts are collected.

(a) Keywords to classify employment contracts



(b) Similarity between matched names

I process the text in employment contracts to extract executive names and non-compete clauses. Natural language processing models are applied to recognize names in the following semantic segments: (i) title of contract, (ii) recipient of an employment letter, (iii) sentences containing keywords such as (the "employee") and (the "executive"), (iv) signatures following key marks such as "/s/". The last segment yields the least precise name recognition and serves as validation for names extracted from other segments. Non-compete clauses are detected if a contract contains keywords related to "non-compete". Durations are extracted from paragraphs containing variations of "non-compete" keywords using the natural language processing models.

Matching non-competes to employment Having executive names, firm identifiers, and employment periods, I match non-compete arrangements to employment records using a name-matching algorithm. For each contract, I search for the closest name match from a set of executives who worked for the same firm in the contract year. Among the 137,795 employment contracts collected, 61,423 contracts can find a close match where the Levenshtein distance similarity ratio is higher than 86%. Figure 1.2b displays the probability distribution of the similarity ratio. Visual inspection of the matched names also confirms very few mistakes in the matching.

Patents I assemble the universe of patents granted in the U.S. since 1926 by combining PatentsView (US Patent and Trademark Office), WRDS US Patents, and Google Patents. Patents are matched to publicly listed U.S. firms following the crosswalks provided by WRDS and Kogan et al. (2017).

Patent data is used for three purposes. The first is to construct the number of newly granted patents per firm and year. Such patent flow measures innovation output. The second is to construct the cumulative number of granted patents per firm and year. Patent stock measures the stock of knowledge and hence the technology ladder of a firm. The third purpose is to measure patent citations between origin and destination firms to infer knowledge diffusion through executive mobility. To accomplish this, I count the number of the destination's new patents granted since hiring an executive which cite the origin's patent stock before executive mobility (destination's post-mobility patent flow cites origin's pre-mobility patent stock). Here the new patents capture knowledge the destination firm learned after hiring the executive. The stock of patents captures knowledge the executive learned from the origin firm. A citation indicates that a citing patent uses similar knowledge in the cited patent. I further complement the measure of knowledge proximity with textual similarities between citing and cited patents (as obtained from Whalen et al. (2020)).

Empirical Analysis

I start this section by documenting inter-firm knowledge diffusion through the mobility of inventors and executives. Then, I examine the innovation-diffusion trade-off in the context of non-compete contracts. There are four main findings:

(i) Inventor mobility diffuses knowledge across firms. After hiring an inventor from a more productive firm, the destination firm more intensively draws upon the knowledge of the origin firm and experiences productivity growth.

(ii) Inventors are compensated for knowledge diffusion. After moving down the firm productivity ladder, inventors experience a growth in earnings.

(iii) Non-compete contracts are associated with increases in firms' RD expenditure.

(iii) Non-compete contracts are associated with decreases in executives' mobility rates.

Worker mobility and knowledge diffusion

Event study. I use an event study framework to document how knowledge diffuses when workers change jobs. I define an event as job-to-job mobility: a worker moves from an origin firm (o) to a destination firm (d) in the same or the subsequent quarter. The event study takes the form:

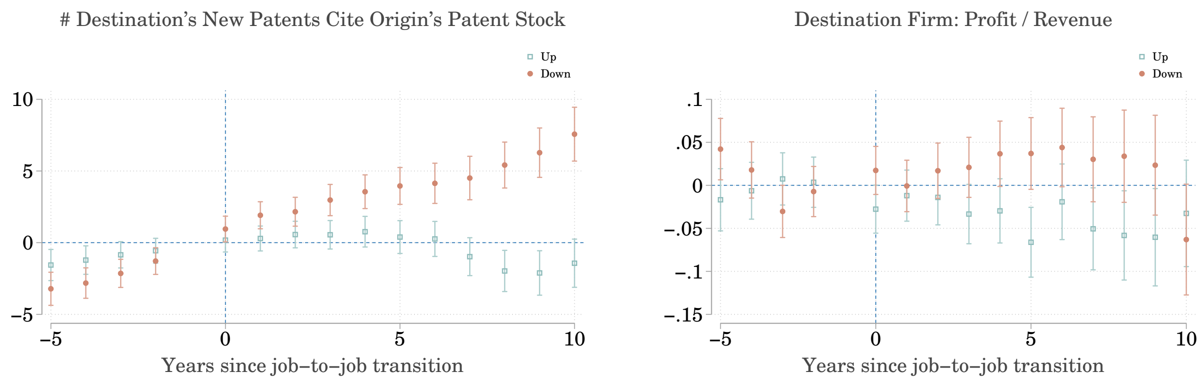
$$y_{od,t} = \sum_{\tau=-3,\tau\neq-1}^{5} \beta_\tau \mathbb{1}\{\text{Mobility}\}_{od,\tau} + \lambda_{od} + \delta_t + \varepsilon_{od,t}$$

On the left-hand side, $y_{jk,t}$ represents knowledge diffusion from origin firm o to destination firm d in calendar year t. Specifically, knowledge diffusion is measured by the number of destination firms' new patents that cite the pre-mobility patent stock of the origin firm. On the right-hand side, $\mathbb{1}\{\text{Mobility}\}_{od,\tau}$ is an indicator for time relative to the mobility event. The indicator $\mathbb{1}\{\text{Mobility}\}_{od,\tau}$ takes value one if mobility occurred τ years before calendar year t. Coefficient β_τ is of key interest. It estimates the dynamic impact of mobility τ years after the move, compared to one year before. The model also includes fixed effects for each pair of origin and destination firms, λ_{od}, and calendar year effects δ_t.

The primary advantage of an event study is that it allows us to visually and flexibly trace knowledge flows around the time of worker mobility. Figure 1.3 reports the event study coefficients

β_τ on event time $\tau \in \{-3, -2, ..., +5\}$. I present the coefficients separately for two groups of workers: inventors on the left panel and executives on the right panel. I distinguish two directions of worker mobility, downward and upward, along the firm productivity ladder. The productivity ladder is ranked on the three-year moving average of productivity before mobility. In particular, I classify job-to-job mobility as downward if the three-year average productivity of destination firms is lower than that of the origin firm.

Figure 1.3: Inter-Firm Patent Citations



(a) Executive Mobility

(b) Profit Rate

Inventors diffuse knowledge and are compensated by knowledge diffusion.

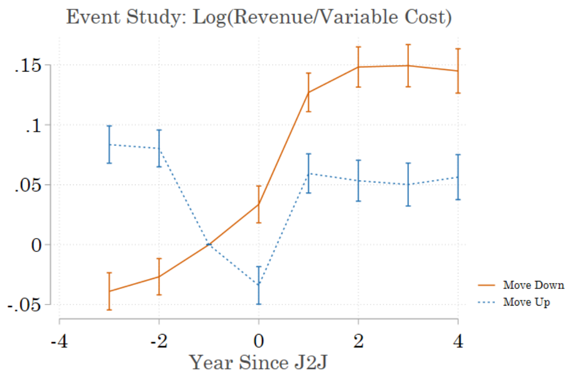
$$\log(\text{TFPR})_{d,t} = \sum_{\tau=-3,\tau\neq-1}^{4} \beta_\tau \mathbb{1}\{\text{Hire Inventor}\}_{d,t-\tau} + \lambda_d + \delta_t + \varepsilon_{d,t}$$

$$\log(\text{Quarterly Earnings})_{i,qrt(t)} = \sum_{\tau=-3,\tau\neq-1}^{4} \beta_\tau \mathbb{1}\{\text{Join Destination}\}_{i,t-\tau} + \lambda_i + \delta_t + \varepsilon_{i,t}$$

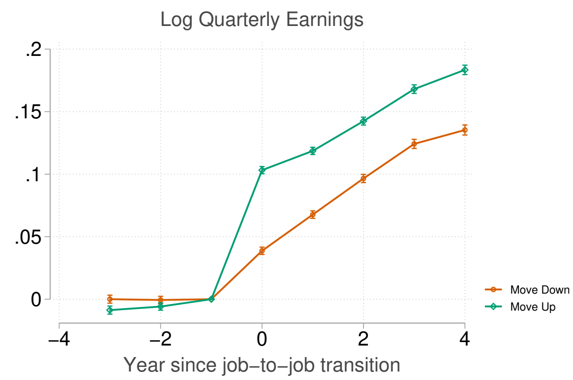
Innovation-diffusion tradeoff of non-compete contracts

Non-compete contracts protect intangibles, substitute for patents.

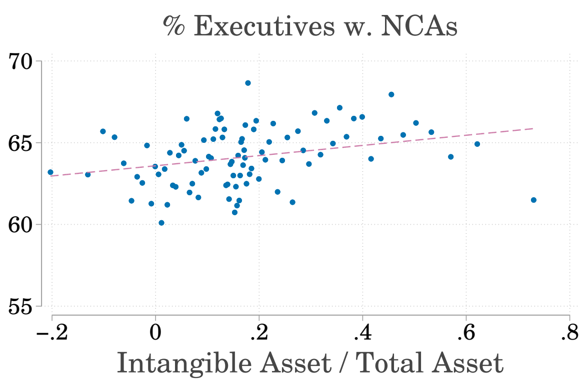
Non-compete contracts encourage R&D.



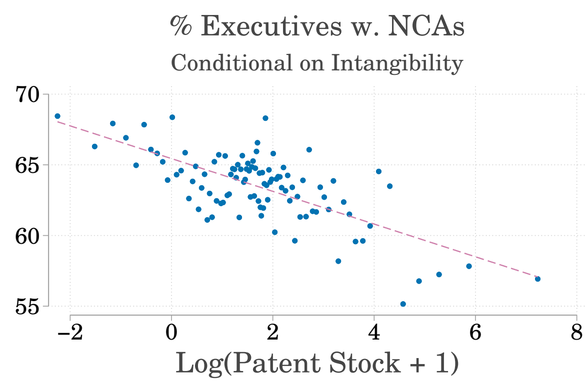
(a) Patent Citation



(b) Profit Rate



(a) Intangibility

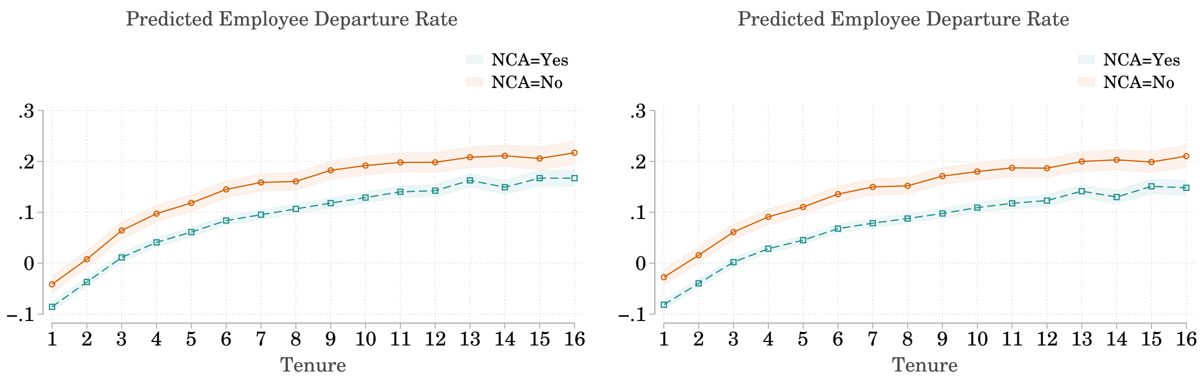


(b) Patent Stock

	Log(R&D)	Log(R&D+1)	R&D ($MM)
NCA $\in [0,1]$	0.080	0.045	37.199
	(0.023)	(0.016)	(13.125)
Lag3. Log(Patent Stock+1)	-4e-5	-2e-5	0.192
	(5e-6)	(4e-6)	(0.004)
NCA $\times$ Lag3. Log. Patent	-2e-5	-2e-5	-0.063
	(6e-6)	(4e-6)	(0.004)
Firm FEs	Y	Y	Y
Year FEs	Y	Y	Y
Obs.	24,884	31,537	31,537

Non-compete contracts discourage job-to-job mobility.

$$\mathbb{1}\{\text{EE Move}\}_{ijt} = \alpha_1\, \mathbb{1}\{\text{NCA}\}_{ijt} \times \text{Tenure}_{ijt} + \alpha_2\, \mathbb{1}\{\text{NCA}\}_{ijt} + \alpha_3\, \text{Tenure}_{ijt}$$
$$+\alpha_4\, \mathbb{1}\{\text{Job Title}\}_{ijt} + \text{Worker}_i\ \text{FE} + \text{Firm}_j\ \text{FE} + \text{Year}_t\ \text{FE} + \varepsilon_{ijt}$$



(a) Separation

(b) Job-to-Job Mobility

1.4 Quantitative Analysis

Having shown that the model's key predictions align qualitatively with the data, we now proceed to quantify the model. Section 1.4 calibrates the model along a BGP equilibrium to match moments from the matched firm-inventor data at the U.S. Census Bureau.

Calibration

Parameterization. I specify the following functional forms. The innovation cost function is $\kappa_{r\&d}(\mu) = \tilde{\kappa}_{r\&d}\mu^{\gamma}$. The matching function is Cobb-Douglas with vacancy elasticity α, i.e., $m(1, \theta_t) = A\theta_t^{\alpha}$. Therefore, a vacant firm meets a worker at rate $\lambda(\theta_t) = A\theta_t^{-(1-\alpha)}$, and a worker meets a vacant firm at rate $\theta_t\lambda(\theta_t) = A\theta_t^{\alpha}$. The entrant knowledge draw follows the Pareto distribution $\mathcal{F}_e(Z, t) = 1 - (e^{-gt}Z)^{\zeta_e}$. The shape parameter ζ_e is constant. The minimum of the support grows as the economy grows and can be normalized to one. I add exogenous exit rate δ_v for vacant firms.

External calibration. I set the discount rate to $\rho = 0.05$. Together with the calibrated growth rate, this rate gives a long-run interest rate of 6%, a reasonable value for the U.S. economy (Cooley, 1995). I assume a quadratic innovation cost function: the R&D scale elasticity $\gamma = 2$. The R&D cost elasticity γ governs firms' sensitivity to R&D returns. Credibly identifying the elasticity parameter is difficult without exogenous variation in innovation cost. I therefore follow Acemoglu, Akcigit, Alp, Bloom, and Kerr (2018) and assume $\gamma = 2$. I set the elasticity of the matching function is to

$\alpha = 0.5$ following Petrongolo and Pissarides (2001). The Pareto shape parameter ζ_e of entrant distribution is pinned down the mean of entrant productivity.

Internal calibration. The remaining parameters are calibrated jointly. I simulate a sample of firms and workers, innovation decisions, worker mobility, and their resulting knowledge evolution. I apply the simulated method of moments (SMM), minimizing the objective function

$$\Big(\boldsymbol{m}(\phi) - \hat{\boldsymbol{m}}\Big)' \Omega^{-1} \Big(\boldsymbol{m}(\phi) - \hat{\boldsymbol{m}}\Big) \quad , \quad \text{where} \quad \phi := \Big\{ p,\ A,\ \tilde{\kappa}_{r\&d},\ \kappa_{entry},\ \delta_\nu \Big\},$$

where $\boldsymbol{m}(\phi)$ is a vector of model-simulated moments and $\hat{\boldsymbol{m}}$ are their data counterpart. The matrix Ω contains squares of the data moments on the main diagonal and zeros elsewhere.

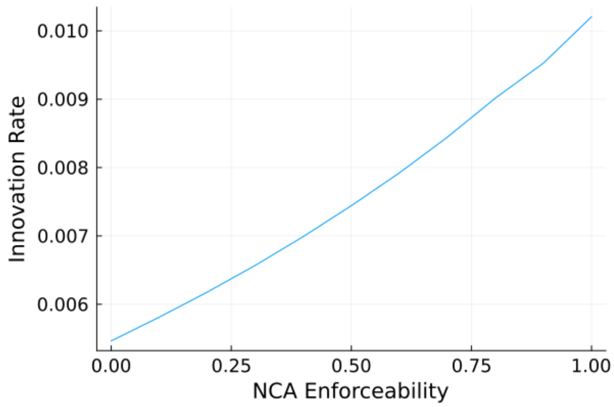
While the parameters are jointly calibrated, I provide a heuristic discussion of the most relevant moments for each parameter. The learning probability is set to match the 3-year average change in $\log(\mathsf{TFPR})$ since hiring an inventor from more productive firms, relative to the $\log(\mathsf{TFPR})$ before hiring. The annual separation rate among inventors infers the matching efficiency.

Parameter	Definition	Value	Moment	Model	Data
p	Learning probability	0.11	$\mathbb{E}[\Delta \log(\mathsf{TFPR}_t)]$ since hiring inventor	9.68%	10.30%
			% Down-the-ladder mobility	42.20%	48.69%
			Annual TFP growth rate	1.06%	1.01%
			Meeting function: $m(1,\theta) = A\theta^{\alpha}$		
α	Matching elasticity	0.50	Petrongolo & Pissarides (2001)	–	–
A	Meeting efficiency	0.25	Annual mobility rate	21.74%	21.81
κ_{entry}	Entry cost	59.45	# Years to hire 1st inventor	6.29	6.05
			Innovation cost: $\kappa_{r\&d}(\mu) = \tilde{\kappa}_{r\&d}\mu^{\gamma}$		
γ	Elasticity	2.00	Acemoglu et. al (2018)	–	–
$\tilde{\kappa}_{r\&d}$	Scale	896	Aggregate R&D / GDP	3.13%	2.67%
ρ	Discount rate	0.05	Cooley et. al (1995)	-	-
ζ_{entry}	Pareto shape of entry dist.	1.14	Mean $\log(\mathsf{TFPR}_t)$ of entrant	0.88	0.88
δ_ν	Exogenous exit rate	0.02	Size adjusted entry rate	1.87%	1.96%
κ_{scrap}	Scrap value (voluntary exit)	18.80	Smooth pasting condition	-	-

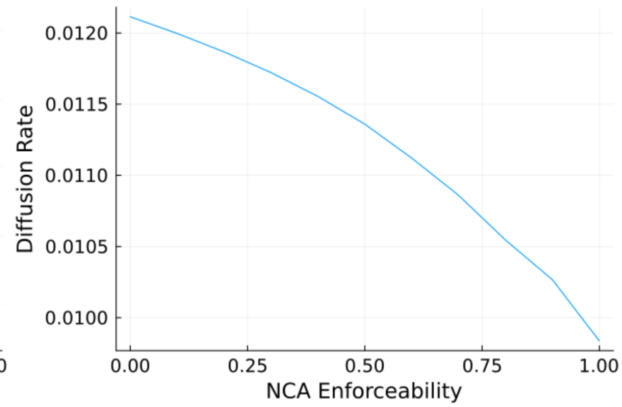
Table 1.2: Externally-calibrated parameters

Comparative Statics. NCA enforcement probability encourages innovation but discourages knowledge diffusion.

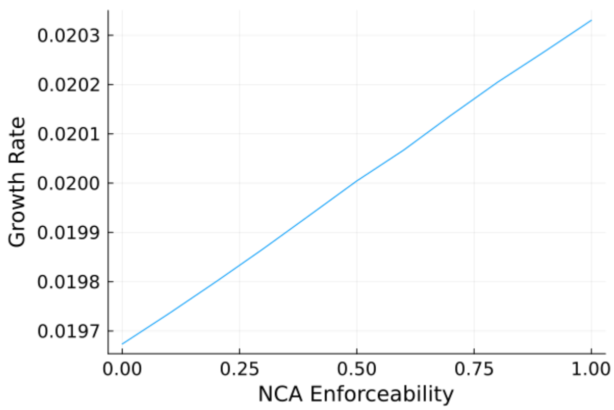
Comparative Statics. Higher NCA enforcement probability is associated with higher aggregate growth rate.



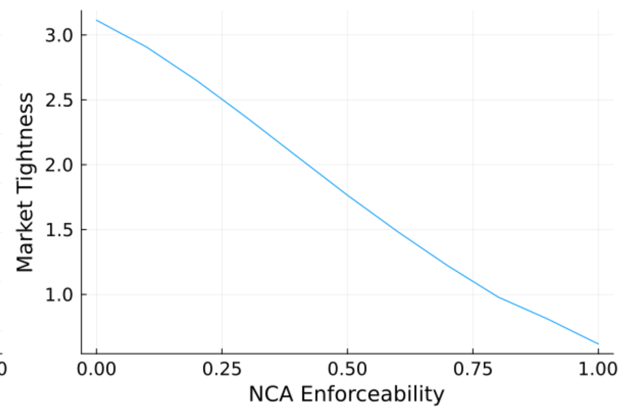
(a) Innovation ↑



(b) Diffusion ↓



(a) Growth Rate



(b) Market Tightness

1.5 Conclusion

This paper contributes to our understanding of how the labor market affects economic growth. First, it offers evidence that worker mobility is an empirically important channel of knowledge diffusion. Second, the paper integrates on-the-job search to an endogenous model of innovation and knowledge diffusion. In contrast with the canonical models that characterize knowledge diffusion as exogenous social learning, this model endogenizes the knowledge diffusion process with endogenous worker mobility. Knowledge diffusion is determined by the voluntary trade of workers and the tightness of labor market. The novel feature, which highlights worker mobility as a channel for knowledge diffusion, yields new implications on the joint dynamics of firms and workers. Worker mobility shapes the evolution of firm productivity and is influenced by the equilibrium distribution of firm productivity. Moreover, the model implies a knowledge adoption motive for poaching workers. And as knowledge is nonrival, workers voluntarily move down the firm productivity ladder.

Third, the paper proposes a theory and constructs new data on non-compete contracts. Non-compete contracts enable a firm to internalize the social returns to innovation. The enforceability of non-competes governs how incumbent and entrant firms divide the surplus from trading workers.

Higher enforceability encourages incumbent innovation but discourages knowledge diffusion by deterring entry. The optimal regulation of non-compete contracts balances the innovation-diffusion trade-offs.

1.6 Appendix

Numeric Algorithm

Discretization. Abusing notation, define the discretized state space, value functions, and distributions

- An equi-spaced grid with n points on the interior, $z := \{z_i\}_{i=1}^n$, and grid spacing Δ_z.
- Vectors for value functions (1.7) (1.8) at grid points: $M := \{M_i\}_{i=1}^n$, $V := \{V_i\}_{i=1}^n$.
- Vectors for probability density functions (**??**) (**??**) at grid points: $f_m := \{f_{m,i}\}_{i=1}^n$, $f_v := \{f_{v,i}\}_{i=1}^n$.

Stationary Equilibrium Algorithm.

- Construct initial guess: value functions $\{M^0, V^0\}$, probability density functions $\{f_m^0, f_v^0\}$, market tightness θ^0, and aggregate growth rate g^0. As the measure of matches is normalized to one, the measure of vacant firms $n_v^0 = \theta^0$.
- Given aggregate growth rate g^t, iterate to convergence on $\{M^t, V^t\}$, $\{f_m^t, f_v^t\}$, θ^t.
  1. Given g^t, $\{f_m^\tau, f_v^\tau\}$, θ^τ, solve HJB equations to obtain $\{M^{\tau+1}, V^{\tau+1}\}$.
  2. Given g^t, $\{M^{\tau+1}, V^{\tau+1}\}$, $\{f_m^\tau, f_v^\tau\}$, solve free entry equation to obtain $\theta^{\tau+1}$.
  3. Given g^t, $\{M^{\tau+1}, V^{\tau+1}\}$, $\theta^{\tau+1}$, solve KF equations to obtain $\{f_m^{\tau+1}, f_v^{\tau+1}\}$.
  4. Check whether $\{M^{\tau+1}, V^{\tau+1}\}$, $\theta^{\tau+1}$, and $\{f_m^{\tau+1}, f_v^{\tau+1}\}$ have converged. If not, go back to 1.
- Given individual behavior, iterate to convergence on aggregate growth rate
  1. Given $\{M^t, V^t\}$, $\{f_m^t, f_v^t\}$, θ^t, update growth rate g^{t+1}

Solving HJB Equations

Discretized HJB Equations

Stack discretized value functions M, V, and U into vector R:

$$R = \begin{pmatrix} M_1 \\ \vdots \\ M_n \\ V_1 \\ \vdots \\ V_n \end{pmatrix}.$$

Let τ be the iteration of the algorithm. The HJB equations (1.7) (1.8) can be jointly discretized as:

$$(r-g)\, R^{\tau+1} = \pi^{\tau} + A^{\tau} R^{\tau+1} \tag{1.9}$$

Vector of flow value π^{τ} is given by:

$$\pi^{\tau} = \begin{pmatrix} e^{z_1} - C_{r\&d}(\mu_1^{\tau}) e^{z_1} \\ \vdots \\ e^{z_n} - C_{r\&d}(\mu_n^{\tau}) e^{z_n} \\ e^{z_1} - \kappa_v \\ \vdots \\ e^{z_n} - \kappa_v \end{pmatrix}.$$

Linear operator A^{τ} can be decomposed into four additive components, describing the evolution of firm knowledge due to innovation (A_1^{τ}), worker mobility (A_2^{τ}), EU transition (A_3^{τ}), and exogenous exit (A_4^{τ}). The construction of A^{τ} will be detailed in Section 1.6.

Boundary Conditions

The HJB equations are subject to boundary conditions of endogenous separation and exit. To incorporate the value-matching and smooth-pasting conditions at the boundary, I rewrite the HJB equation (1.9) in terms of an HJB variational inequality (HJBVI):

$$\min\left\{[(r-g)\, I - A^{\tau}]\, R^{\tau+1} - \pi^{\tau}\ ,\ B^{-1} R^{\tau+1} - \underline{R}\right\} = 0,$$

$$B^{-1} = \begin{pmatrix} I_{n\times n} & -I_{n\times n} \\ 0 & I_{n\times n} \end{pmatrix} \quad , \quad B = \begin{pmatrix} I_{n\times n} & I_{n\times n} \\ 0 & I_{n\times n} \end{pmatrix} \quad , \quad \underline{R} = \begin{pmatrix} [\underline{M}]_{n\times 1} \\ [\underline{V}]_{n\times 1} \end{pmatrix}$$

Define $X := B^{-1} R^{\tau+1} - \underline{R}$. The HJBVI equation can be written as

$$\min\left\{[(r-g)\, I - A^{\tau}]\, BX + [(r-g)\, I - A^{\tau}]\, B\underline{R} - \pi^{\tau}\ ,\ X\right\} = 0,$$

and can be solved as a linear complementarity problem (LCP)

$$\begin{aligned} X^{T}\left([(r-g)\, I - A^{\tau}]\, BX + [(r-g)\, I - A^{\tau}]\, B\underline{R} - \pi^{\tau}\right) &= 0 \\ [(r-g)\, I - A^{\tau}]\, BX + [(r-g)\, I - A^{\tau}]\, B\underline{R} - \pi^{\tau} &\geqslant 0 \\ X &\geqslant 0 \end{aligned} \tag{1.10}$$

Discretization with Upwind Finite-Differences

To compute the HJB equations (1.7) (1.8), I need to approximate the differential operator numerically. I use the upwind finite difference method following Achdou, Han, Lasry, Lions, and Moll (2022). The idea is to use the forward difference approximation whenever the drift of the state variable is positive, and the backward difference approximation whenever it is negative. For joint value M_i of a firm-worker match, define

$$\textit{Forward difference}: \qquad \partial_{z,F} M_i := \frac{M_{i+1} - M_i}{\Delta_z}$$

$$\textit{Backward difference}: \qquad \partial_{z,B} M_i := \frac{M_i - M_{i-1}}{\Delta_z}$$

The drift $\mu - g$ of state variable is derived from the dynamic decision rule (1.9) and approximated with

$$\textit{Forward}: \qquad \mu_{F,i} - g = (c')^{-1}\left(e^{-z_i}\,\partial_{z,F} M_i\right) - g$$

$$\textit{Backward}: \qquad \mu_{B,i} - g = (c')^{-1}\left(e^{-z_i}\,\partial_{z,B} M_i\right) - g$$

Define the forward and backward Hamiltonians:

$$\textit{Forward}: \qquad H_{F,i} = -c(\mu_{F,i}) + e^{-z_i}\,(\mu_{F,i} - g)\,\partial_{z,F} M_i$$

$$\textit{Backward}: \qquad H_{B,i} = -c(\mu_{B,i}) + e^{-z_i}\,(\mu_{B,i} - g)\,\partial_{z,B} M_i$$

Using the upwind scheme, the derivative of the value function is approximated with [11]

$$\partial_z M_i \approx \qquad \partial_{z,F} M_i\,\mathbb{1}_{F,i} + \partial_{z,B} M_i\,\mathbb{1}_{B,i} + e^{z_i} c'(g)\,\mathbb{1}_{C,i} \tag{1.11}$$

where $\mathbb{1}_{\{\bullet\}}$ denotes the indicator function, $\mathbb{1}_{F,i}$, $\mathbb{1}_{B,i}$, $\mathbb{1}_{C,i}$ are indicators given by:

$$\begin{aligned}
\textit{Forward}: \quad \mathbb{1}_{F,i} &= \quad \mathbb{1}_{\{\mu_{F,i} > g\}}\mathbb{1}_{\{\mu_{B,i} > g\}} + \mathbb{1}_{\{H_{F,i} \geqslant H_{B,i}\}}\mathbb{1}_{\{\mu_{F,i} \geqslant g\}}\mathbb{1}_{\{\mu_{B,i} \leqslant g\}} \\
\textit{Backward}: \quad \mathbb{1}_{B,i} &= \quad \mathbb{1}_{\{\mu_{F,i} < g\}}\mathbb{1}_{\{\mu_{B,i} < g\}} + \mathbb{1}_{\{H_{F,i} < H_{B,i}\}}\mathbb{1}_{\{\mu_{F,i} \geqslant g\}}\mathbb{1}_{\{\mu_{B,i} \leqslant g\}} \\
\textit{Central}: \quad \mathbb{1}_{C,i} &= \quad \mathbb{1}_{\{\mu_{F,i} \leqslant g\}}\mathbb{1}_{\{\mu_{B,i} \geqslant g\}}
\end{aligned}$$

The optimal innovation decision is

$$\mu_i^{\tau} = (c')^{-1}\left(e^{-z_i}\,\partial_z M_i\right) \tag{1.12}$$

For the second-order derivative, I use a central difference approximation:

$$\partial_{zz} M_i \approx \frac{M_{i+1} - 2M_i + M_{i-1}}{(\Delta_z)^2}$$

In HJB equation (1.8), the drifts $(-g)$ are negative. I calculate the approximate $\partial_z V_i$ with backward difference operator:

$$\partial_z V_i \approx \frac{V_i - V_{i-1}}{\Delta_z}.$$

[11] For notational simplicity, define the differential operator ∂ such that $\partial_z = \frac{\partial}{\partial z}$ and $\partial_{zz} = \frac{\partial^2}{\partial z^2}$.

Linear operator.

The linear operator A^{τ} has four components, describing the evolution of firm knowledge due to innovation (A_1^{τ}), worker mobility (A_2^{τ}), exogenous separation (A_3^{τ}), and exogenous exit (A_4^{τ}).

$$A^{\tau} = A_1^{\tau} + A_2^{\tau} + A_3^{\tau} + A_4^{\tau} = \begin{bmatrix} A_{1mm}^{\tau} + \beta\, A_{2mm}^{\tau} + A_{3mm}^{\tau} & \beta\, A_{2mv}^{\tau} - A_{3mm}^{\tau} \\ (1-\beta)A_{2vm}^{\tau} - A_{3vv}^{\tau} & A_{1vv}^{\tau} + (1-\beta)A_{2vv}^{\tau} + A_{3vv}^{\tau} + A_{4vv}^{\tau} \end{bmatrix}$$

Linear operator: innovation.

$$\textit{Innovation :} \qquad A_1^{\tau} = \begin{bmatrix} A_{1mm}^{\tau} & 0 \\ 0 & A_{1vv}^{\tau} \end{bmatrix}$$

$$A_{1mm}^{\tau} = [columns - width = 1.2cm] a_{B,1} + a_{C,1} a_{F,1} 0 \dots 0 a_{B,2} a_{C,2} a_{F,2} \dots 0 \vdots \ddots \ddots \ddots \vdots 0 \dots a_{B,n-1} a_{C,n-1} a_{F,n-1} 0 \dots$$

where for $\forall\ i = 1, 2, 3, ..., n$:

$$\begin{aligned} a_{F,i} &= \frac{(\mu_{F,i} - g)\, \mathbb{1}_{F,i}}{\Delta_z} + \frac{\sigma_m^2}{2\Delta_z^2} \\ a_{C,i} &= -\frac{(\mu_{F,i} - g)\, \mathbb{1}_{F,i}}{\Delta_z} + \frac{(\mu_{B,i} - g)\, \mathbb{1}_{B,i}}{\Delta_z} - \frac{\sigma_m^2}{\Delta_z^2} \\ a_{B,i} &= -\frac{(\mu_{B,i} - g)\, \mathbb{1}_{B,i}}{\Delta_z} + \frac{\sigma_m^2}{2\Delta_z^2} \end{aligned}$$

Linear operator: worker mobility.

$$\textit{Worker Mobility :} \qquad A_2^{\tau} = \begin{bmatrix} \beta\, A_{2mm}^{\tau} & \beta\, A_{2mv}^{\tau} \\ (1-\beta)A_{2vm}^{\tau} & (1-\beta)A_{2vv}^{\tau} \end{bmatrix}$$

$$A_{2mm}^{\tau} = \lambda(\theta)\theta[columns - width = 2cm] - \sum_{j=2}^{n} \widetilde{d}_{1,j}\, f_{v,j}\, \widetilde{d}_{1,2}\, f_{v,2} \dots \widetilde{d}_{1,n}\, f_{v,n} (1-p)\widetilde{d}_{2,1}\, f_{v,1} - (1-p) \sum_{j=1}^{1} \widetilde{d}_{2,j}\, f_{v,j} - \sum_{j=3}^{n}$$

$$A_{2mv}^{\tau} = \lambda(\theta)\theta[columns - width = 1cm](1-\tau) \sum_{j=1}^{n} \widetilde{d}_{1,j}\, f_{v,j} - \widetilde{d}_{1,2}\, f_{v,2} - \widetilde{d}_{1,3}\, f_{v,3} \dots - \widetilde{d}_{1,n}\, f_{v,n} - \widetilde{d}_{2,1}\, f_{v,1} (1-\tau) \sum_{j=1}^{n} \widetilde{d}_{2}$$

where indicator $\widetilde{d}_{i,j} := \mathbb{1}\left\{p\ M_{\max\{i,j\}}^{\tau} + (1-p)M_j^{\tau} - M_i^{\tau} - V_j^{\tau} + (1-\tau)V_i^{\tau} > 0\right\}$.

$$A_{2vm}^{\tau} = \lambda(\theta)[columns-width = 2cm](1-p)\sum_{j=2}^{n}\widehat{d}_{1,j}\, f_{m,j} - (1-p)\widehat{d}_{1,2}\, f_{m,2} \ldots - (1-p)\widehat{d}_{1,n}\, f_{m,n} - \widehat{d}_{2,1}\, f_{m,1}\sum_{j=1}^{1}\widehat{d}_{2,$$

$$A_{2vv}^{\tau} = \lambda(\theta)[columns-width = 2.7cm] - \sum_{j=1}^{n}\widehat{d}_{1,j}\, f_{m,j}(1-\tau)\widehat{d}_{1,2}\, f_{m,2}(1-\tau)\widehat{d}_{1,3}\, f_{m,3} \ldots (1-\tau)\widehat{d}_{1,n}\, f_{m,n}(1-\tau)\widehat{d}_{2,1}\, f$$

where indicator $\widehat{d}_{i,j} := \widetilde{d}_{j,i} = \mathbb{1}\left\{pM^{\tau}_{\max\{i,j\}} + (1-p)M^{\tau}_i - M^{\tau}_j - V^{\tau}_i + (1-\tau)V^{\tau}_j > 0\right\}$.

Linear operator: exogenous separation and quit.

$$\textit{Exogenous Separation :} \qquad A_3^{\tau} = \begin{bmatrix} A_{3mm}^{\tau} & -A_{3mm}^{\tau} \\ -A_{3vv}^{\tau} & A_{3vv}^{\tau} \end{bmatrix} \quad , \quad \textit{Exogenous Quit :} \qquad A_4^{\tau} = \begin{bmatrix} 0 & 0 \\ 0 & A_{4vv}^{\tau} \end{bmatrix}$$

$$A_{3mm}^{\tau} = \begin{pmatrix} -\delta_m & 0 & \ldots & 0 \\ 0 & -\delta_m & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -\delta_m \end{pmatrix} , \quad A_{3vv}^{\tau} = \begin{pmatrix} -\frac{\delta_m}{\theta} & 0 & \ldots & 0 \\ 0 & -\frac{\delta_m}{\theta} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -\frac{\delta_m}{\theta} \end{pmatrix} , \quad A_{4vv}^{\tau} = \begin{pmatrix} -\delta_v & 0 & \ldots & 0 \\ 0 & -\delta_v & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -\delta_v \end{pmatrix} .$$

Summary of the algorithm.

To sum up, the algorithm for finding a solution to the HJB equation is as follows. Make an initial guess $R^0 := \left\{\{M_i^0\}_{i=1}^n, \{V_i^0\}_{i=1}^n, \{U_i^0\}_{i=1}^n\right\}$. For each iteration $\tau = 0, 1, 2, ...$, follow

1. Compute $\partial_z M_i$ using (1.11)

2. Compute μ^{τ} from (1.12)

3. Find $R^{\tau+1}$ from (1.10)

4. If $R^{\tau+1}$ is close enough to R^{τ}: stop. Otherwise, go to step 1.

Solving KF Equations

Upwind finite difference. To compute the KF equations (**??**) (**??**), I again use the upwind finite difference method. Stack the discretized probability density functions f_m, f_v into vector f:

$$f = \begin{pmatrix} f_{m,1} \\ \vdots \\ f_{m,n} \\ f_{v,1} \\ \vdots \\ f_{v,n} \end{pmatrix}.$$

Let Δ be step size, τ be iteration of the algorithm, x be the dampening parameter,

$$\begin{aligned} D^{\tau}\ f^{\tau+1} &= 0 \\ \widehat{f}^{\tau+1} &= x f^{\tau+1} + (1-x) f^{\tau} \\ D^{\tau+1} &= D(\widehat{f}^{\tau+1}) \end{aligned} \tag{1.13}$$

Transition matrices. Following the construction of A^{τ}, transition matrix D^{τ} has three components, describing the law of motions due to innovation (D_1^{τ}), EE transition (D_2^{τ}), and EU transition (D_3^{τ}).

$$D^{\tau} = D_1^{\tau} + D_2^{\tau} + D_3^{\tau} + D_4^{\tau} = \begin{bmatrix} A_{1mm}^{\tau} + \beta\, A_{2mm}^{\tau} + A_{3mm}^{\tau} & \beta\, A_{2mv}^{\tau} - A_{3mm}^{\tau} \\ (1-\beta) A_{2vm}^{\tau} & A_{1vv}^{\tau} + (1-\beta) A_{2vv}^{\tau} + A_{4vv}^{\tau} \end{bmatrix}^{\mathsf{T}}.$$

The transition matrices for innovation is the transpose of the derivative matrices from the HJB equation, i.e., $D_{1mm}^{\tau} = \left(A_{1mm}^{\tau}\right)^{\mathsf{T}}$, $D_{1uu}^{\tau} = D_{1vv}^{\tau} = \left(A_{1vv}^{\tau}\right)^{\mathsf{T}}$.

Solving Market Tightness

$$\begin{aligned} &\left[\rho I - \left(\frac{N_v^{\tau}}{N_v^{\tau+1}}\right)^{\vartheta} (A_{1vv}^{\tau} + (1-\beta) A_{2vv}^{\tau} + (1-\beta_u) A_{3vv}^{\tau} + A_{4vv}^{\tau})\right] V^{\tau+1} \\ &= \widetilde{\pi} + \left(\frac{N_v^{\tau}}{N_v^{\tau+1}}\right)^{\vartheta} \left[(1-\beta_0) A_{2vm}^{\tau} + (1-\beta_u) A_{3vm}^{\tau}\right] M^{\tau} + (1-\beta_u) A_{3vu}^{\tau} U^{\tau} \end{aligned}$$

$$(f_e)^{\mathsf{T}} V^{\tau+1} = \kappa_e$$

2 STRATEGIC RESTRAINT: WHEN DO HUMAN-CAPITAL-INTENSIVE COMPANIES CHOOSE (NOT) TO USE NONCOMPETE AGREEMENTS?

2.1 Introduction

Exploring the role of restrictive legal levers in the management of knowledge and human capital is one of the traditional questions studied by scholars in strategic management. Firms often use restrictive clauses in employment contracts and patent enforcement to lower the risk of expropriation of their valuable knowledge by competitors (Kim & Marschke, 2005; Agarwal, Ganco, & Ziedonis, 2009; Starr, Balasubramanian, & Sakakibara, 2018a). Since employee mobility drives knowledge spillovers across firms (Rosenkopf & Almeida, 2003; Fallick, Fleischman, & Rebitzer, 2006; Tzabbar, 2009; Kim & Steensma, 2017; Sevcenko & Ethiraj, 2018; Tzabbar & Cirillo, 2020), management of knowledge outflows frequently focuses on human capital.

In this context, noncompete agreements (NCAs) have been studied extensively as a way to reduce mobility to rivals (Marx, 2011; Marx, Strumsky, & Fleming, 2009), lower the likelihood that employees found competing startups (Starr, Balasubramanian, & Sakakibara, 2018a), reduce knowledge spillovers to competing firms (Marx, Singh, & Fleming, 2015), improve value appropriation (Younge & Marx, 2016; Starr, Ganco, & Campbell, 2018b; Starr, Prescott, & Bishara, 2021), and enhance collaboration between workers (Seo & Somaya, 2021).[1] Further, NCAs incentivize firms to invest in employees' human capital through training (Meccheri, 2009; Starr, 2019), share confidential information (Garmaise, 2011), and pursue riskier R&D projects (Conti, 2014). Given their practical relevance and theoretical interest, NCAs have become a poster child for the literature on the management of knowledge flows through mobility.

While NCAs can be detrimental to employees because they restrict the employees' outside options, the value of NCAs for firms has been generally viewed as positive (Marx, Strumsky, & Fleming, 2009; Younge & Marx, 2016; Starr *et al.*, 2018a, b). That is hardly surprising given that the laws that enable firms to use NCAs are intended to help firms retain their intellectual property and encourage investments associated with stronger property rights (Alchian & Demsetz, 1973; Ostrom & Hess, 2008). A question not addressed in prior work is, if NCAs are beneficial to firms, why don't all firms use them?[2] Do NCAs have downsides for some firms using them? Which firms *choose not to use NCAs* and why?

Our study aims to explore firms' choice to opt-out of NCA use even when the NCAs are used

[1] A noncompete agreement is a clause in an employment contract under which an employee agrees not to join a competitor or start a competing firm after exit from the focal employment, usually for a pre-determined amount of time and over a geographical area (Rubin & Shedd, 1981; Garmaise, 2011; Samila & Sorenson, 2011; Conti, 2014; Starr, Balasubramanian, & Sakakibara, 2018a).

[2] There is limited evidence on the firm-level use of NCAs. Starr et al., (2021) estimated that around 18% of the entire US workforce is subject to NCAs.

within the focal industry. By doing so, we introduce a novel explanation of the heterogeneity of NCA use among competitors. There is emerging evidence about the heterogeneity of NCA use across workers. For instance, Starr *et al.* (2021) surveyed workers and found that many workers are not bound by NCAs even when such covenants are popular among firms in the focal industry and enforceable in the focal state. They also report that being bound by an NCA positively correlates with the worker's human capital and employer size.[3] However, together with the work examining the effects of state-level variation in the NCA legislation and its enforceability (Marx *et al.*, 2009; Younge & Marx, 2016; Starr *et al.*, 2018a, 2018b, 2021; Seo & Somaya, 2021), these studies leave open the question of the reasons behind this firm-level heterogeneity.

We develop a theoretical framework describing when firms opt out of using NCAs. The core of our argument is that while NCAs may improve employee retention, they may also lower the firm's ability to attract workers. If a firm is less concerned about knowledge expropriation by exiting workers while it still needs to attract talented workers (e.g., because talent constitutes its core resource), it may decide not to use NCAs, even when NCAs are enforceable and used by competitors. Our theory posits that even within industries and states, there may be a significant firm-level heterogeneity in NCA use due to the tension between talent attraction and retention.

We utilize a novel survey implemented in collaboration with PayScale, a US data and software company focusing on compensation analytics. PayScale routinely collects employment-related data from its client firms. We attached the NCA-related questions to their periodic firm-level survey (the respondents are primarily human resource managers and executives). The survey allowed us to observe the actual use of NCAs at the firm level, and we also asked questions about the reasons for opting out of NCAs.

Descriptively, our data reveal significant heterogeneity in the use of NCAs. On average, within industries, around 31% of the surveyed firms use NCAs for all workers, 40% for some workers, and 28% for none. We also find that among firms using NCAs for none or some of their workers, 15% of them report not using NCAs (at least for some workers) because doing so makes it harder to attract talented workers. This percentage increases to 30-35% in human capital-intensive industries such as 'architecture and engineering' or 'PR and marketing.'

In a regression analysis, where we condition on industry and state, firms that rank talent as the key resource differentiating them from competitors (relative to other resources) *and* are not the leading firms in their respective industries are more likely *not* to use NCAs.[4] These firms are

[3]Other studies that focus on individuals are Colvin and Shierholz (2019), who provide descriptive information about the use of NCAs, and Johnson and Lipsitz (2022), who study NCA use in the context of hair salons.

[4]We conceptualize leading vs. non-leading firms as a general proxy for within-industry differences in quality. Leading firms are more likely to have resources that are valuable to competitors and are subject to expropriation risks. These resources include knowledge or social capital embedded in employees. Similarly, we conceptualize the relative reliance on talent as a general proxy for reliance on skilled human capital relative to other resources. While these variables are based on survey questions, we employ validation and robustness tests using external datasets such as Compustat and CPS for a subsample of observations where a match is possible. The results hold irrespective of whether we control for

more likely to report that they do not use NCAs precisely because they would hurt their ability to attract talented workers. In addition, we find that opting out of NCA use among firms that rely on talent is associated with easier filling of high-skill technical positions for non-leading firms relative to leaders. NCA use is thus more penalizing when attracting technical workers for non-leaders relative to leaders when their reliance on talent is high. Further, we observe that opting out of NCA use as a differentiation strategy is more pronounced in industries that rely less on patents, which further underscores knowledge leakage concerns as an underlying mechanism.

In a supplemental analysis, we show that firms opting out of NCAs are likely to use other practices to retain workers (Huselid, 1995; Foss, Pedersen, Fosgaard, & Stea, 2015). To assess the sensitivity of our results to alternative explanations, we employ a range of robustness tests, including recently developed tests for the sensitivity of regression estimates to omitted variable bias (Cinelli & Hazlett, 2020; Cinelli, Ferwerda, & Hazlett, 2020). We conclude from these tests that our findings result from firms making strategic choices as opposed to patterns previously reported across states, industries, or individuals (Marx *et al.*, 2009; Starr *et al.*, 2018a, b).

The study has theoretical, managerial, and policy implications. It implies that we may need a more refined theory of how mobility frictions interact with firm strategy. Specifically, the availability of frictions enables strategic differentiation contingent on firm characteristics. We contribute to the human capital literature on the use of legal levers (Agarwal *et al.*, 2009; Starr *et al.*, 2018b) by developing a framework highlighting the tension between the ability of firms to attract and to retain talent. By uncovering the firm-level heterogeneity, we also contribute to the literature on NCAs by complementing its focus on state-level enforceability (Conti, 2014; Starr *et al.*, 2018a) or worker-level heterogeneity (Starr *et al.*, 2021, Rothstein & Starr, 2021). From a managerial perspective, our study may guide whether, when, and how managers should use NCAs. Our study is also highly relevant in the context of recent effort by the federal government to ban NCAs due to their detrimental effect on workers.[5] Since NCAs are governed by state laws, and the proposed initiatives face significant pushback from some states, banning NCAs without a federal legislative overhaul may not be possible. However, the public discourse may make skilled workers more aware and proactive when dealing with NCAs. Such movement may, in turn, lead more firms to voluntarily opt out of using NCAs to attract skilled workers. Consequently, the mechanism that we describe in our study could play an important role in decreasing the extent to which NCAs are used by firms even when the national effort to explicitly ban NCAs fails.

2.2 Theoretical Background

A long-standing line of research has focused on examining the management of knowledge embodied in human capital (Coff, 1997; Rosenkopf & Almeida, 2003; Singh & Agrawal, 2011; Carnahan,

firm size.

[5]https://www.ftc.gov/legal-library/browse/federal-register-notices/non-compete-clause-rulemaking

Agarwal, & Campbell, 2012; Kim & Steensma, 2017; Sevcenko & Ethiraj, 2018; Tzabbar & Cirillo, 2020). The fundamental problem is that human capital is free to leave (Coff, 1997; Campbell, Ganco, Franco, & Agarwal, 2012; Raffiee, 2017). While the knowledge (such as trade secrets, intellectual property, or client information) created by employees through employment generally belongs to the firm (Klepper, 2001; Agarwal *et al.*, 2009; Campbell, Coff, & Kryscynski, 2012), knowledge is much harder to protect against expropriation than other resources (Arrow, 1962). Employers often utilize a variety of knowledge safeguards, such as patents or restrictive clauses in employment contracts to inhibit competitors' access to their valuable knowledge (Kim & Marschke, 2005; Starr *et al.*, 2018a). Since mobility represents a key channel through which knowledge spillovers and expropriation occur (Rosenkopf & Almeida, 2003; Agarwal *et al.*, 2009), knowledge safeguards are often focused on improving retention (Ganco, Ziedonis, & Agarwal, 2015; Starr *et al.*, 2018a), deterring employees from joining competitors (Marx *et al.*, 2009), or inhibiting knowledge use after departure (Agarwal *et al.*, 2009).

Scholars studying the management of knowledge flows through mobility have often focused on the role of NCAs (Garmaise, 2011; Starr *et al.*, 2018a, b; Seo & Somaya, 2021). NCAs belong to a group of post-employment restrictive covenants that employees may sign as part of their employment contract. NCAs restrict employees from joining or starting a competing firm (as defined by the NCA) after they leave within a specified geographical area and for a specified duration. While the allowed scope of NCAs varies, some form of the clause is enforceable in almost all US states, with the most notable exception being California (Gilson, 1999) and, very recently, New York.[6] However, NCAs have been shown to discourage mobility even when non-enforceable (Prescott & Starr, 2022).[7] Relative to the enforcement of patents and other forms of intellectual property (IP) by litigation, NCAs represent an effective tool because they protect all knowledge that employees possess from use by rivals (Balasubramanian, Chang, Sakakibara, Sivadasan, & Starr, 2022). Violations also tend to be readily observable, and if enforceable, these violations may be easier to prove in the courts. Prior research has shown that NCAs improve employee retention (Lipsitz & Starr, 2022; Balasubramanian *et al.*, 2022), are positively associated with employees' mobility to different states or different fields (Marx *et al.*, 2009), and lower the likelihood of employees starting competing businesses (Starr *et al.*, 2018a). By increasing retention, NCAs also encourage employers to invest in human capital (Meccheri, 2009; Starr *et al.*, 2018a, b).[8]

Despite the extensive study of NCAs, we lack an understanding of why some firms do *not*

[6]https://ag.ny.gov/sites/default/files/non-competes.pdf accessed on July 23, 2023.

[7]Many firms include the NCA provision in contracts even in states where the NCAs are not enforceable (Starr, 2019). Not all employees may be informed about the enforceability of the NCAs in the state and signing this provision as part of the employment contract may be sufficient to discourage mobility (Prescott & Starr, 2022).

[8]Prior research has also shown mixed effects by NCAs on wages (Garmaise, 2011; Starr, 2019; Starr et al., 2018a). This is most likely driven by the presence of competing effects. While NCAs lower the value of employees' outside options, which allows the employer to appropriate more value in the form of lower wages, employers may be legally required to provide additional compensation for signing the NCA. Further, greater investments in human capital may make the individual better trained, increasing their value on the labor market even with the NCA in place. As a result, the wage effects of NCAs are ambiguous and prior work has shown mixed results (Garmaise, 2011; Starr, 2019; Starr et al., 2018a).

use NCAs, even in states where the practices are enforceable and in industries where they are common. It raises the possibility that the benefits of using NCAs may not be universal, and there may be downsides to including NCAs in employment contracts. Our key proposition is that while NCAs improve the ability to retain existing workers, they may worsen the ability of firms to attract prospective talented workers. Some firms may sacrifice an improved ability to retain workers in exchange for an improved ability to attract them. This logic also implies that making NCAs available to firms creates differentiation opportunities.

We rely on survey data to unpack such firm-level heterogeneity and propose a novel theoretical explanation for why competitors differ in their use of NCAs. The body of existing work was developed by examining the implications of state-level variation in the enforceability of NCAs. Complementing these studies, a smaller stream documents the variations in NCA use at the worker level. These studies rely on worker surveys representative of the US labor force (Starr *et al.*, 2021, Rothstein & Starr, 2021) or specific groups of employees such as executives (Kini, Williams, & Yin, 2020; Shi, 2021), physicians (Lavetti, Simon, & White, 2020), or hair salon employees (Johnson & Lipsitz, 2020). Less is known, however, about firm-level heterogeneity in the use of NCAs and its drivers.[9]

We start by describing a general conceptual framework capturing how using NCAs affects a firm's ability to attract and retain talent and how such a mechanism depends on firm characteristics.[10] We examine the patterns observed in the data, including what predicts NCA use, how firms respond to the question of why firms do not use NCAs, and how the predictors of NCA use correlate with the hiring and retention of workers (when comparing firms that use and do not use NCAs). Finally, we discuss how the patterns relate to the conceptual grounding and how they inform prior work. We do not test specific predictions because our findings inform our theoretical framework. Consistent with several studies in the recent literature (Eesley & Lee, 2020; Agarwal, Ganco, & Raffiee, 2021), our approach to the analysis is abductive (Heckman & Singer, 2017; Goldfarb & King, 2016).

NCAs and the Ability of Firms to Retain and Attract Talent

By limiting mobility to competitors, NCAs can isolate firms' valuable resources by constraining their flow to competitors. However, the NCAs also significantly restrict workers' options for future employment, and some employees may be unwilling or hesitant to sign them (Rubin & Shedd, 1981; Rothstein & Starr, 2021; Balasubramanian *et al.*, 2022). The "career-chilling" effect of NCAs may be more salient to workers who value career flexibility or to whom the restrictions may be particularly

[9] Starr et al. (2021) report a positive correlation between firm size and NCA use. However, their focus is on the individual-level variation of NCA use and they do not explain or explore variation across firms, which is the subject of our study. The authors also do not identify the lower ability to attract talent as a potential downside of using NCAs. Further, we describe below that individual-level differences in NCA use as reported by Starr et al. (2021) could not explain our results. We thank an anonymous reviewer for helping us to delineate differences between our study and prior work.

[10] While we focus on NCAs when developing the theoretical arguments, the logic may apply to other restrictive practices as well. We revisit the potential differences in the discussion.

harmful, such as skilled workers. Using NCAs can thus be detrimental to firms that value skilled human capital. Using NCAs can enhance the retention of current workers while lowering the ability to attract prospective workers. This creates tension for the focal firm.

How does such tension vary across firms, and how do firms decide to use or not to use NCAs with their workers? Notably, the calculus related to the ability to attract and retain workers is separate from the explicit costs associated with NCAs. For instance, the explicit costs include compensating differentials in the form of higher wages that employees receive in return for signing an NCA. As discussed above, the extant empirical evidence shows that the effect on wages is null or ambiguous because signing an NCA also puts downward pressure on wages (Garmaise, 2011; Starr *et al.*, 2018a). Other costs are non-wage costs, such as those associated with drafting and incorporating the covenants into the employment contracts and enforcing the NCAs through litigation (enforcing an NCA is typically cheaper than enforcing a patent [Kesan & Ball, 2006]). Such direct costs likely apply to all rivals in a similar fashion.[11] Given our interest in explaining within-industry cross-firm heterogeneity in the use of NCAs, we focus on how the calculus with respect to the ability of firms to attract and retain workers varies with firms' competitive positions and key resources (i.e., within a state, which determines the NCA enforceability and applies to all firms in the state, and within an industry, which broadly defines the set of competitors).

We posit that the incentive to use NCAs is generally larger for industry leaders than for non-leaders for the following reasons. First, NCAs prevent the expropriation of valuable knowledge by competitors. The benefit of NCAs for the focal firm depends on the nature of skills and knowledge embedded in human capital that the firm needs to protect from expropriation. Knowledge residing in leading firms is likely more valuable to rivals than the knowledge from non-leader firms, as leaders may accumulate knowledge related to new technologies or valuable clients (McElheran, 2015). Rivals may desire to hire employees of leading firms because of the knowledge they possess or because they are likely to be more skilled and productive than those of their rivals (Campbell *et al.*, 2012a).[12] Even in less technologically intensive industries, employees of leaders may have more valuable social capital that may be expropriated by rivals (Phillips, 2005). The leading firms may invest more heavily in training their workers, and these investments cannot be realized if the workers leave. Consequently, using employment practices such as NCAs that improve retention and restrict knowledge outflows may be more critical for leaders.

It is helpful to note that, from the workers' perspective, working for leading firms may be attractive even if the firms ask them to sign NCAs.[13] This is because the leading firms may provide

[11] In our analysis, we also condition on firm size, which should serve as a proxy for resources that firms have available for the enforcement of NCAs. We also include a control for NCA prevalence among competitors, which serves as a proxy for whether NCAs are a common practice in the focal industry.

[12] Note that even leading firms that rely on talent less critically as their source of competitive advantage relative to other resources may attract more productive and skilled workers than competitors. Consequently, this argument is separate from the discussion of talent below.

[13] We can also think about the firm's decision to seek new hires as a demand side and workers' willingness to join as a

higher compensation, more opportunities for learning and careers, and higher status. Although the ability to attract talented workers is important for all firms, leading firms (relative to non-leaders) may be less concerned about the impact of NCAs on their ability to attract talent because leading firms may still be desirable places for workers despite NCAs.

From a non-leading firm's perspective, however, the impact of NCAs on the ability to attract external talent may be more critical than their impact on the retention of existing employees. Non-leaders generally have lower bargaining power in the labor market and may be unable to offer comparable compensation, status, and advanced learning opportunities relative to leading firms. Their ability to attract talented workers is likely lower than that of leading firms. Moreover, while employee exit is still detrimental to the human capital of non-leading firms, these firms likely worry less about the leakage of their knowledge to competitors. Their knowledge is more likely to represent the industry average rather than cutting-edge. Hence, non-leaders (relative to leaders) may face more difficulties when attracting skilled workers and be less concerned about knowledge leakage to rivals. Such calculus may favor opting out of NCA use for non-leaders relative to leaders.

From the standpoint of workers who value career flexibility, firms *not* requiring NCAs may become attractive. Thus, opting out of NCA use can be an effective differentiation strategy for non-leading firms. These differences between leader and non-leader firms are also in line with previous literature arguing that non-leaders may need to implement strategies that mitigate or neutralize the dominant firms' competitive advantages while avoiding head-on retaliation (Mascarenhas, 1986; Ito, 1997), and leaders are more likely to use strategies to prevent knowledge spillovers to competitors than non-leader firms (Albino, Garavelli, & Schiuma, 1998; Paniccia, 1998; Boschma & Lambooy, 2002).

The different emphasis on attracting versus retaining workers between leaders and non-leaders may be amplified by the extent to which firms rely on talented workers as a source of differentiation. Talented workers may likely be more critical in some industries than others, such as less capital-intensive and more service-based contexts. But their importance as a source of differentiation may also vary within industries. While leaders may have, on average, more productive workers, and more valuable knowledge that rivals can expropriate, they may also hold other valuable resources such as brand and various capital-intensive assets. Thus, the relative importance of talent as a source of differentiation may not always be higher for leading firms. Consequently, it is meaningful to consider reliance on talent as a separate construct from a firm's leadership position. Extensive prior work in strategy has shown that firms, even within industries, vary in how they utilize resources to create competitive advantage (Porter, 1985; Barney, Wright & Ketchen Jr., 2001). Importantly, how much firms rely on talent likely affects how NCAs affect their ability to attract versus retain talent.

When leading firms rely more heavily on talent relative to other resources, competitors' potential

supply side. For instance, our arguments imply that the supply is less sensitive to NCA use by leading firms (relative to non-leading firms) and that reliance on talent increases the demand.

expropriation of their human capital becomes more threatening to their competitive advantage. Such threats increase their incentive to use NCAs. Talented workers may be willing to join these leading firms *despite* having to sign NCAs. Consequently, if human capital is more critical to their businesses relative to other resources, NCAs are more useful for leaders because potential knowledge leakage becomes more threatening and, at the same time, they may be less concerned with the negative impact of NCAs on their ability to attract new workers.

The situation is likely reversed for non-leading firms. Non-leaders relying on talent need to attract and retain high-quality workers, while the human capital may not be inherently more valuable to other firms. Thus, such firms may not need NCAs to prevent the leakage to rivals, while they may still worry about the impact of NCAs on their ability to attract skilled workers, which is amplified as talent becomes more critical for the firm. Consequently, non-leading firms relying heavily on talent may opt out of NCAs. Further, if NCAs are common among the leaders, non-leaders may advertise their opt-out in the labor market. For instance, Daniel Hertzberg, the CEO of CAD-design firm Onshape,[14] argued in an article for Boston Globe that, although using NCAs is a standard practice in Massachusetts, the company is following other tech firms, such as Acquia and RunKeeper (all are non-leaders) that voluntarily eliminated NCAs for their workers. As we mention in the opening quote of our study, he claimed that “we want people who want to be here, not ones who feel trapped." Capturing the importance of being different, he also described that “we’ve been contacted by many talented people we’d like to hire but who are restricted by noncompete agreements. After they endure their waiting period, it would feel hypocritical to then ask them to sign the same kind of draconian agreement at my company."[15] This example illustrates that non-leading companies that need to attract talented workers may differentiate by opting out.

2.3 Empirical Analysis

Data and sample

Our data are derived from an employer survey conducted in 2017 by PayScale, a US-based data analytics company focused on gathering information about human resources (HR), mainly compensation practices. PayScale deployed the survey to HR managers and top executives during November and December 2016. Given that the data gained through the surveys constitute a core product for PayScale, they focus on achieving data quality, accuracy, and reliable sourcing of the data. The survey asked questions on hiring practices, compensation practices, and the use of post-employment

[14] According to an industry report by Enlyft tracking the use of various CAD products and technologies, Onshape has a non-leading market position with a market share of less than 5% in the Computer-aided Design & Engineering category and rank 71 out of 109 in the firms offering CAD products. https://enlyft.com/tech/computer-aided-design-engineering accessed on 10/11/2022.

[15] https://www.bostonglobe.com/opinion/2015/06/27/onshape-ceo-john-mceleney-noncompetes-hurt-workers-and-their-employers/6NbXbI5jhZpl5wyvc28FSI/story.html accessed on 10/12/2022.

restrictive covenants such as NCAs.[16]

The data contain responses from 7,698 global employers participating in the survey. The respondent organizations are diverse, including private and public firms, public institutions, universities and schools, hospitals, and non-profit organizations. We limit our sample to private and public firms (3,610 observations dropped) headquartered and located in the US (given large differences in the relevant legal frameworks outside the US; 1,278 observations removed). From the remaining 2,810 firms, including both Fortune 500 companies and small and medium-sized businesses, we further remove respondents who were either unsure of their NCA use or chose not to answer the NCA questions (1,007 observations removed).[17] The NCA questions are more likely to be missing for larger firms (which we condition out by controlling for firm size). To further mitigate the potential bias stemming from missing observations, we employ a robustness test that imputes missing NCA-use variables.

The restrictions imposed above result in a sample of 1,803 observations.[18] In our regressions, the sample size varies across models with different dependent variables due to the varying extent of missing variables. The sample somewhat overrepresents larger firms compared to the US firm population (based on the 2017 County Business Patterns), but significantly less so than Compustat.[19] In addition to raw estimates, we use iterative proportional fitting to create weights to match our sample to the population along with size, industry, and state. Then, we re-estimate our main specifications using weighted least squares.

Variables

NCA use: We create categorical and continuous variables capturing NCA use based on two survey questions. The first one is: "Which employees at your organization are subject to non-compete agreements (Prohibited from joining or starting a competing organization)?" The respondents chose from four categories: 'All employees,' 'Some employees,' 'No employees,' and 'Don't know.' For those who chose 'Some employees,' the survey provides a follow-up question: "To the best of your knowledge, what percentage of all employees within the organization have signed non-competes?" The respondents can select the answer from five choices: '1-20%,' '21-40%,' '41-60%,' '61-80%,' and

[16]The full survey contains 125 questions. The survey focuses on compensation practices (data on compensation practices is the main product of PayScale) so most questions focus on how firms compensate their workers (why firms give pay raises, pay structure, incentives, administration of pay, how job classifications are created, pay grades and ranges, variable vs. fixed pay, compensation strategy, how firms measure performance, transparency of pay, performance reviews, payroll system, etc.).

[17]Appendix Table A1 compares firms that answered the NCA questions with those that did not answer them. In Table A1 Model 3, we include firm-level controls and industry and state fixed effects, and we do not find noticeable differences in our key independent variables ("non-leader" and "talent").

[18]In 99% of observations, a single person responded for each firm. In 1% of observations, more than one respondent submitted the survey because different establishments of the same firm were registered separately as clients with PayScale. The results are robust to the exclusion of either one or both observations for each of these firms. As shown in Appendix Table A2, 83.3% of respondents are in managerial or higher occupations, and over half (54.8%) have an HR-related job. The respondents likely have an accurate knowledge of their firms' NCA use.

[19]Appendix Table A3 compares the firm size distributions for PayScale, Compustat, and the County Business Patterns.

'81-100%.' The categorical NCA-use variable is a dummy coded as one if the response to the first survey question is 'All employees', zero if 'Some' or 'No employees,' and missing if 'Don't know'.[20] The continuous NCA-use variable is measured by the fraction of employees subject to NCAs using the follow-up survey question. We take the midpoint of each category (e.g., 10 for '1-20%') for those who chose 'Some employees' and treat 'All employees' as 100% and 'No employees' as 0%.

Industry leader/non-leader. This variable is constructed based on the response to the question: "Is your organization #1 in its industry?" The response can be 'Yes,' 'No,' or 'I don't know.' We construct a binary variable, *Non-leader*, coded as one if their choice is 'No' and zero if their choice is 'Yes.' The choice of 'I don't know' is treated as missing (26% of the sample). While this measure is subjective, most of the prior literature relies on secondary measures inferred from performance data.[21] For the subset of public firms, we match our sample with Compustat to examine the validity of the survey-based leadership measure. The survey-based measure appears to be a reasonable proxy for the leadership position based on the Compustat sample.[22]

Talent as the key differentiator from competitors. The variable is constructed based on the question: "Which of the following sets your company apart from competitors the most?" This single-selection question has six possible choices: 'Larger client or customer lists,' 'Talented employees,' 'Innovative products,' 'Best-in-class service,' 'Better at improving employee skills,' 'Other (please specify).' For the main analysis, we construct a binary variable, *Talent*, coded as one if the response is 'Talented employees' and zero if other choices. This aggregation is informed by the patterns observed in the data. We validate the measure using the Current Population Survey (CPS) in the robustness test section. It shows that the measure exhibits a strong correlation with skilled human capital intensity. In a robustness test, we also use a survey question capturing the amount of training new hires receive as a proxy for the importance of talent. While the results remain consistent, we suggest interpreting them with caution because the amount of training may depend on the use of NCAs (Starr *et al*., 2018b).

NCA nonuse reason. To the respondents who answered that their organization either does not use NCAs at all or uses them only for some workers, the survey follows with the question about the reasons for not using NCAs (at least for some workers). The question is: "Why doesn't your organization use non-competes (multiple selections allowed)?" There are six possible answers to this question.[23] With our focus on attracting and retaining talent, we create a binary variable coded as one if the respondents chose 'Non-competes make it hard to attract talented employees,' and

[20] In the robustness section, we assess the robustness of our results using several alternative operationalizations of the NCA use variables.

[21] For instance, measures based on the threshold values of R&D intensity (Berry, 2006), size, sales (Ito, 1997; Ito & Pucik, 1993; McElheran, 2015), or market share (Berry, 2006) have been used to proxy for industry leadership.

[22] Due to the smaller number of public firms, we are unable to examine the main analysis using this subsample and the Compustat variables. We discuss the validation of the measure in more detail in the robustness section.

[23] The possible choices consist of 'Non-competes are not commonly used in the industry,' 'Non-competes make it hard to attract talented employees,' 'Loss of employees to competitors is not a big concern,' 'Not familiar with what non-competes are,' 'Non-competes are not legally enforceable in my state,' and 'Other (please specify).'

zero otherwise. We also construct another binary variable coded as one if 'Loss of employees to competitors is not a big concern' was chosen, and zero otherwise.

Ability to Attract Talent, Ability to Retain Talent. To explore how the use of NCAs relates to the ability of firms to attract and retain talent, we rely on multiple survey questions. As a proxy for the ability to attract talent, we first use the question: "Do you have any positions that have been open for six months or more?" Companies that are better able to attract talent should have fewer vacant positions. We create a binary variable based on the response of 'Yes' or 'No'. Among those responding 'Yes,' the survey follows with: "What kind of positions do you have a hard time filling? (check all that apply)." There are nine possible choices that are not mutually exclusive. As a proxy for higher knowledge intensity, we focus on managerial jobs ('Management' and 'Executive Level') and technical jobs ('IT' and 'Engineering'), and create dummy variables based on each category.[24]

Control variables. Our controls include variables identified in prior work as potentially driving NCA use. We utilize industry and state fixed effects for all analyses. Prior literature shows that NCAs tend to be more common in states that enforce them and for workers in technical sectors (Starr *et al.*, 2021; Balasubramanian *et al.*, 2022). In the robustness section, we employ subsample analyses based on states where NCAs are enforceable and in high-tech industries only. The PayScale uses 28 industry categories instead of standardized industry classes such as SIC or NAICS.[25] We include several additional control variables derived from survey questions: *firm size, NCAs common among competitors in local markets*, *primary deliverable*, the *share of low-wage employees*, and *the respondent job function*. *Firm size* is a categorical variable based on the question: "How many full-time employees are in your organization?" The respondents choose from: "1-99 employees," "100-749 employees," "750-4,999 employees," and "Over 5,000 employees." Starr *et al.* (2021) report that individual workers in larger firms (over 5,000 employees) are more likely to be bound by NCAs relative to smaller firms. Our results are robust to the inclusion or exclusion of firm size. *NCAs common among competitors in local markets*, which captures NCA use within the local labor market, is based on the response to the question: "To the best of your knowledge, how common are non-competes in your local market among competitors?" The respondents choose 'Very uncommon,' 'Uncommon,' 'Common,' or 'Very common.' We construct a binary variable coded as one if the response is 'Very common' or 'Common' and zero if 'Very uncommon' or 'Uncommon.' *Primary deliverable* is based on the response to the question: "What is your organization's primary deliverable?" The possible choices consist of 'a tangible product or products,' 'a service or services,' 'knowledge and information,' or 'something else (please specify).' The type of deliverable may correlate with differences in the human capital of the firm's employees. For *the share of low-wage employees* (who are less likely to be subject to the NCAs [Starr *et al.*, 2021]), we utilize the question: "What percentage of full-time

[24]The other possible choices include 'Customer Service,' 'Sales,' 'Marketing,' 'Finance,' and 'Other (please specify).'

[25]Twenty one percent of respondents did not select a specific industry from the list and entered 'other.' To reduce noise, we tried to assign firms to industries based on searching the company profiles. This reduced the proportion of firms in the 'other' category to 8.9%. The results are robust to the exclusion of this category.

employees in your organization earn less than $47,000 per year?" There are six choices: 'None,' '1-20%,' '21-40%,' '41-60%,' '61-80%,' and '81-100%.' We create a continuous measure by taking the midpoint of each category and treating the 'None' answer as zero. Finally, we control for the job function of the survey respondent, as the respondent's ability to answer specific details may vary with their position. The variable of the *respondent job function* is categorical based on the question: 'What is your primary job function?' The functions are aggregated into three categories: 'HR and compensation,' 'Executive (COO, CEO, etc.),' and 'Others.'[26]

Estimation Method

We employ a series of OLS regressions and linear probability models for ease of interpretation. Our claims are not causal, and the key objective of the analysis is to isolate the estimated relationships from alternative explanations unrelated to our conceptual framework. Throughout our specifications, we rely on industry and state fixed effects as well as the set of control variables to uncover the firm-level heterogeneity within industries and states. To corroborate our findings and validate our survey-based measure, we employ a series of robustness tests using different specifications and subsamples, and external data sources such as Compustat, Glassdoor, and CPS. To further assess our results' sensitivity to alternative explanations, we employ recently developed sensitivity tests to assess how strongly omitted variables would need to be associated with our explanatory and outcome variables to overturn the results (Cinelli & Hazlett, 2020). In all our specifications, standard errors are clustered at the industry-by-state level. All results remain robust if we use two-way clustering on industry and state instead.[27]

Results

Table 1 reports the summary statistics and pairwise correlations for the variables we use in the analyses. The categorical variables are dichotomized for ease of interpretation. Table 2 describes the rate of NCA use for the key variables. In our sample, 71.7% of firms use NCAs for at least some of their employees, and 31.4% use them for all employees. Consistent with Starr *et al.* (2021), we observe that larger firms with 100 or more employees tend to adopt NCAs (77.4%) more often than smaller firms with 1-99 employees (66.7%), while the larger firms are likely using them only for some employees (53.0%). In contrast, smaller firms are more likely to use NCAs for all their employees (37.6%) rather than only some (29.1%). Unconditional adoption rates of NCAs for industry leaders are 73.7%, and for non-leaders, 70.9%. Firms reporting talented employees as what sets them apart from competitors are less likely to adopt NCAs (65.3%), compared to 73.1% of firms that report that factors other than talent set them apart from competitors. The prevalence of NCAs in local markets where firms operate is largely associated with the NCA use of focal firms.

[26]The category 'Others' includes 'Finance/Accounting', 'Sales', 'Technology', 'Marketing', 'Operations', 'Consultant', 'Other (please specify)'.

[27]The results using two-way clustering are available upon request.

90.3% of firms that report NCAs are common among their competitors in their local markets use NCAs for their workers, whereas only 44.7% of firms that report NCAs are uncommon among their competitors in their local markets use NCAs.

Table 3 summarizes the use of NCAs within and across industries. The highest prevalence of NCAs is in Marketing & PR, Technology, and Business & Management (40.3-63.2% of firms require NCAs from all workers), and the lowest prevalence is in Customer Service and Real Estate (48.5 and 59.3% of firms require no NCAs at all). Thus, industries that rely more on talented workers and clients (Marketing/PR, Business & Management) and technological knowledge (Technology, Biotech & Science) have a higher likelihood of firms using NCAs.[28] The table also reveals a significant heterogeneity within industries, including knowledge-intensive industries (the correlation between the reliance on talent and NCA use is negative within industries). For instance, in Accounting/Finance, 33.6% of firms require NCAs from all workers, while 53.8% do not require NCAs at all, or less than one-fifth of their workforce is required to sign them. A similar pattern is present in Medical/Healthcare: 22.4% of firms require all employees to sign, while 58.9% require no NCA or only less than one-fifth of their workforce.

The last column of Table 3 reports the reason for not using NCAs in the focal industry (or using them only for some workers): "NCAs make it hard to attract talent." The highest proportion of firms that list this reason for not using NCAs is in Architecture/Engineering and Marketing/PR (34.7% and 33.3% among firms not using NCAs at all or using them only for some workers), followed by Accounting/Finance (19.4%). This is consistent with our argument that some firms relying heavily on skilled workers choose *not* to use NCAs as a differentiation strategy, even in industries where many firms use NCAs heavily (as mentioned above, 63.2% of firms in Marketing/PR, 33.6% in Accounting/Finance, and 26.3% in Architecture/Engineering use NCAs for all workers).

We proceed by exploring what explains NCA use within each industry and state. Table 4 reports regression results where the dependent variables are NCA use (a dummy for all workers and the fraction of workers having NCAs), and the key independent variable of interest is the interaction between the industry non-leader status and the reliance on talent as a differentiator from competitors (i.e., "talent").[29] The table reporting the coefficients for all controls is in Appendix Table A5. The estimation sample is reduced from 1,803 to 1,232 (68.3%) due to missing survey responses on the

[28]The knowledge intensive industries have, on average, somewhat higher rates of NCA use and the reliance on talent positively correlates with industry level of knowledge intensity (described in the robustness section). Note that this does not contradict the negative unconditional correlation between NCA use and the reliance on talent (Tables 1 and 2). This is because the unconditional correlation between NCA use and the reliance on talent depends on their correlation both within and across industries. The results on Table 4 show that, within industries, the correlation between NCA uses and the reliance on talent is negative and large for non-leaders.

[29]In Table A4, we report the number of observations in each cell defined by industry leader (vs. non-leader), talent (vs. non-talent), and the fraction of workers subject to NCAs.

independent variables.[30] We call the remaining 1,232 observations *the NCA estimation sample*.[31]

Based on Model 1 of Table 4, which only includes the main effects of talent and industry non-leaders, non-leaders are 3.6 percentage-points less likely to use NCAs for all workers (p-value=0.190), and firms that report talented employees as the key differentiator from the competition are 5.7 percentage points less likely to use NCAs for all workers (p-value=0.081). As we will see in the following models, the negative coefficient on talent as a predictor for NCA use is driven by non-leading firms (i.e., it becomes positive for leading firms). Model 2 adds the key variable of interest, the interaction term of talent and industry non-leaders. Non-leader firms that report talent as the key differentiator from competition are 9.3 percentage points *less* likely to use NCAs than non-leaders relying on other resources, and the estimate on the interaction term is -0.151 (p-value=0.060). In contrast, leader firms that report talent as the key differentiator from competition are 5.8 percentage points *more* likely to use NCAs than leaders that report other resources as the key differentiator, though the estimate is noisy (p-value=0.404). Model 3 looks at the variation within industries and states. The coefficient on the interaction in Model 3 is not only consistent but larger in magnitude. Non-leaders relying heavily on talented employees are 12.0 percentage points less likely to use NCAs than non-leaders relying on other resources, and the coefficient on the interaction term is -0.190 (p-value=0.017), whereas leader firms relying heavily on talented employees are 7.0 percentage points more likely to use NCAs relative to leader firms relying on other resources (p-value=0.132). In Model 4, with additional controls, the corresponding coefficient on the interaction term is -0.189 (p-value=0.015).[32] To illustrate the size of these estimates, Panel A of Figure 1 plots the predicted likelihood of having all workers signing NCAs across leaders vs. non-leaders and talent vs. non-talent (based on Model 4 in Table 4). There is no significant difference across leaders vs. non-leaders in the likelihood of using NCAs for all workers if they rely on something other than talent as the key differentiator. In contrast, if talent is the key differentiator from competition, then non-leader firms are 20.5 percentage points less likely than leader firms to use NCAs for all workers.[33]

Models 5-8 in Table 4 report the results using the fraction of employees subject to NCAs instead of a dummy for all workers with NCAs. Around 5% of the remaining respondents (66 observations) did not provide an answer for the NCA-use fractions, so the number of observations used here is 1,166. Based on the most saturated model in Model 8, non-leaders relying on talent as the

[30]The largest reduction occurs due to the industry non-leader dummy. Removing observations with missing values or the values recorded as "I don't know" lowers the sample size by 483 observations (26.8% of the sample). The second largest reduction stems from the variable NCAs common among competitors in local markets" (61 observations, or 3.3% of the sample). This is likely because some respondents lacked the knowledge to answer this question.

[31]In the analysis, for each dependent variable, we use the sample from the fully specified model (i.e., all variables are non-missing).

[32]The inclusion of non-leader, talent, and their interaction increases the R2 of the model from 0.187 to 0.197.

[33]As we show in the robustness section, the results remain unchanged if we exclude the 'Some' category or if we aggregate with the 'All' category the firms that use NCAs for at least 50% of their workers. When we aggregate 'Some' with 'All' in the dependent variable or aggregate firms below the 50% threshold with 'All', the coefficients on the interaction term continue to be negative but the errors become larger. This suggests that non-leader firms that opt-out of NCAs use do so to differentiate themselves from leaders that use NCAs extensively.

key differentiator are using NCAs for 9.8 percentage points *lower* proportion of workers relative to non-leaders relying on other resources. In contrast, leaders relying heavily on talent as the key differentiator are using NCAs for 8.5 percentage points *higher* proportion of workers relative to leaders relying on other resources (p-value=0.198). The coefficient on the interaction term is estimated at -0.183 (p-value=0.016).[34] Like above, Panel B of Figure 1 shows that the difference in NCA use in terms of the fraction of workers subject to NCAs is only salient when firms rely on talented employees as the key differentiator.[35]

Among firms that do not use NCAs at all or use them only for some employees, we further explore their reasons for not using NCAs (at least for some workers). Specifically, we examine how firms' industry positions and the reliance on talent relate to their concerns about attracting talented employees as the reason for the NCA non-use. Table 5 reports the regression results (coefficients on all controls are reported in Appendix Table A7). Here, we use a subsample of 691 respondents (56.1% of the NCA estimation sample) who do not use NCAs at all or use them only for some workers and responded to the questions about the reasons for the NCA non-use. The dependent variable is a dummy for the response, "NCAs would make attracting talented workers difficult." In Model 1, only including non-leaders and talent, non-leaders are 4.0 percentage points more likely to report the ability to attract talent as the reason for not using NCAs than leaders, and firms relying heavily on talent are 3.0 percentage points less likely to do so than firms relying on other resources. In Model 2, adding the interaction term, industry leaders relying heavily on talented employees are 15.1 percentage points less likely to see the ability to attract talent as the reason for not using NCAs relative to leaders relying on other resources (p-value<0.001). Thus, when leaders relying on talented employees do not use NCAs, their motivations likely stem from other reasons: NCAs are legally unenforceable, NCAs are not common among its industry members, or firms are not familiar with NCAs. For non-leaders, however, such associations reverse; among firms reporting talented workers as the key differentiator, industry non-leaders (relative to leaders) are 16.8 percentage points more likely to report that they do not use NCAs because the NCAs would make hiring talent difficult (p-value for the interaction term=0.003; Model 2). The results remain consistent across specifications including industry and state fixed effects (Model 3) as well as other firm-level controls (Model 4).[36] Panel C in Figure 1 displays the predicted likelihoods based on Model 4.[37] When we observe non-leaders that rely on talent not using NCAs, it is more likely because they worry about

[34]The inclusion of non-leader, talent, and their interaction increases the R2 of the model from 0.260 to 0.268.

[35]Appendix Table A6 shows the disaggregated results of NCA-use for the full set of differentiating factors.

[36]In Model 4, the inclusion of non-leader, talent, and their interaction increases the R2 from 0.176 to 0.182.

[37]In Appendix Table A8, we employ analogous regressions where the reason for not using NCAs is a lack of concern for losing talent. Our theory posits that fewer concerns about the expropriation of knowledge by competitors can be a reason for opting out of using NCAs. The dependent variable is a dummy for the "loss of employees is not a serious concern." The results are consistent with our theory. For instance, in Model 2, the coefficient on the interaction between non-leaders and talent is a 26.6 percentage points increase (p-value<0.001). However, the interpretation is more difficult because lower concern about losing talent may be due to other reasons: the firm may have strong bargaining power in the labor markets, the labor supply may be abundant, or the risk of expropriation may be small. Appendix Table A9 shows the disaggregated results of NCA-nonuse reasons for the full set of differentiating factors.

their ability to attract talented workers relative to leaders that rely on talent. These results provide evidence that some firms strategically opt out of using NCAs because they are concerned about their ability to attract talented workers.

Next, we explore whether and how the non-use of NCAs is associated with a firm's actual ability to attract talent. In Table 6, we use the survey questions about whether firms face difficulty filling competitive job positions that are vacant for a longer period. Here, the sample consists of those who completed job position questions (832 respondents, or 67.5% of the NCA estimation sample). Our primary interest is in how the ease of filling vacancy positions is different across firms using and not using NCAs, depending on their industry positions and their reliance on various resources as a source of differentiation. For ease of interpretation, we first divide the sample based on the reliance on talented workers or other differentiators and then separately estimate the coefficients on the interaction for non-leaders and the fraction of workers *not* subject to NCAs (i.e., the extent of firms *not* using NCAs). This estimation is strictly correlational.[38]

In Table 6, Models 1-3 are based on the subsample of firms relying on talented workers as the critical differentiator, while Models 4-6 are based on the subsample of firms relying on other differentiators. All the models include the full set of controls and industry and state fixed effects. In Model 1, the coefficient on the interaction term for non-leaders and NCA non-use fraction is estimated as a 19.0 percentage points decrease in the likelihood of firms having long-term vacancy positions in the talent-reliant sample (p-value=0.418), while such a coefficient is 3.1 percentage points increase for the non-talent-reliant sample in Model 4 (p-value=0.751). To get a sense of the estimate, Figure 2 Panel A plots the estimated effects of NCA non-use across the groups of leaders/non-leaders and talent/non-talent based on Models 1 and 4. Comparing those using NCAs for all workers and not using NCAs at all among non-leaders that rely heavily on talent, those not using NCAs at all are 9.7 percentage points *less* likely to have a long-term job opening, while the corresponding estimate for leaders is 11.6 percentage points *increase* in likelihood. These estimates are noisy, but the signs and magnitudes of the coefficients are consistent with our main arguments.

We further disaggregate the types of jobs that firms have difficulty filling. Models 2 and 5 show the results for the likelihood of having a long-term job opening for a managerial position, whereas Models 3 and 6 show the results for the likelihood of having a long-term job opening for a technical position. For both types of jobs, when talented workers are critical resources, industry non-leaders are less likely to encounter difficulty filling those vacant positions by opting out of NCAs. In particular, the coefficient on the interaction term in Model 3 represents 36.2 percentage points decrease in the likelihood of long-term vacancy in a technical job (p-value=0.089), compared to the corresponding estimate of -0.061 in Model 6 (the non-talent sample; p-value=0.435). Panel B (and C) in Figure 2 illustrate the differences in the likelihoods between firm groups, indicating that

[38]Further, because NCA use is endogenous, as we showed that it depends on the industry non-leader and talent variables, the estimated coefficient of the interaction term does not represent simple derivatives of the two individual terms.

non-leaders relying heavily on talent are 9.9 (13.1) percentage points less likely to have a managerial (technical) long-term job opening if they do not use NCAs at all (compared to those using NCAs for all workers), whereas those relationships reverse for industry leaders. In sum, though correlational and suggestive, the regression results from Table 6 and the plots in Figure 2 do not contradict the key arguments. Opting out of NCA use appears to be associated with easier filling of high-skill positions (conditional on relying on talented workers). NCA use is, thus, more penalizing when attracting workers for non-leaders relative to leaders.

We also implement a supplementary analysis of how firms' decision *not* to use NCAs to attract talent is associated with differences in organizational climate (Schneider, Ehrhart, & Macey, 2011). We use the survey question: "Rate your level of agreement with the following statements," which has six different items. Each answer is based on a five-point Likert scale ('Strongly disagree,' 'Disagree,' 'Neither agree nor disagree,' 'Agree,' 'Strongly agree'). We focus on two items (most others focus on compensation), 'There is frequent, two-way communication between managers and employees' and 'Employees at my organization feel appreciated at work.'[39] Since the reasons for the NCA nonuse are only available for respondents who answered that their organization either does not use NCAs at all or uses them only for some workers, we restrict the sample to these firms. Table A10 shows the regression results where the dependent variables are respondents' evaluations of frequent communication with top management and employees' general job satisfaction. Panels A and B in Figure A1 display the corresponding predicted likelihoods. Overall, these results suggest that when firms do not use NCAs at all, and the reason for not using NCAs is their ability to attract talent, respondents rate communication with managers and job satisfaction higher relative to other firms.[40] These results may imply that firms that opt out of using NCAs because they are concerned that NCAs would inhibit their ability to attract talent may resort to alternative ways of retention such as improving the work environment.

Finally, we examine how the estimated effects vary with the IP environment. If our patterns are partly driven by concerns about knowledge leakage, as we argue theoretically, we should observe the associations to be more pronounced when the ability to protect IP through alternative means such as the availability of patent protection is weaker. This is because the risk of knowledge leakage is likely larger when firms lack other potential means of protecting knowledge, such as patents. To examine this, we explore the heterogeneity in NCA use and the reason for the NCA

[39]The other four items consist of: 'Compensation drives employee engagement at my organization,' 'Employees at my organization feel they are paid fairly,' 'My organization has a transparent pay process,' and 'Employees at my organization have a great relationship with their direct managers."

[40]We also explore various type of incentives such as individual/team incentive bonuses, retention bonuses, hiring bonuses, and other available metrics such as the extent to which compensation drives employees' engagement, and the extent to which employees feel they are paid fairly. The regression results for incentives are reported in Appendix Table A11, and those for other items related to work environments in the same survey question as Table A10 are reported in the Appendix Table A12. The only result representing sizeable associations is team incentive bonuses (Models 3 and 4 in Table A11, which weakly suggests that team-based incentive schemes may effectively reduce individual-level turnover, and thus be used for retention when NCAs are not used. However, as mentioned earlier, the effects of NCAs on compensation are notoriously difficult to estimate and interpret so this result needs to be taken with caution.

non-use as they vary with the patenting intensity of the industry. We rely on the DISCERN database (Arora, Belenzon, & Sheer, 2021) that links the USPTO patents between 1980-2015 to Compustat, to calculate industry-level patent intensity (based on two-digit NAICS). Then, we interact the patent intensity variable with the non-leader and talent variables to estimate how the effects (reported in Tables 4 and 5) vary with patent intensity. In Table A13, we use a dummy for high vs. low patent-intensive industries by equally dividing the industries into both categories. In Table A14, we use a continuous measure of patent intensity, the log-transformed patent stock. We lose some statistical power in the estimations, likely due to noisy matching between the 2-digit NAICS codes and the PayScale industry classifications. Still, the results are consistent with the argument that the relationships between our key independent variable (the interaction between non-leader and talent) and NCA use are attenuated in high patent-intensive industries relative to low patent-intensive industries. For instance, in Table A13 Model 2, in low patent-intensive industries and where talent is a critical resource, non-leaders are 32.7 percentage points less likely to use NCAs for all workers relative to leaders. In high patent-intensive industries, this difference decreases to 15.3 percentage points. Further, in low patent-intensive industries, a lack of concern about retaining talent is more likely to be reported by non-leaders as a major reason for not using NCAs relative to leaders (34.5 percentage points difference, Model 8). This difference decreases to 6.5 percentage points in high patent-intensive industries. Similar attenuations are observed when we use a continuous measure of industry-level patent intensity (Appendix Table A14). The concerns about knowledge leakage are likely critical for leaders when they lack alternatives to protect IP and help to drive differences in NCA use for leaders vs. non-leaders.

Robustness Tests: Alternative Specifications

We perform additional analyses to assess the robustness of the results. We employ weighted regressions on firm size, industries, and states, use alternative operationalization of NCA and talent variables, and assess the robustness of the regression estimates in different subsamples. We also assess the sensitivity of our regression estimates to potential confounders.

First, we assess how the representativeness of the survey sample affects our main findings. We use iterative proportional fitting ("raking") to create the weights based on firm size, industry, and state (Kalton & Flores-Cervantes, 2003). The weighting answers the question of how the estimates may change if the survey sample is perfectly representative of the population of US firms in terms of size, industry, and state.[41] The analysis is in the Appendix and the results are reported in Appendix Table A15. The findings reported in Tables 4 and 5, as well as the findings from Table 6 related to technical occupations remain robust and consistent in the weighted regressions.

Next, we examine the robustness of our main findings to using alternative measures of NCA use (using different aggregation for the binary variable and different cutoffs for the continuous variable),

[41]We obtain the data from the 2017 County Business Patterns for the population of US firms.

alternative proxy for talent (investments in training of new hires), subsample analysis (knowledge-intensive sectors and excluding states with unenforceable NCAs), and analysis using multiple imputations of missing variables. The analysis and the results are described in the Appendix (Appendix Tables A16-A20). The results remain consistent.

Finally, we conduct the sensitivity analysis proposed by Cinelli and Hazlett (2020). The gist of the approach is to estimate how strongly any potential confounders need to be associated with both treatment and outcome to negate the observed relationships. The analysis is described in the Appendix and reported in Appendix Table A21. Based on this analysis, we conclude that potential confounders, such as unobserved individual characteristics, would need to explain far more variation in the dependent variable than observed firm-level characteristics such as firm size to overturn our NCA results. A similar conclusion can be derived for NCA use measured by the fraction of workers and the NCA non-use reason. Consequently, we conclude that omitted variables are unlikely to drive our findings.

Validation of Key Explanatory Variables and Further Mechanism Checks

Since our key variables rely on subjective evaluations by responding managers, we examine the validity of the variables using external sources: Glassdoor, CPS, and Compustat.

First, we check our firm size measure in PayScale using Glassdoor.[42] We scraped the information on employment from the Glassdoor website and merged it with the PayScale dataset by matching a company name, headquarter state, and industry. 92.7% of companies in PayScale with employer names are matched with Glassdoor. Appendix Table A22 examines the correlation between the size measures in PayScale and Glassdoor. It suggests that the PayScale size measure broadly aligns with the Glassdoor size measure, as one additional employee in PayScale is associated with a 1.25 increase in size in Glassdoor.

Next, we examine how our talent measure is associated with the average levels of workers' human capital for each industry and state. The test is inherently noisy, but a positive correlation provides an additional validation check. We link our data with the Current Population Survey (CPS) in the periods 2007-2016 to obtain workers' education levels at the industry level (2-digit NAICS) and the industry-by-state level. The CPS sample was pooled across periods, and the education variables were aggregated to produce three measures: a dummy for a bachelor's degree or higher, a dummy for a master's degree or higher, and a continuous measure of years of education. Regression results are reported in Appendix Table A23. In Panel A, the three measures of industry-level human capital—the ratio of workers with bachelor's degrees or higher, the ratio of workers with master's degrees or higher, and the years of education—are all strongly positively correlated with our talent

[42]Glassdoor is an employer review and recruiting website where both current and former employees voluntarily and anonymously review their companies. The employees are incentivized to leave reviews through a "give-to-get" policy, whereby contributors gain access to the information submitted by others.

measure. Similar results are derived when the education level is measured at the industry-state level (in Panel B).

To examine whether the subjective evaluation of industry leaders and non-leaders in the survey aligns with the objective measures of industry leadership, we match the firm names in our sample with Compustat.[43] From 1,803 observations in the survey, we identify 119 Compustat matches (112 unique firms). We then perform regressions where the independent variable is the industry leader dummy, and the dependent variables are the dummies for the top 10%, top 5%, and top 5 firms in the number of employees, total assets, sales, and net income within the industries defined by 4-digit SICs. The results are reported in Appendix Table A24. Due to its subjective nature, the loosely defined industries and markets, and the small size of the matched sample, our industry leadership measure is not perfectly aligned with the objective financial performance. Still, we find that, for instance, industry leaders are 9.7-11.7 percentage points more likely to be top five firms of any of the metrics within the same 4-digit SIC industries (Panel C).

Finally, it is useful to note that our arguments hinge on the assumption that prospective workers are aware and sensitive to the firms' use of NCAs. We assume that at least some talented prospective workers pay attention to whether a firm uses NCA and, if it does, prefer to select another employment opportunity. While we cannot validate this assumption in the context of our survey, there is emerging evidence in the literature supporting this notion. For instance, Prescott and Starr (2022) report that about 70% of workers in knowledge-intensive industries who are subject to NCAs are aware of the implications of what they signed. More recently, Cowgill, Freiberg, and Starr (2024) implemented a large-scale randomized field experiment with knowledge workers while randomizing whether the NCA clause is present in the employment contract (and how salient it is). Their findings appear consistent with our assumptions. For instance, they find that for highly skilled workers (defined as those making $40/hour), the presence of an NCA in an employment contract leads to a 7% lower likelihood of accepting a job offer (if the NCA is non-salient). This increases to 17% for a more salient NCA. They also track how much time prospective workers spend on reading each section of the employment contract and find that about 70% of highly skilled workers do not skip reading the section with the NCA (in the non-salient version) supporting the notion that such workers are aware and paying attention to NCAs.

2.4 Discussion and Conclusion

While extensive prior literature has conceptualized NCAs as beneficial for firms (Marx *et al.*, 2015; Starr *et al.*, 2018b), there may be downsides to using NCAs for some firms. Our study examines why some firms may opt out of using NCAs, even if they are available and are used by their competitors.

[43]We employed fuzzy matching of company names using the STATA command matchit, followed by manual checks to detect false positives.

We provide a conceptual explanation highlighting the tension between the ability to attract and retain talent, leading to firm-level heterogeneity in using or not using NCAs.

We find that non-leaders that rely more heavily on talent are less likely to use NCAs (relative to leaders that rely on talent). Importantly, these firms are more likely to respond that they opt out of NCAs because NCAs can lower their ability to attract talented workers. Further, by opting out of NCAs, the focal firms are better able to fill vacant skilled positions such as engineers. We also find that firms opting out of NCAs due to their need for talent are likely to have workers who are more satisfied and communicate better with managers. Lastly, the patterns are more pronounced in less patent-intensive industries, underscoring that concerns about knowledge leakage may drive the results.

Our theoretical framework developed above aligns with these findings. In a competitive context, for non-leading firms, acquiring highly skilled employees may be more important than protecting their existing knowledge from expropriation. For firms that report talent as the key differentiator, attracting a highly skilled and motivated workforce may be particularly critical. To the extent that NCAs diminish the attractiveness of these firms, NCAs undercut the firms' key advantage. While talent may be essential for non-leaders, leading firms may rely on valuable complementary resources such as brands, dominant distribution channels, or client relationships in addition to skilled workers. Relative to non-leaders, leading firms may have cutting-edge knowledge and IP and, thus, may worry more about the diffusion of such knowledge to their competitors. Consequently, for leading firms, retaining workers may be more critical, and such firms may opt for more extensive NCA use. Our analysis indicates that opting out of NCAs to attract talent correlates with a work environment where workers are more satisfied and communicate better with management. This may imply that opting out of NCAs is part of a broader human capital strategy oriented toward the attraction and retention of skilled employees. The firms that implement such strategies may be seen as creating a work climate where employees prefer to stay relative to competitors that employ legal levers to retain their workers.

While the explanation of our findings applies to NCAs, it may be useful to discuss several broader implications. Our logic may apply to other restrictive practices that are widespread, effective, and require firm commitment (e.g., non-poaching or non-solicitation agreements [Balasubramanian, Starr, and Yamaguchi, 2023]). In contrast, it may be more difficult for some firms to opt out of patent enforcement as a differentiation strategy. Patent litigation results from a potential ex-post infringement instead of a violation of an employment contract. Thus, the opt-out strategy may be seen as "cheap talk" as it does not require commitment at the time of hiring. We leave the investigation of how our results extend to other restrictive clauses for future work. Further, our findings indicate that non-leader firms primarily differentiate from leading firms that use NCAs very aggressively and use NCAs for all or most of their workers (see Appendix Table A17). This may be because such leading firms are highly visible, and it may be easier to draw a contrast with such firms when recruiting talent. Such dynamics may extend to other restrictive practices.

Our study has several limitations that open avenues for future work. The usual tradeoffs stem from our use of survey data. In return for a high level of granularity and detail, some of the data may reflect subjective assessments of the respondents. For example, whether a firm is an industry leader or not may be open to interpretation by respondents, leading to noise or bias. While the fact that most of the survey respondents were in managerial positions or HR jobs (see Appendix Table A1) may partially alleviate this concern, we sought to validate our leadership measure by matching the public firms in our sample with Compustat. We generally found a good correspondence between our proxy and other measures of industry leadership (see Appendix Table A24). Still, it will be helpful for future work to replicate our results using a broader range of measures.

The survey-based focus may affect the sample's representativeness and the generalizability of the results. While PayScale sent their survey to a representative cross-section of firms in terms of size and industry, it is possible that the response rate may be a function of some variables of interest in our model. We have used raking weights matching on size, industry, and states, and imputation methods for missing NCA questions to mitigate potential biases stemming from the selection in the survey participation and the absence of observations (Appendix Tables A15, A19, and A20). However, it is still possible that failing and low-performing firms may be less likely to respond to the PayScale survey. Those firms may face severe difficulties attracting talented workers whether they use NCAs or not. Given the low benefits of not using NCAs to attract talented workers, failing firms may opt for using NCAs and settle for less competitive workers. Considering this possibility, our results may not be generalizable to low-performing and failing firms, and future work should examine this population in more detail.

Further, the objective of our analysis was primarily descriptive while focusing on ruling out alternative explanations and spurious patterns. Our core objective is to explain the variation in the strategic behavior (i.e., opting out of NCAs). An ideal design would require strict exogeneity of a firm's leadership position, how much it relies on skilled human capital, and of the matching between human capital and firms. Implementing such a design in the context of our data is impossible. For instance, one may be concerned that a firm's choice to use or not use NCAs may be driven by the level of human capital it possesses. Non-leaders may tend to use NCAs less simply because they can hire only lower-level human capital. While our finding that firms choose not to use NCAs due to the lower ability to attract workers mitigates this concern, future studies could consider the quality of human capital the firms can acquire.

Our analysis is also cross-sectional in nature. Future work may focus on the temporal aspects of NCA use. For instance, it would be useful to examine how firms create and develop their NCA strategies over time. Such an approach would also open avenues for incorporating other theoretical perspectives that deviate from our rational choice framework such as the behavioral theory of the firm.[44] Future work may also examine the interaction between the state-level variation in NCA enforceability and the firm-level NCA use.

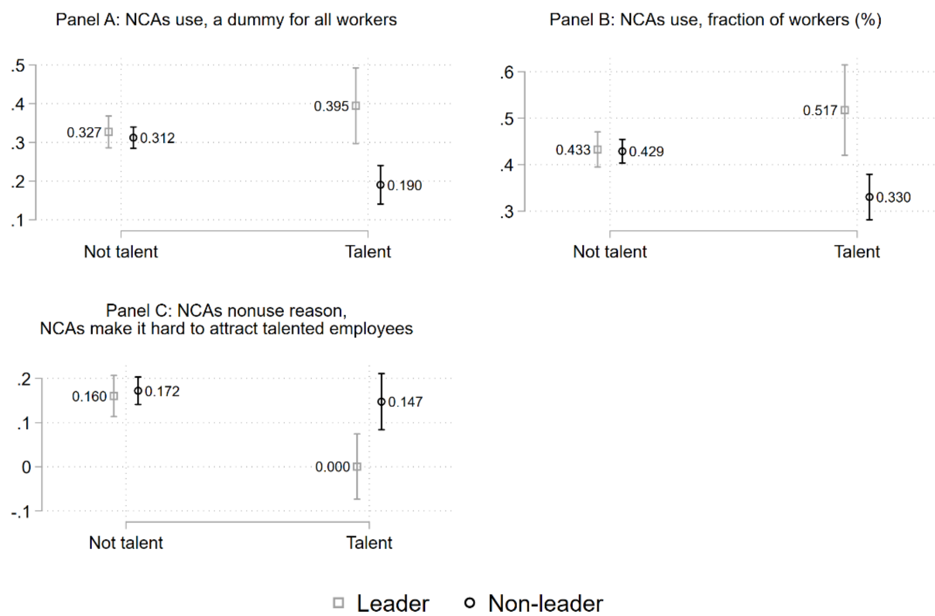
Our study provides one of the first investigations of firm-level heterogeneity in NCA use, and we make several contributions to the literature. We contribute to the literature on strategic human capital and mobility frictions by developing a novel explanation for the firm-level heterogeneity in the use of NCAs within industries. Relatedly, it is useful to connect our findings to the broader discussions in the field on the specificity of human capital (Campbell *et al.*, 2012a). Prior work has emphasized that mobility frictions allow firms to lock employees into their organizations, improving both their value creation and value appropriation associated with human capital (Starr *et al.*, 2018a). Our study presents an important complementary view. Improving value appropriation and “de facto" firm-specificity of human capital may be at the expense of the quality of talent that firms can hire. In the context of frictions, we should focus more on examining the impacts of frictions on hiring and not only on knowledge expropriation and existing workers.

We also contribute to the literature on NCAs. Most of the existing work examined the relationship between the state-level differences in the enforceability of NCAs and various individual- and firm-level outcomes. We shift attention to a more granular analysis at the firm level and highlight the importance of NCA nonuse. We show that such a shift has important ramifications by revealing significant heterogeneity in firms’ actual use of NCAs. Examining firm-level behaviors more directly will reduce the impact of confounding factors and potentially improve the reliability of analyses in future studies. Future work on NCAs may collect more granular data and go beyond examining legislative changes at the state level only. We hope to stimulate more research in this fruitful and important area.

[44]We note that such analysis would focus on the within-firm differences using firm fixed effects. Using firm fixed effects is not relevant in our context because it would remove the primary source of heterogeneity.

2.5 Appendix

Figure 1: Predicted likelihood of NCAs use and reasons for NCAs nonuse



Notes: The regression models for calculating predicted likelihoods are based on Table 4 Model 4 (Panel A), Model 8 (Panel B), and Table 5 Model 4 (Panel C).

Figure 2: Marginal effects of NCA non-use on the ability to attract talent

Notes: The regression models for calculating predicted likelihoods are based on Table 6 Models 1 and 4 for Panel A, Models 2 and 5 for Panel B, and Models 3 and 6 for Panel C.

Table 1: Summary statistics and pairwise correlations

	Variable	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	NCA use: a dummy for all employees	.31	.46	1													
2	NCA use: a dummy for all or some employees	.72	.45	.43	1												
3	NCA use: fraction of workers	.43	.44	.91	.63	1											
4	Non-leader	.69	.46	-.04	-.03	-.03	1										
5	Talent	.19	.39	-.02	-.07	-.01	.09	1									
6	Firm size: a dummy for 100 employees or more	.47	.50	-.14	.12	-.11	-.12	-.06	1								
7	NCA use is common in local market	.60	.49	.26	.50	.38	.00	-.01	.02	1							
8	Share of low-wage employees	.35	.27	-.11	-.05	-.13	.01	-.09	.02	-.10	1						
9	Respondent job function: HR and compensation	.28	.45	-.03	-.06	-.03	-.04	.05	-.16	-.02	.02	1					
10	Respondent job function: Executive	.13	.34	-.02	-.03	-.02	.02	.01	-.25	-.01	.03	-.25	1				
11	NCAs make it hard to attract talent	.15	.36	.	.03	-.03	.04	.00	.04	.02	-.06	.03	-.03	1			
12	Long-term vacant position	.31	.46	-.03	.00	-.02	.02	-.05	.11	.02	-.02	-.02	-.01	.08	1		
13	Difficulty in filling managerial jobs	.08	.28	.01	.04	.01	.04	-.03	.08	.01	-.05	.00	-.02	-.01	.46	1	
14	Difficulty in filling technical jobs	.14	.35	-.02	.06	.00	-.07	-.02	.18	.04	-.08	-.05	-.08	.06	.62	.33	1

Table 2: Propensity to use NCAs for No, Some and All employees

	P(No emp.)	P(Some emp.)	P(All emp.)
All firms	28.3	40.3	31.4
Firm size: 1-99 employees	33.2	29.1	37.6
Firm size: 100 or more employees	22.6	53.0	24.4
Industry non-leader (Non-leader = 1)	29.0	42.2	28.7
Industry leader (Non-leader = 0)	26.3	40.9	32.8
Reliance on talent (Talent = 1)	34.7	35.6	29.7
Reliance on other resources (Talent = 0)	26.8	41.2	31.9
NCA use is common in local markets	9.8	48.6	41.7
NCA use is uncommon in local markets	55.2	28.1	16.6

Notes: All variables in the table are either dummies or originally categorical but dichotomized.

Table 3: NCA use and reasons for not using them across industries

	Ratio of employees with NCAs (%)							HHI	N	“NCAs make it hard to attract talent" (ratio of firms not using NCAs at all or using them only for some workers)
	0	1-20	21-40	41-60	61-80	81-99	100			
Marketing & PR	18.4	10.5	2.6	5.3	0.0	0.0	63.2	0.447	38	33.3
Technology	20.1	13.4	4.0	4.0	1.3	3.6	53.6	0.350	224	14.8
Business & Management	24.7	16.9	3.9	6.5	6.5	1.3	40.3	0.262	77	19.4
Biotech & Science	24.4	17.1	7.3	0.0	7.3	4.9	39.0	0.254	41	5.0
Manufacturing	19.4	29.4	7.7	4.0	2.0	2.3	35.1	0.256	299	11.8
Accounting & Finance	31.1	22.7	4.2	2.5	4.2	1.7	33.6	0.266	119	17.2
Transportation	27.5	25.0	12.5	7.5	0.0	0.0	27.5	0.235	40	11.1
Energy/Utilities	38.6	15.9	9.1	2.3	0.0	6.8	27.3	0.262	44	10.0
Architecture & Engineering	38.2	21.1	6.6	3.9	1.3	2.6	26.3	0.266	76	34.7
Other	35.3	23.5	5.9	6.5	1.3	2.0	25.5	0.253	153	17.0
Customer Service	48.5	15.2	6.1	0.0	3.0	3.0	24.2	0.322	33	14.3
Food, Beverage & Hospitality	36.4	34.8	1.5	0.0	3.0	0.0	24.2	0.314	66	12.2
Medical & Healthcare	27.1	31.8	9.9	4.2	1.0	3.6	22.4	0.237	192	8.7
Retail	38.3	28.4	6.2	2.5	1.2	2.5	21.0	0.276	81	15.1
Real Estate	59.3	20.4	3.7	1.9	0.0	0.0	14.8	0.416	54	13.2

Notes: Among 1,803 observations where NCA questions are not missing, only industries with over 30 non-missing observations are displayed. HHI is calculated based on the seven categories of employee ratios with NCAs. The industries are ordered by the average NCA use in descending order. The last column shows the percentage of firms that indicated talent attraction as the reason for not using NCAs among firms not using NCAs at all or using them only for some workers.

Table 4: NCAs use as depending on leadership and reliance on talent

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Dependent Variable	NCA use: a dummy for all employees				NCA use: fraction of employees			
Non-leader ∗ Talent		**-0.151**	**-0.190**	**-0.189**		**-0.125**	**-0.174**	**-0.183**
		(0.080)	**(0.079)**	**(0.077)**		**(0.080)**	**(0.081)**	**(0.076)**
Non-leader	-0.036	-0.013	0.006	-0.015	-0.024	-0.006	0.015	-0.004
	(0.027)	(0.031)	(0.030)	(0.030)	(0.026)	(0.029)	(0.028)	(0.028)
Talent	-0.057	0.058	0.070	0.068	-0.057	0.038	0.071	0.085
	(0.033)	(0.069)	(0.069)	(0.068)	(0.033)	(0.069)	(0.070)	(0.066)
Constant	0.337	0.322	-0.013	-0.066	0.447	0.434	-0.017	-0.139
	(0.026)	(0.027)	(0.090)	(0.094)	(0.025)	(0.026)	(0.090)	(0.087)
Controls	No	No	No	Yes	No	No	No	Yes
Industry FE; State FE	No	No	Yes	Yes	No	No	Yes	Yes
Observations	1,232	1,232	1,232	1,232	1,166	1,166	1,166	1,166
R-squared	0.004	0.007	0.112	0.197	0.003	0.006	0.131	0.267
Mean of DV	0.302	0.302	0.302	0.302	0.419	0.419	0.419	0.419

Notes: Linear probability models in Models 1-4 and OLS in Models 5-8. The dependent variable in Models 1-4 is based on: "Which employees at your organization are subject to non-compete agreements (Prohibited from joining or starting a competing organization)?" It is coded as one for 'All employees' and zero for 'Some employees' or 'No employees,' while the 'Don't know' answer is missing. In Models 5-8, we use the follow-up question: "To the best of your knowledge, what percentage of all employees within the organization have signed non-competes?" There are five possible choices: '1-20%,' '21-40%,' '41-60%,' '61-80%,' and '81-100%.' We take the midpoint of each category for 'Some employees' and treat 'All employees' as 100% and 'No employees' as 0%. Standard errors (in parentheses) are clustered at the industry-by-state level.

Table 5: Using NCAs makes it difficult to attract talent

Dependent Variable	Model 1	Model 2	Model 3	Model 4
	Reason for NCAs non-use: NCAs make it hard to attract talented employees			
Non-leader ∗ Talent		**0.149**	**0.134**	**0.135**
		(0.049)	**(0.068)**	**(0.069)**
Non-leader	0.040	0.019	0.014	0.012
	(0.030)	(0.034)	(0.034)	(0.034)
Talent	-0.030	-0.151	-0.154	-0.160
	(0.033)	(0.028)	(0.054)	(0.055)
Constant	0.137	0.151	0.380	0.435
	(0.026)	(0.028)	(0.200)	(0.197)
Controls	No	No	No	Yes
Industry FE; State FE	No	No	Yes	Yes
Observations	691	691	691	691
R-squared	0.003	0.007	0.171	0.182
Mean of DV	0.159	0.159	0.159	0.159

Notes: Linear probability models. The sample only consists of the respondents who answered ‘Some employees’ or ‘No employees’ to the NCA-use question. The dependent variable is based on the question about the reasons for not using NCAs. The question is: “Why doesn’t your organization use non-competes (multiple selections allowed)?" There are six possible answers to this question: ‘Non-competes are not commonly used in the industry,’ ‘Non-competes make it hard to attract talented employees,’ ‘Loss of employees to competitors is not a big concern,’ ‘Not familiar with what non-competes are,’ ‘Non-competes are not legally enforceable in my state,’ and ‘Other (please specify).’ The dependent variable is coded as one if the respondents chose ‘Non-competes make it hard to attract talented employees,’ and zero otherwise. Standard errors (in parentheses) are clustered at the industry-by-state level.

Table 6: Ability to fill vacant positions

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Subsample	Firms relying on talent as the key differentiator (Talent = 1)			Firms relying on other resources as the key differentiator (Talent = 0)		
Dependent Variables	Long-term vacant position	Difficulty filling managerial jobs	Difficulty filling technical jobs	Long-term vacant position	Difficulty filling managerial jobs	Difficulty filling technical jobs
Non-leader ∗ NCA non-use percentage	**-0.190**	**-0.148**	**-0.362**	**0.031**	**0.024**	**-0.061**
	(0.234)	**(0.180)**	**(0.211)**	**(0.096)**	**(0.061)**	**(0.078)**
Non-leader	0.412	0.245	0.357	0.016	0.019	-0.019
	(0.179)	(0.121)	(0.169)	(0.068)	(0.043)	(0.049)
NCA non-use percentage	-0.006	-0.055	0.253	-0.005	-0.024	0.062
	(0.211)	(0.155)	(0.171)	(0.077)	(0.047)	(0.064)
Constant	0.133	-0.186	-0.179	0.194	-0.012	0.220
	(0.417)	(0.165)	(0.466)	(0.336)	(0.096)	(0.285)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE; State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	147	147	147	685	685	685
R-squared	0.506	0.433	0.547	0.116	0.100	0.157
Mean of DV	0.252	0.075	0.129	0.317	0.083	0.152

Notes: Linear probability models. Controls include firm size, the dummy for the prevalence of NCAs in local labor market, primary deliverable, Share of low-wage employees and the respondents' job functions. In Models 1 and 4, the dependent variable is based on the response of 'Yes' or 'No' the question: "Do you have any positions that have been open for six months or more?" Among those responding 'Yes,' the survey follows with: "What kind of positions do you have a hard time filling? (check all that apply)." There are nine possible choices that are not mutually exclusive: 'Management,' 'IT,' 'Customer Service,' 'Sales,' 'Executive Level,' Marketing,' 'Engineering,' 'Finance,' and 'Other (please specify).' The dependent variable in Models 2 and 5 is based on the choice for managerial jobs ('Management' and 'Executive Level'), while the dependent variable in Models 3 and 6 is based on the choice for technical jobs ('IT' and 'Engineering'). Standard errors (in parentheses) are clustered at the industry-by-state level.

3.1 Introduction

Scarcity leads to competition. It applies to educational opportunities as well. To get into a top university, being smart is not enough. In the US, wealthy parents can send their kids to expensive private schools or live in a better school district. The latter choice is also costly in the form of higher rents or property taxes. In many other countries, such as Turkey and South Korea, parents compete through private after-school tutoring, which prepares students for college entrance exams. In 2010, Korean families spent 10.7% of their income on such informal education for each student (OECD 2012). The industry has grown exceedingly large. The expenditure on private tutoring amounts to 1.8% of Korean national GDP and 54.6% of the annual budget for public education in 2009 (Statistics Korea, 2010). The public has been complaining about the skyrocketing tutoring expenditure. However, many rounds of reforms proves unsatisfactory.

The Korean case is not unique. In many countries, enrollments of elite universities are inelastic, and students take tutoring to catch up, keep up, or get ahead of their peers in admission. What complicates the design of policies is that individual spending on tutoring, or, more generally, on education, not only depends on observable and unobserved individual characteristics (e.g., preference), but also responds to and influences other students' spending. Another difficulty is due to the dual function of educational investment: it generates genuine human capital and signals in the admission tournament. Take a tax on private tutoring for example. It is likely to decrease tutoring spending. However, it also depresses human capital formation before college and affects the ranking orders and, thus, the college and labor market outcomes of students. The optimal policy need to take these quantitative implications into account.

This paper studies policy designs in a structural model. The model features an admission tournament, in which households purchase tutoring service to compete for the fixed capacities of selective colleges. I also allow tutoring service to influence human capital formation and signals production at the same time. A structural approach is necessary for counterfactual policy evaluations mainly for three reasons. First, the determinants of individual tutoring spending, such as the distribution of tutoring spending and the admission cutoffs, are equilibrium objects that are not invariant to policy changes. Second, it helps quantify the two roles of educational investment, which is needed in the quantitative analysis. Last, the structural model can be used for a wide range of policy experiments, including the ones that have never been tried.

The model captures several salient features of the high-stake college admission process. In the model, selective colleges have a fixed supply of seats, and the admissions is a rank-order tournament depending on the relative signals (Lazear and Rosen, 1981). Students are heterogeneous in ability, family wealth, and preferences for colleges and choose the level of private tutoring investment, which raises one's signals and human capital. The admission probability is determined in equilibrium such that given admission cutoff, the number of seats in selective colleges is equal to the number of attendees. The human capital distribution, admission cutoff, and college assignment are all equilibrium outcomes.

The model quantifies the dual function of tutoring investment. On the one hand, it produces genuine human capital which can be useful in labor market, as emphasized by Becker (1962) and Mincer (1974). On the other hand, it leads to better signals and, hence, admission to a higher ranked college. As the admission depends on the rank of signals, human capital is generally over invested in response to competitive pressure. The two channels can have quite different policy implications. If tutoring works only through the human capital channel, then the existence of a tutoring market is generally good, and the necessary policy intervention is to subsidize tutoring market or to provide students with credit. If private tutoring works only through the signaling channel, then we have two consequences. First, the positional externality of the admission tournament implies over-investment of private tutoring. Subsidy or cheaper credit may only exacerbate the wasteful investment. Second, with the presence of borrowing constraints, the very existence of the tutoring market can propagate the advantage of wealthy families, and subsidy or cheaper credit may offset some of that distortion in the student-college assignment. It is not true that we can do better by simply cracking down the tutoring market. Any reform about it must balance the two functions and pay attention to distributional consequences.

I estimate the model with a nationally representative sample in South Korea, the Korean Educational Longitudinal Study of 2005. The information on tutoring expenditure, academic performance and post-college outcome allows me to separate the effects of tutoring on human capital and on signals. The unobserved preference for colleges can be revealed from one's tutoring choice conditional on her initial academic performance. I find that both functions of tutoring are economically and quantitatively important.

As a policy experiment, I explore the implications of taxing and subsidizing the tutoring market. The experiment helps understand how peer competition shapes tutoring expenditure, and to what extent tutoring magnifies achievement gap in an admission tournament. Subsidizing the tutoring services increases the overall human capital. Households with medium income are most responsive to the price reduction under subsidy, by increasing their tutoring expenditure. Tax has the opposite effects: tutoring expenditure and human capital get lower. I further evaluate the impact of information. I find that reduced signal noise incentivizes the tutoring expenditure of high-ability students, whereas discourages the low-ability students due to the more "rigid" ranking order. High-ability students are benefited in college assignment.

This paper is related to a small but growing literature on endogenous pre-college human capital formation in an admission tournament. These studies focus on student effort under various admission policies. For example, Bodoh-Creed and Hickman (2018) and Grau (2018) study various forms of affirmative action rules, using B&B data and Chilean administrative data respectively. Arslan (2018) emphasizes the role of preference over colleges using Turkey college admissions data. These studies do not consider borrowing constraint – or to say, they interpret the cost as disutility. Myong (2018) investigates the effects of different scholarship on student effort, and the borrowing constraint is assumed on attending private high schools and colleges. Domina (2007) finds that more access to scholarship in universities leads to increased attendance of advanced courses in high school. As mentioned above, the tutoring services can be bought from the market, so that they are quite different from utilitarian effort costs – the borrowing constraint can interact with wealth levels of students and play an important role in the admission tournament. The first contribution of this paper is to quantitatively analyze variable educational expenditure (i.e., tutoring) with borrowing constraint in an admission tournament. Note that studying tutoring service has another advantage: It is observable in data.

The second contribution of this paper is to separate and quantify the two outcomes of educational investment (including effort): genuine human capital and signals in the admission tournament. Existing studies use academic achievement (such as GPA, test scores) as a proxy for both pre-college human capital and signals. This is fine under two alternative assumptions: First, the two are perfectly correlated. Second, pre-college human capital plays no role in subsequent studies and work. If either assumption is true, or to say, if we do not take for granted that one of these two assumptions is satisfied, then we should allow genuine human capital to be different from signals. That is what this paper plans to do. I model the production of signals and pre-college human capital. The two production functions can be separately identified with the information of post-college outcomes.

The paper also contributes to the literature on private tutoring. One strand of the studies have investigated the effect of tutoring on academic outcome and obtained mixed results (Dang, 2007, for Vietnam; Gurun and Millimet, 2008, for Turkey; Ono, 2007, for Japan; Ryu and Kang, 2013, for Korea; Zhang, 2013, for China). Another strand of literature examines the policy impacts on tutoring expenditure in Korea, and find little to no effect (Choi and Choi, 2016; Kim and Lee, 2010). One of the limitations of these studies is as follows: even if we understand what policies can reduce the overall tutoring spending, we are still not sure whether they improve welfare. A structural model can help quantitatively evaluate a wide range of policy candidates, including taxing private tutoring, expansion of selective universities, and adjusting admission policies. These experiments should be not only interesting for Korea but also a broad set of countries with high-stake exams and active private tutoring market.

This paper is organized as follows. Section 2 lays out the model and estimation strategy. Section 3 describes the institutional background and data. Section 4 presents the estimation results of the baseline model, including parameter estimates and model fit. A few counterfactual policy experiments are displayed in Section 5. Section 6 concludes.

3.2 Background and Data

Institutional Background

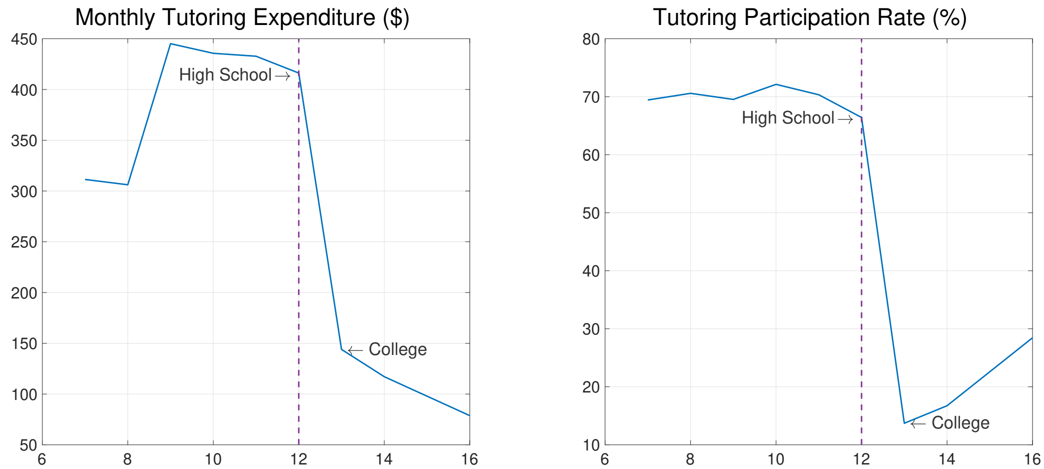
The academic rat race among Korean high-school students for college admissions is an annual competition for seats at a diverse set of tertiary institutions from most prestigious universities to two-year colleges. College rankings are fairly well-agreed upon and stable over time. Graduating from prestigious universities brings substantial economic and non-economic premiums. For example, the top three Korean colleges, accommodating 1% of college graduates, account for 74% of the CEOs (Lee, 2007), 63.7% of senior officials and 58.1% of congressmen in South Korea (Chae, Hong, and Lee 2005).

The College Scholastic Ability Test (CSAT) scores play a key role in college admission. All high school students who intend to attend colleges must pass the annual national CSAT. Near-perfect CSAT scores are required at top three colleges. As of 2010, this national test consisted of 5 sections: Korean language arts, mathematics, English, social studies/science, and the second foreign languages. Students are informed of the scores and percentile rankings of all subjects before application. Admission quotas are pre-specified and determined by Korean Ministry of Education. Colleges have an explicit formula, including weight, to calculate the final score in admission. There are two rounds of admission each year: early decision and regular admission. The early decision is based on a combination of high school records, CSAT scores, extracurricular activities, recommendation letters and interview, while regular admission relies exclusively on the ranking of CSAT scores.

Throughout high school, students exert time and money to prepare for standardized tests. The use of tutoring is prevalent and primarily for academic purposes. In secondary education, 90% of tutoring expenditure is for academic purpose, among which 92% is spent on the commonly administered subjects in CSAT: Korean language arts, mathematics and English (2010 Survey of Private Education Expenditure). Since poor parents are not able to foot the bills for private tutoring, the heavy reliance on private tutoring in Korea creates an inequitable distribution of education resources.

Tutoring expenditure drops dramatically after college attendance. Figure 1 shows the tutoring participation rate and the average monthly expenditure among the participants. Around 70% of students in secondary school take tutoring. The participation rate drops 50% after attending college. The average monthly tutoring expenditure falls to one-third. The dramatic decline cannot be easily rationalized by the human capital motive alone, suggesting that the incentives to compete for good colleges can play an important role in tutoring decision.

Figure 1: Motivating Fact



Data

The Korean Educational Longitudinal Study of 2005 is used in analysis. The KELS 2005 is a longitudinal survey that began in 2005 with a nationally representative sample of 6,908 Korean seventh graders (first year middle school). The survey follows the cohort annually before 2012 and biennially afterwards. The data includes information on the students' academic performance as measured by GPA and standardized test scores, tutoring expenditure, high school characteristics, family background, college attendance, and students' perceived labor market earnings. Information on initial academic performance and family background allows me to model the incentives facing households when they make tutoring decision. Pre-college test scores and post-college outcomes help disentangle the two effects of tutoring: producing genuine human capital and generating signals in the admission tournament.

The national College Scholastic Ability Test (CSAT) is is a standardized test that examines individuals' abilities for entering a university. It is held once per year and is made up of five sections. Korean language arts, mathematics and English are mandatory subjects that account for more than 60% of total points in CSAT. Science and Second Foreign Language sections are elective. Students choose one subject of each elective section depending on the majors they plan to apply for. I focus on the three mandatory subjects in measuring the initial academic performance, CSAT score and tutoring expenditure because CSAT scores of mandatory subjects are comparable across individuals. Besides, test scores of elective subjects are highly correlated with that of mandatory subjects with correlation coefficient 0.8.

In my analysis, colleges are grouped into three tiers with Tier 1 representing the top 15% of college seats. KELS 2005 dataset does not contain direct measure of college quality. I use the lowest CSAT score of students admitted through regular admission as a proxy for college quality. As the regular admission process is solely based on CSAT scores, the lowest score is a meaningful reflection of admission cutoff and hence college quality. I further assume that signals are commonly observed and evaluated in the same way by all colleges. The lack of college names limits one's ability to link a university identifier to its actual admission policy.

I focus on students who attend colleges right after graduating from high school, which represent 76% of the full sample. This number is higher than but still comparable to the national level of college enrollment rate 70% in the same year (Korea Educational Development Institute, 2017). High school graduates entering into labor market (6.64% of full sample) or retaking the national exam for college admissions (12.97% of full sample) are excluded from the analysis. I further restrict my analysis to a sample of 1300 students who have complete information on standardized test scores (grade 7, 12), tutoring expenditure (grade 7-12), household income (grade 7-12), high school and college characteristics.

Table 1 provides descriptive statistics for the estimation sample and the full sample. Initial test score, CSAT and GPA are normalized to have mean zero and unit standard deviation among the full sample. The mean test scores in the estimation sample is positive and has a standard deviation below one. This primarily reflects that the low-performing students are less likely to report their CSAT scores and are more likely to work right after high school graduation. Students in restricted sample attend better high schools and take slightly more tutoring, largely because I require students to attend colleges right after high school. Household income is measured by average monthly income after excluding the education spending on siblings. The distributions of household income and expected wage in estimation sample resemble that in full sample. Although the estimation sample includes more high-performing students than would a nationally representative sample, it is those students who actively take tutoring to compete for elite colleges, and are more responsive to policy changes regarding tutoring market and college admission. The estimation sample still covers a broad range of key players in the admission tournament.

Table 1: Descriptive Statistics

	Estimation		Full Sample		
	Mean	Std. Dev.	Mean	Std. Dev.	Obs.
	Panel A. Households Characteristics				
Male (%)	35.85		52.37		6908
Initial Test Score (Grade 9)	0.4152	0.9188	-7.0e-8	1	6622
CSAT Score (Grade 12)	0.1727	0.9286	-1.4e-7	1	3857
High School GPA	0.2052	0.8801	-1.1e-7	1	4827
HH Income ($100/Month)	41.667	38.899	40.639	40.231	5100
	Panel B. School Characteristics				
High School Quality	0.6169	0.1855	0.5326	0.2280	5354
Tier 1 College (%)	15.00		12.55		3514
Tier 2 College (%)	56.23		52.65		
Two-Year College (%)	28.77		34.80		3514
Expected Wage ($100/Month)	17.428	5.6258	17.244	5.8078	2923
	Panel C. Choice Variables				
Tutor Expense ($100/Month)	2.8950	2.7458	2.7087	3.1289	5100

Unit: $ in 2010.

In survey, students are asked wage expectation in the year they graduate from college. There are six options of expected pre-tax annual income categories to choose from. Students are also asked actual monthly pre-tax income after they get employed. But because the most recent survey data made available is collected in 2014, which is the senior year for four-year college students, the actual earnings are not observed for the majority of the sample. To investigate how accurate the wage perceptions are, I compare the wage expectation with the actual wage, focusing on a sample of two-year college graduates. Table 2 presents the distribution of expected monthly wage. The median expected wage falls into category $(\$1105, \$1473]$. This is consistent with the median actual wage $1238, and mean actual wage $1245 of the same sample. While these data are based on small sample sizes and only two-year college group, they are still informative and suggesting that expectations data are predictive of actual realizations.

Table 2. Expected Monthly Wage of 2Yr College Graduates

Exp. Wage ($)	$\leqslant 737$	$(737, 1105]$	$(1105, 1473]$	$(1473, 1842]$	$(1842, 2210]$	> 2210
Percent (%)	2.53	30.38	39.24	22.15	4.43	1.27
Observations	4	48	62	35	7	2

Unit: $ in 2010.

3.3 Model

Environment

There is a continuum of households of unit mass. Each household has a child in high school, and is endowed with $X_i = (A_i, q_i, y_i, \nu_i)$. Ability A_i represents the child's stock of skills at the beginning of high school, and is perfectly measured by initial test score. q_i is the high school quality, y_i is household income. ν_i represents household i's preference for Tier 1 colleges. Individual preference ν_i is private information, while the population distribution is common knowledge.

Colleges are categorized into three tiers: high quality four-year (Tier 1), low quality four-year (Tier 2), and non-selective two-year colleges (Tier 3), with total mass one. Four-year colleges have higher quality than two-year colleges, which is measured by wage return. Anyone can attend two-year colleges, whereas the admission process for four-year colleges can be competitive because of capacity constraints. Colleges in the same tier are identical for a household. All households agree on the ranking of colleges. Colleges wish to admit the best students possible but genuine human capital is private information. They rank students based on commonly observed set of signals including test scores. Once a student enters college, her belonging household makes no other decisions: the college is an absorbing state.

The model starts from 1st year high school. Household i with $X_i = (A_i, q_i, y_i, \nu_i)$ chooses how much to spend on private tutoring, while taking into account how much they value colleges, and how tutoring decisions will affect their admission chances. At the end of high school, human capital is produced, signals crucial for college admission are generated. Students are assigned to colleges based on the rank-order of signals.

Admission

The admission policy is a combination of CSAT score, high school GPA and quality, other factors that are unobserved to econometrician. The admission criteria is pre-specified as

$$s_i = \delta_1 CSAT_i + \delta_2 GPA_i + \delta_3 q_i + \delta_4 A_i + \delta_5 A_i^2 + \delta_6 y_i + \delta_7 y_i^2 + \xi_i. \tag{3.1}$$

Here $CSAT_i$ represents scores in College Scholastic Ability Test, a national test held once per year. GPA_i is high school GPA, q_i is high school quality measured by the school's advancement rate into college. $\delta_4 A_i + \delta_5 A_i^2$ captures the student ability that is not reflected in end-of-high school test scores but is correlated with initial performance (A_i). This ability component can be observed by colleges through recommendation letter and during interview in the early admission. $\delta_6 y_i + \delta_7 y_i^2$ captures unmeasured admission signals such as extracurricular activities which matter for early admission decision. Households from wealthy background can afford and often spend significant amount of money building application packages including extracurriculars. ξ is a random matching shock commonly observable to colleges during admission process, but not to the student while making tutoring decision. ξ is assumed to be normally distributed for tractability, and is normalized as $\xi \sim N(0,1)$ as s is scale-free.

Signals Generation

Test scores are generated from

$$\begin{aligned} CSAT_i &= \gamma_{10} + \gamma_{11} e_i + \gamma_{12} e_i^2 + \gamma_{13} A_i + \gamma_{14} A_i^2 + \gamma_{15} q_i + \epsilon_{1i} \\ GPA_i &= \gamma_{20} + \gamma_{21} e_i + \gamma_{22} e_i^2 + \gamma_{23} A_i + \gamma_{24} A_i^2 + \gamma_{25} q_i + \epsilon_{2i} \end{aligned}, \tag{3.2}$$

where A_i is the initial ability measured by test score at the beginning of high school, e_i is the monthly tutoring expenditure during high school. Parameters $\gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}$ describe the signaling channel, in which tutoring improves test scores and thus admission chances. γ_{25} captures the "small-pond-big-fish" effect because GPA is not comparable across high schools. ϵ_1, ϵ_2 are independent shocks in scores generating process and follow normal distribution.

Preference

Households value current consumption, wages in labor market, non-pecuniary benefits from attending Tier 1 colleges. For tractability, the three components are assumed to be additively separable:

$$u_i = \ln(y_i - e_i) + \nu_i E_{\epsilon_1,\epsilon_2,\xi} I_{1i}(e_i) + \beta E_{\epsilon_1,\epsilon_2,\xi,\varepsilon_c,\varepsilon_w} \ln(w_i). \tag{3.3}$$

Here y_i is the household income available for household consumption and the student's education (excluding the education spending on siblings). Parameter β captures the importance of labor market payoff. ν_i represents one's discounted non-pecuniary utility value from attending Tier 1 colleges. The non-pecuniary benefits may include social status, alumni network etc. ν_i follows is assumed log-normally distributed, under which each individual strictly prefers Tier 1 colleges over other college tiers.

$I_{1i} \in \{0,1\}$ indicates whether student i gets accepted into Tier 1 colleges. At the time of choosing tutoring, household i formulates probability $E_{\epsilon_1,\epsilon_2,\xi} I_{1i}(e_i)$ of attending Tier 1 colleges. It is the probability that the signal surpasses admission cutoff c_1:

$$\begin{aligned} E_{\epsilon_1,\epsilon_2,\xi} I_{1i}(e_i) = Pr \quad & \{\delta_1 (\gamma_{10} + \gamma_{11} e_i + \gamma_{12} e_i^2 + \gamma_{13} A_i + \gamma_{14} A_i^2 + \gamma_{15} q_i + \epsilon_{1i}) + \delta_3 q_i + \delta_4 A_i \\ + \delta_5 A_i^2 + \delta_6 y_i + \delta_7 y_i^2 \quad & + \delta_2 (\gamma_{20} + \gamma_{21} e_i + \gamma_{22} e_i^2 + \gamma_{23} A_i + \gamma_{24} A_i^2 + \gamma_{25} q_i + \epsilon_{2i}) + \xi_i \geqslant c_1\} \end{aligned}. \tag{3.4}$$

Expected Wage

Logarithm of labor market entry wage is given by

$$\ln(w_i) = \rho_1 e_i + \rho_2 e_i^2 + \rho_3 A_i + \rho_4 A_i^2 + \rho_5 q_i + \sum_{j=1}^{3} r_j I_{ji} + \varepsilon_{ci} + \varepsilon_{wi}, \tag{3.5}$$

where r_j denotes the monetary payoff from attending colleges of Tier $j \in \{1, 2, 3\}$. ε_{ci} refers to the human capital shock realized during college, ε_{wi} is the wage shock realized in labor market. $\varepsilon_{ci}, \varepsilon_{wi}$ are assumed independent from all information one has prior to college entrance and with zero mean. Parameters ρ_1, ρ_2 capture the marginal productivity of tutoring expenditure in producing genuine human capital. After the realization of human capital shock in college ε_{ci}, a student forms expectation on her labor market outcome:

$$E_{\varepsilon_w} \ln(w_i) = \rho_1 e_i + \rho_2 e_i^2 + \rho_3 A_i + \rho_4 A_i^2 + \rho_5 q_i + \sum_{j=1}^{2} r_j I_{ji} + \varepsilon_{ci}. \tag{3.6}$$

Expected wage is surveyed at the end of college. This subjective expectation reflects a student's perceived monetary return to college. As the tutoring choice is jointly determined by perceived monetary return and non-pecuniary preference, observing expectations allows making more accurate inference on individual preference. For the purpose of estimating preference parameters, it would be ideal to have information on the expectations agents hold at the time of making tutoring choice. But due to data limitation, the end-of-college expectations are the closest approximations.

The validity of using end-of-college expectations hinges on two assumptions. Firstly, students report their expectation truthfully. This assumption is implicitly made when using any survey data and is not specific to expectations data. Second, students do not systematically change their beliefs on college premium (r_j) during college. Since the idiosyncratic match quality realized during college is embedded in the error term ε_{ci}, the changing expected college premium is mainly driven by college-specific information shocks and dropout decisions. Given the great emphasis on education and low dropout rate in Korea, it is not a strong assumption that households are well informed of the college premium when making tutoring decision.

The expected wage is also used to identify the human capital and wage formation. To get unbiased estimates, the expectations are required to be predictive of actual realizations. The wage comparison made in Table 2 suggests that the wage expectations are informative of the actual realizations.

Equilibrium

Households compete for the fixed and pre-determined amount of slots in selective colleges. The colleges capacity, production technology, admission criteria, and the joint distribution of household endowments are common knowledge prior to households' choices of tutoring. In a large contest with a continuum of households and under rational expectation, households can anticipate the correct admission cutoffs without uncertainty. Given admission cutoffs and with borrowing constraint, each household chooses the optimal tutoring expenditure:

$$\max_{e_i \geqslant 0} \ln\left(y_i - e_i\right) + \nu_i E_{\epsilon_1,\epsilon_2,\xi} I_{1i}\left(e_i\right) + \beta E_{\epsilon_1,\epsilon_2,\xi,\varepsilon_c,\varepsilon_w} \ln\left(w_i\right). \tag{3.7}$$

Consistent with the lack of financial loans designed for pre-college education, there is no borrowing possible to finance the tutoring cost. And consistent with the generous provision of financial aids in college, households are assumed to have access to perfect financial market during and after college. Therefore, the utility from workforce monetary payoff can be expressed as a function of expected present value of the lifetime earnings, that is, $\beta E_{\epsilon_1,\epsilon_2,\xi,\varepsilon_c,\varepsilon_w} \overset{[}{t} = 1]T \sum \frac{\ln\left(w_{it}\right)}{\left(1+r\right)^t}$. Since the wage measures is observed only once and in early career, assumptions have to be made on how wages evolve over the life cycle and across college types. The current version of wage equation (5) is time-invariant, which at least, implicitly assumes that the growth rate of wage is the same for all colleges.

The first-order condition gives:

$$\beta\left(\rho_1 + 2\rho_2 e_i\right) + \left(\beta r_1 + \nu_i\right) \frac{\partial E_{\epsilon_1,\epsilon_2,\xi} I_{1i}\left(e_i\right)}{\partial e_i} + \beta r_2 \frac{\partial E_{\epsilon_1,\epsilon_2,\xi} I_{2i}\left(e_i\right)}{\partial e_i} \leqslant \frac{1}{y_i - e_i}. \tag{3.8}$$

At the margin, households are trading off the tutoring cost with the future benefits of improving admission chances and obtaining human capital. The admission cutoff c_j of Tier $j \in \{1, 2\}$ colleges is determined by market clearing condition:

$$\int P\left(c_{j+1} \leqslant s \leqslant c_j\right) d\mathcal{F}\left(A, q, y, \nu\right) = \kappa_j, \tag{3.9}$$

where κ_j is the capacity of Tier j colleges. The number of admitted students is equal to the number of college seats, conditional on households' optimal tutoring choices.

Estimation Strategy

Under the assumptions that random terms $\{\epsilon_{1i}, \epsilon_{2i}, \xi_i\}$ are independent and normally distributed, I estimate the signal generation equations (2) by ordinary least squares, and the admission equation (1) with ordered probit. The expected wage equation (6) is estimated by maximum likelihood because expected wages are categorical in data. Let π_{ik} be the probability of expected wage being in category $(\varpi_{k-1}, \varpi_k]$, then

$$\begin{aligned}\pi_{ik} &= \Pr\left(\ln\left(\varpi_{k-1}\right) \leqslant E_{\varepsilon_w} \ln\left(w_i\right) \leqslant \ln\left(\varpi_k\right)\right) \\ &= \Phi\left(\ln\left(\varpi_k\right) - \rho_1 e_i - \rho_2 e_i^2 - \rho_3 A_i - \rho_4 A_i^2 - \rho_5 q_i - \textstyle\sum_{j=1}^{2} r_j I_{ji}\right) \\ &\quad -\Phi\left(\ln\left(\varpi_{k-1}\right) - \rho_1 e_i - \rho_2 e_i^2 - \rho_3 A_i - \rho_4 A_i^2 - \rho_5 q_i - \textstyle\sum_{j=1}^{2} r_j I_{ji}\right)\end{aligned}$$

Here $\Phi(\cdot)$ represents the cumulative normal distribution function with standard deviation σ_c. The log likelihood of observed wage expectations for a dataset with N observations and 6 wage categories can be written as

$$\mathcal{L}_1\left(\overrightarrow{\rho}, \overrightarrow{r}, \sigma_c\right) = \sum_{i=1}^{N} \ln\left(\Pr\left(E_{\varepsilon_w} \ln\left(w_i\right) \mid e_i, A_i, q_i, \vec{I_i}\right)\right) = \sum_{i=1}^{N} \sum_{k=1}^{6} \mathbb{1}\left(k\left(i\right) = k\right) \cdot \ln\left(\pi_{ik}\right).$$

The parameters of the utility function (3) are also estimated by maximum likelihood. Let $\lambda_i\left(e_i\right)$ be the likelihood of the household i choosing the observed tutoring investment. Conditional on initial endowment $\{A_i, q_i, y_i\}$, tutoring investment e_i is determined by non-pecuniary preference ν_i.

$$\lambda_i\left(e_i\right) = \Pr\left(e_i \mid A_i, q_i, y_i\right) = \Pr\left(\nu_i\left(e\right) \mid A_i, q_i, y_i, e_i\right).$$

The mapping $e \mapsto \nu_i(e)$ can be derived from the first order condition (8):

$$\nu_i\left(e\right) \leqslant \frac{\dfrac{1}{y_i - e_i} - \beta\left(\rho_1 + 2\rho_2 e_i\right) - \beta r_2 \dfrac{\partial E_{\epsilon_1,\epsilon_2,\xi} I_{2i}\left(e_i\right)}{\partial e_i}}{\dfrac{\partial E_{\epsilon_1,\epsilon_2,\xi} I_{1i}\left(e_i\right)}{\partial e_i}} - \beta r_1,$$

where "$\leqslant$" holds at corner solution $e = 0$. Therefore, $\lambda_i\left(e_i\right)$ can be further written as

$$\lambda_i\left(e_i\right) = \begin{cases} \Psi\left(\dfrac{\dfrac{1}{y_i - e_i} - \beta\left(\rho_1 + 2\rho_2 e_i\right) - \beta r_2 \dfrac{\partial E_{\epsilon_1,\epsilon_2,\xi} I_{2i}\left(e_i\right)}{\partial e_i}}{\dfrac{\partial E_{\epsilon_1,\epsilon_2,\xi} I_{1i}\left(e_i\right)}{\partial e_i}} - \beta r_1\right), & \text{if } e_i = 0 \\ \varphi\left(\dfrac{\dfrac{1}{y_i - e_i} - \beta\left(\rho_1 + 2\rho_2 e_i\right) - \beta r_2 \dfrac{\partial E_{\epsilon_1,\epsilon_2,\xi} I_{2i}\left(e_i\right)}{\partial e_i}}{\dfrac{\partial E_{\epsilon_1,\epsilon_2,\xi} I_{1i}\left(e_i\right)}{\partial e_i}} - \beta r_1\right), & \text{if } e_i > 0 \end{cases}$$

The preference parameters can be estimated from the log likelihood of observed tutoring investment:

$$\mathcal{L}_2\left(\beta, \mu, \sigma_\nu\right) = \sum_{i=1}^{N} \ln\left(\lambda_i\left(e_i\right)\right).$$

3.4 Estimation Results

Signals and Wage

Table 3 reports the estimates for the signals and wage equations. Tutoring expenditure exhibits diminishing marginal return. A \$100 change of monthly tutoring expenditure e from median level leads to 0.08 standard deviation change of CSAT score, and 0.02 standard deviation change in GPA. The marginal effect of tutoring on CSAT is 4 times as its effect on GPA. Tutoring plays a more important role on CSAT scores.

Tutoring produces genuine human capital. Conditional on ex-post college assignment, spending an extra \$100 on tutoring from median values every month can raise wage by 2.09%, equivalent to \$34 monthly wage gain for a mean wage earner. Given the wage gain over the life cycle, the labor market return to tutoring is sizable. Note that the human capital return could be overestimated driven by the omitted individual college quality. Conditional on college tiers, the omitted college quality is likely to be positively correlated with tutoring expenditure. This omitted variable bias can be resolved by adding college fixed effects. However, adding college tiers implies that the quality rankings are less likely to be agreed upon by all households, because households may have idiosyncratic preference for college characteristics such as location and amenity. This will complicate the rank-order tournament setup and create computational burden in solving equilibrium outcomes.

There are two channels that tutoring can affect wage: producing genuine human capital, improving admission probability. Figure 2 shows the relative magnitude of the two channels. The horizontal axis represents the quantile rank of household income and ability. The vertical axis describes the ratio of the marginal effect through improving admission probability to the total marginal effect. As the ratio is below 0.5 for all, tutoring impacts wage mainly through the production of human capital.

Ability takes a greater role in generating CSAT and GPA than in affecting wages. One standard deviation change from median initial test score can boost up CSAT by 0.55 standard deviation, but only raise wage by 2.76%. The estimates are consistent with a common view that previous academic performance helps one enter a better college, but post-college wage is mainly driven by the human capital in college, which is accrued primarily through college quality.

Figure 2: Admission and Human Capital Channels in Wage Determination

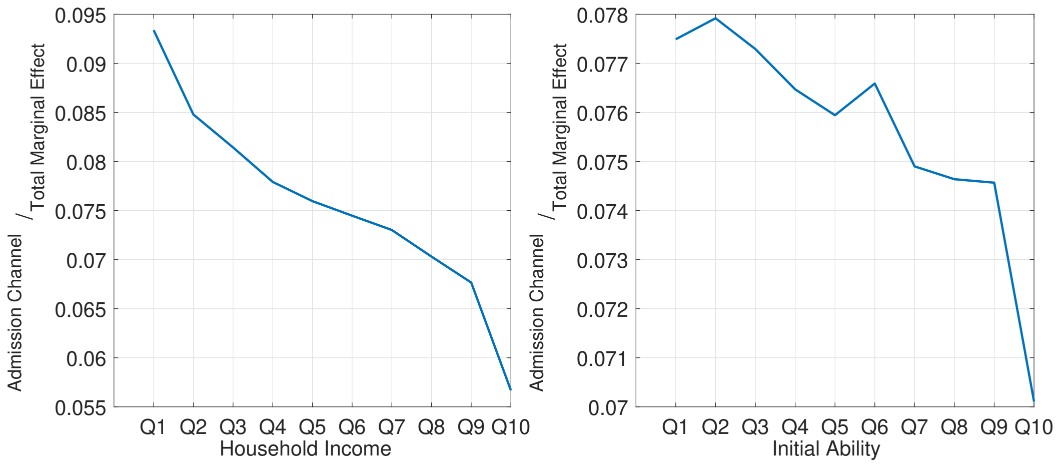


Table 3. Signals and Wage Parameters

	CSAT		GPA		ln(Wage)	
Variables	Coef.	St. Dev.	Coef.	St. Dev.	Coef.	St. Dev.
Tutor e	0.1119	0.0168	0.0356	0.0194	0.0250	1.89e-7
Tutor e^2	-0.0049	0.0015	-0.0027	0.0018	-0.0007	1.64e-8
Ability A	0.3848	0.0253	0.2358	0.0293	0.0092	2.22e-8
Ability A^2	0.1740	0.0221	0.1501	0.0255	0.0190	1.37e-7
HS Quality q	0.9055	0.1096	-0.8232	0.1268	0.0444	1.69e-7
Tier 1 I_1					0.2071	4.06e–7
Tier 2 I_2					0.0925	4.12e-7
Constant	-0.9685	0.0736	0.4028	0.0852	2.6119	1.37e-7
Std. Error	0.7060		0.8167		0.3013	1.15e-8

Admission

Estimates for admission criteria are displayed in Table 4. The marginal impact of one standard deviation change in CSAT is larger than the marginal effect of one standard deviation change in GPA. Conditional on CSAT score, household income and initial performance explain a substantial proportion of variation in admission outcomes. This is consistent with the fact that about 50% of students in sample enter colleges through early admission, where other criteria, such as essays and letters of recommendation, extracurricular activities, aptitude examinations or interviews.

Incorporating initial performance to admission equation weakens the marginal impacts of CSAT and GPA, and thus, the estimated effect of tutoring expenditure on admission chances. But as many tutoring centers help students prepare for application packages and college interviews in early admission, the model may underestimate the impact of tutoring on admission probability. The agents may be more responsive to college competition incentives than the model would suggest.

Table 4. Admission Preference Parameters

Parameter	Description	Value	St. Dev.
δ_1	admission weight on $CSAT$	0.4483	0.0597
δ_2	admission weight on GPA	0.3404	0.0456
δ_3	weight on high school quality	0.4616	0.2068
δ_4	weight on initial performance g	0.0935	0.2038
δ_5	weight on initial performance g^2	0.2033	0.0393
δ_6	weight on household income y	0.0062	0.0022
δ_7	weight on household income y^2	-1.4e-5	6.6e-6
c_1	admission cutoff of Tier 1 colleges	2.2474	0.1683
c_2	admission cutoff of Tier 2 colleges	0.1422	0.1523

Preference

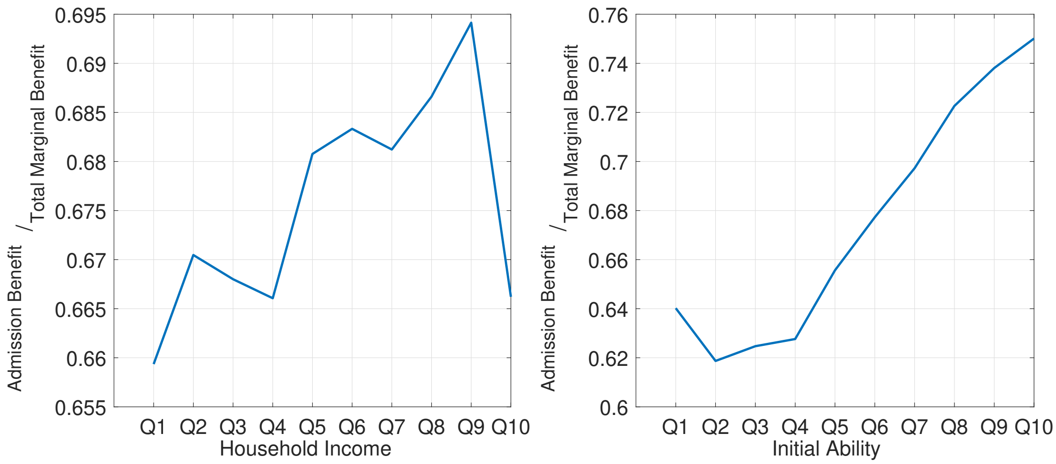
Table 5 describes the preference parameters. The estimates of the structural parameters indicate that while expected wage is a statistically significant determinant of the tutoring expenditure, they play a rather small role in the choice. Figure 3 compares the relative magnitudes of the marginal utility benefit of tutoring through wage $\frac{\partial E_{\epsilon_1,\epsilon_2,\xi,\varepsilon_c,\varepsilon_w}\beta \ln(w)}{\partial e}$ and the marginal benefit through college admission $\frac{\partial \nu_i E_{\epsilon_1,\epsilon_2,\xi} I_1(e)}{\partial e}$. The vertical axis displays the ratio of college competition incentive to the total marginal benefit of tutoring expenditure. The high ratio implies that competition for Tier 1 colleges is the driving force for tutoring investment. The competition incentive is stronger for high ability students. As a counterfactual exercise, I shut down the competition channel, so that tutoring investment cannot impact college assignment. Households significantly lower their tutoring expenditure, with the majority of households not even purchasing tutoring. This exercise suggests over-production of human capital in competing for prestigious colleges.

Table 5. Admission Preference Parameters

Parameter	Description	Value	St. Dev.
β	preference for $\log(\text{wage})$	0.0882	1.9e-5
μ	preference for Tier 1, $\ln(\nu) \sim N(\mu, \sigma_\nu^2)$	0.8073	2.2051
σ_ν	preference for Tier 1, $\ln(\nu) \sim N(\mu, \sigma_\nu^2)$	3.5169	4.9611

Non-pecuniary preference for Tier 1 colleges play a major role in tutoring choice. This is consistent with the substantial non-economic premiums of graduating from an elite college. But admittedly, the non-pecuniary preference ν may also captures pecuniary benefits, such as wage

Figure 3: Wage and Admission Incentives



growth, that are associated with Tier 1 colleges. Note that w_i measures one's wage expectation at the beginning of career. If graduates from Tier 1 colleges enjoy lower unemployment rate and higher wage growth, by construction, those benefits will be contained in preference term ν.

Model Fit

Figure 4 depicts the model predictions by household income, high school quality and initial academic performance. The tutoring expenditure of households from the top 5% income group is over-predicted. There are two possible explanations for this over-prediction. First, the model does not allow borrowing and lending before college. In the model, the opportunity cost of purchasing tutoring is the lost consumption. However in reality, households are trading off current consumption, human capital return, with financial return. It is the wealthy households who hold more financial asset and receive higher financial return. The heterogeneous asset return may explain the declining marginal tutoring spending with income. Second, the model does not consider the time constraint. Faced with binding time constraint, although wealthy households can afford and are willing to purchase more tutoring, students may not have extra time to learn.

There is a potential tension between model fit and the usefulness of the model for counterfactual analysis. To help improve model fit, mostly the convex relationship between the admission rate and initial ability in the data, I include initial ability and its quadratic term in the admission criteria (1). It captures the idea that initial ability may be observed by colleges through recommendation letters and interviews in the early admission process. But adding initial ability into the admission criteria also mechanically reduces the importance of tutoring in admission.[1] To make model fit better, more weight is given to initial ability, a predetermined variable, which effectively makes admission probability insensitive to policy changes. In addition, the current model still underpredicts the Tier

[1] Another issue is that initial ability should be useful for only a fraction of students who have a chance of early admission.

1 admission probability for students with top-decile ability (the third panel of the second row in Figure 4). Alternatively, one could replace initial ability and its quadratic form with a quadratic term of CSAT to help fit the convexity without putting too much weight on predetermined variables. But that comes at a cost: the admission probability would lose closed form solutions, due to the error term in CSAT in the quadratic term, so that estimation is computationally demanding and is therefore not pursued in the current version.

Figure 4: Model Fit

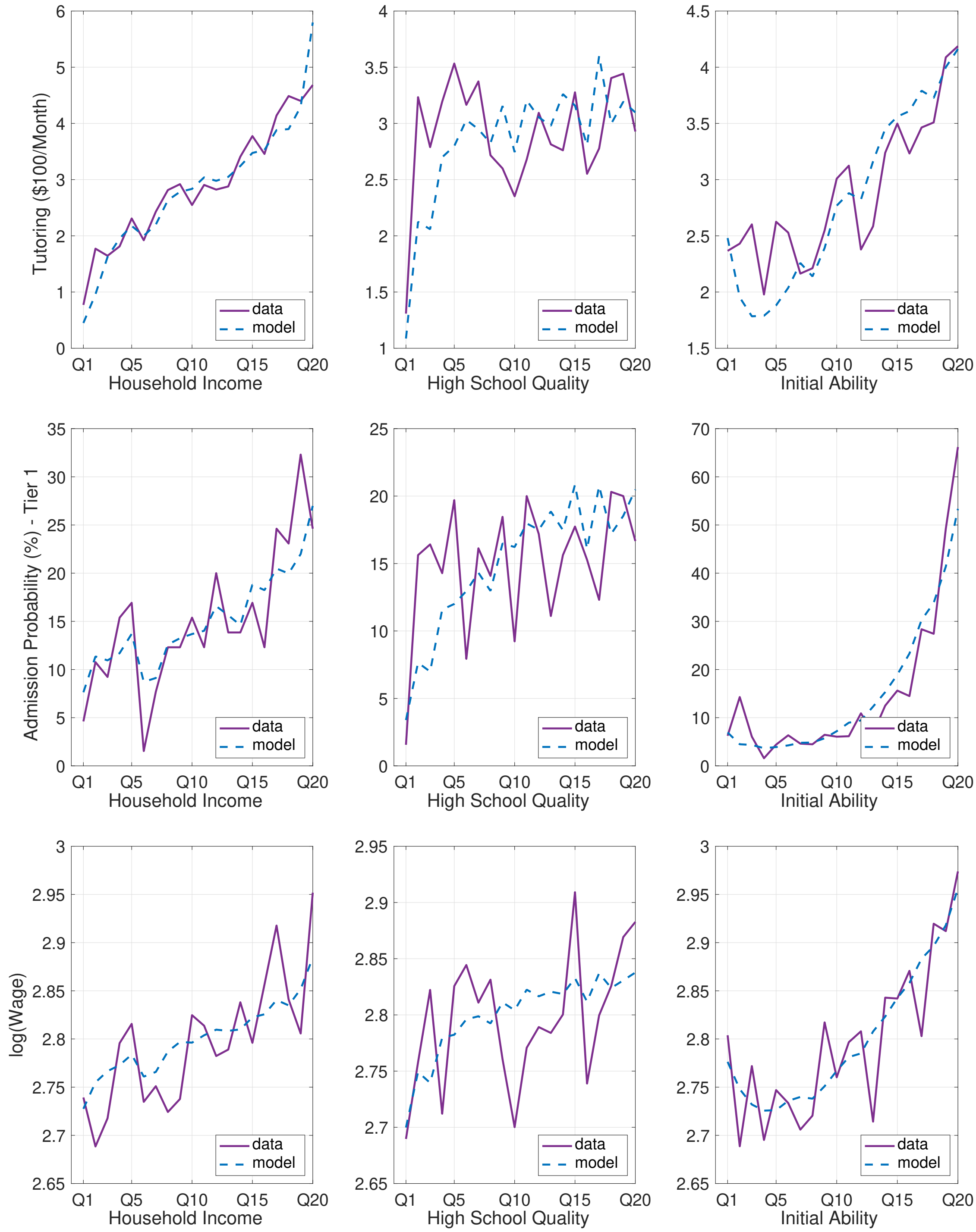
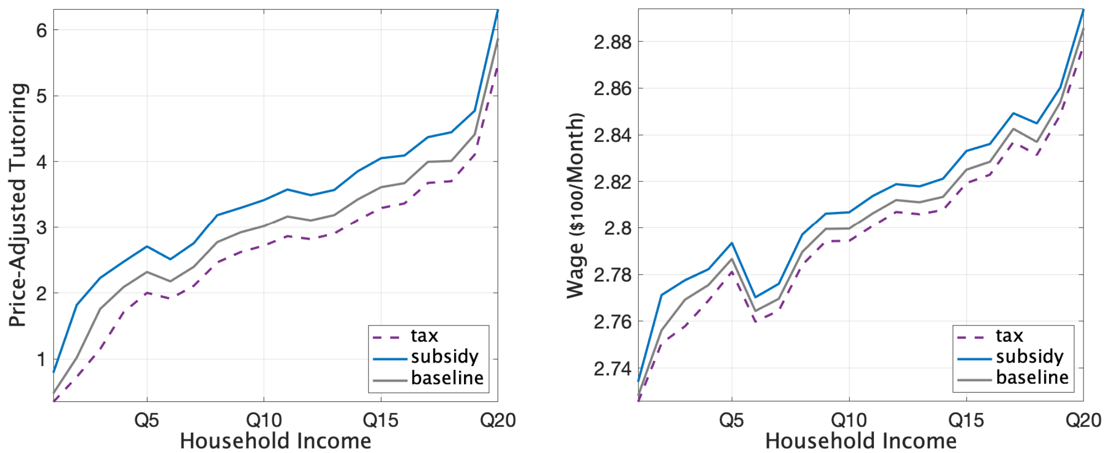


Figure 5: Counterfactual Tax and Subsidy



3.5 Counterfactual Experiments

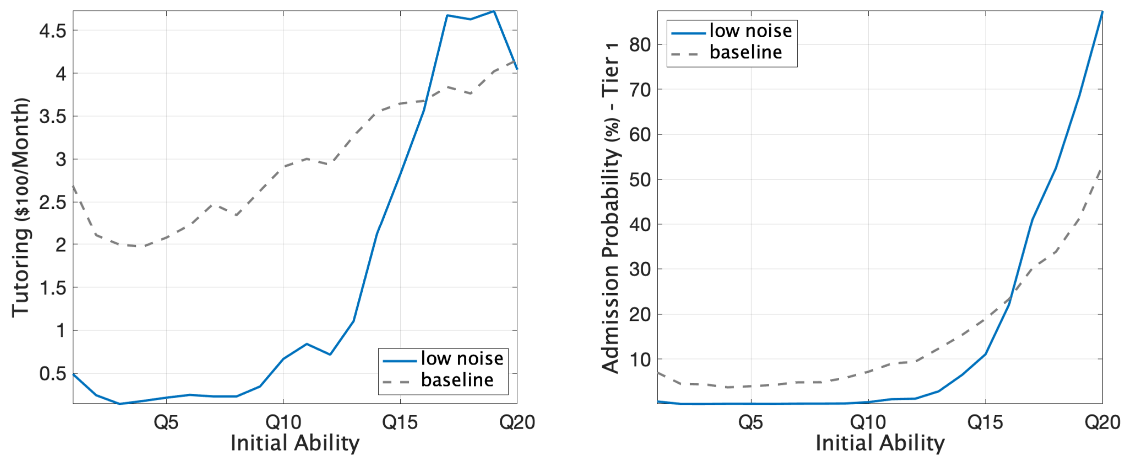
Tax and Subsidy

I now explore how tax and subsidy in tutoring market affect the tutoring incentives and equilibrium outcomes. It is theoretically ambiguous how households adjust their tutoring choices in response to price change. When price is high, human capital return to tutoring spending declines, but admission chances may improve because small increases in signal result in the student surpassing a larger fraction of competitors. Below I present counterfactuals where the price of tutoring increases (decreases) by 30% due to tax (subsidy).

Figures 5 presents the heterogeneous effects on tutoring expenditure and expected wage. The tutoring expenditure has been adjusted by price so that it measures the units of tutoring service purchased. One might think that because the top-income households purchase the most tutoring service, their expenditure should decrease the most with a proportional tax. However, it is the middle-income households' expenditure that respond the most in the experiment. Note that the middle-income households are more likely to be constrained, so that the tax is more "expensive" because tutoring service implies a greater loss in marginal utility. In addition, the unconstrained top-income households face less competition, so their marginal return of tutoring is higher.

The human capital accumulated gets lower, while admission probability as a function of income is almost unchanged, the latter perhaps due to the lack of responsiveness to competition incentives. Last, note that the distributional effects are across every ability level, so on average the admission probability as a function of ability is not changed. A subsidy would have the opposite effect.

Figure 6: Counterfactual - No Admission Noise



Reduce Admission Noise

In this experiment, I evaluate the importance of signal noise by assuming away the random matching shock ($\xi = 0$). The admission is determined by

$$s_i = \delta_1 CSAT_i + \delta_2 GPA_i + \delta_3 q_i + \delta_4 g_i + \delta_5 g_i^2 + \delta_6 y_i + \delta_7 y_i^2. \tag{3.10}$$

The declining noise-to-signal ratio provides households more certainty when making tutoring decision. Now the distributional effect is mostly across ability. Reduced noise lowers the probability that a low-ranking student is perceived as a high-ranking student. Therefore, the high-ability students can more effectively to purchase tutoring service to defend their positions in the ranking order. On the other hand, the low-ability students have less incentives of doing so due to the more "rigid" ranking order. These effects are reflected on Figure 6.

Across the income distribution, households decrease spending while the middle- and bottom-income decrease the most, the latter of which is perhaps due to the discouraged competition incentives as high-ability students can more easily stand out. As a result, it is the students of high ability and from high income family benefit in admission.

3.6 Conclusion

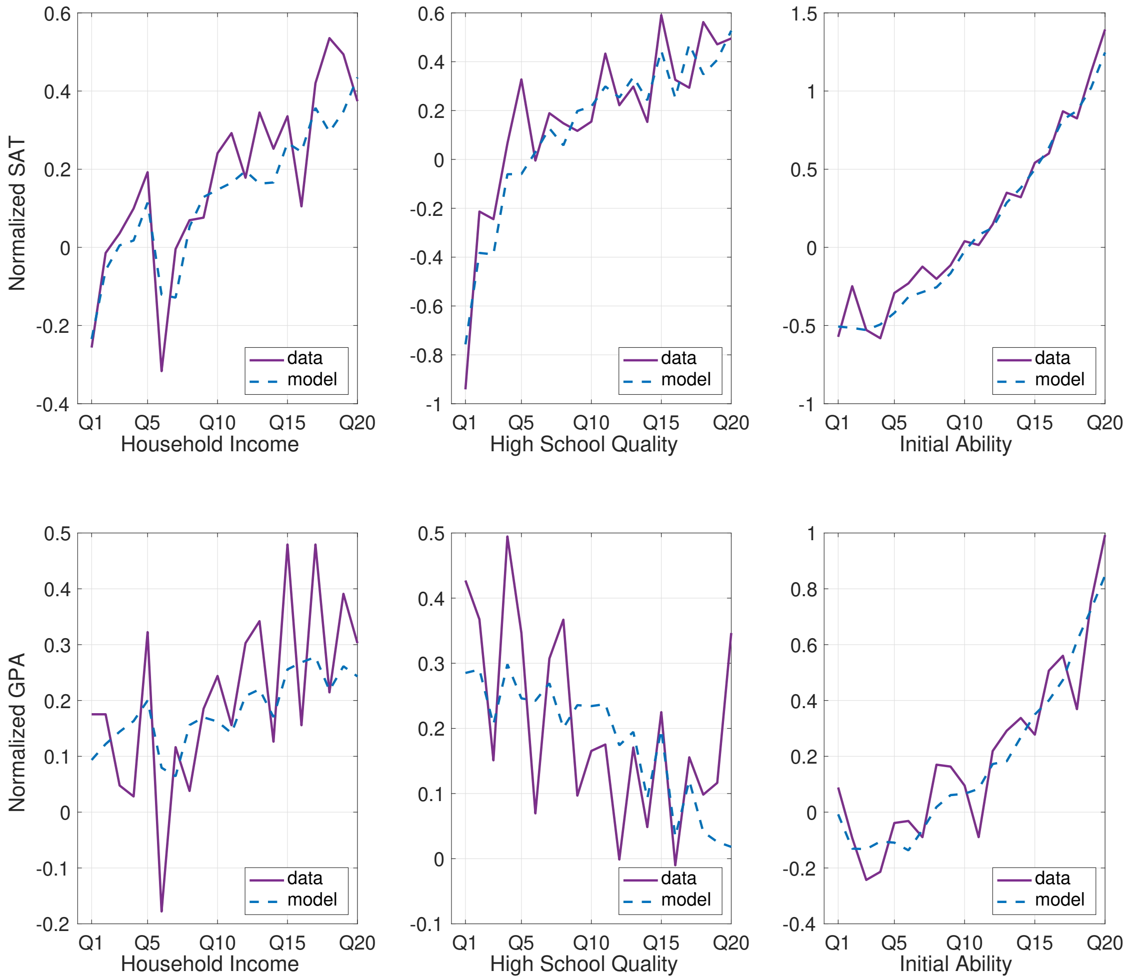
This paper develops a structural model to study pre-college educational investment (i.e., tutoring) in a college admission tournament. Methodologically, there are mainly two contributions. First, it allows educational investment to separately affect human capital accumulation and signals production. Second, it quantitatively studies the educational spending decision in a admission tournament. I find that tutoring produces genuine human capital, but also results in over-production driven by competition for prestigious colleges. The response of tutoring spending with respect to college admission is economically and quantitatively important.

As a result, the model provides policy implications. For example, conventional wisdom says a tax on educational investment should universally reduce investment and human capital by everyone. In this model, however, I find that the reduction of expenditure is most prominent among students from middle-income families. This is because the middle-income households are more likely to be constrained, so that the tax is more "expensive" because tutoring service implies a greater loss in marginal utility. Furthermore, I explore the impact of signal noise. Reduced noise incentivizes the tutoring expenditure of high-ability students, whereas discourages the low-ability students due to the more "rigid" ranking order. The admission chances of high-ability students get improved.

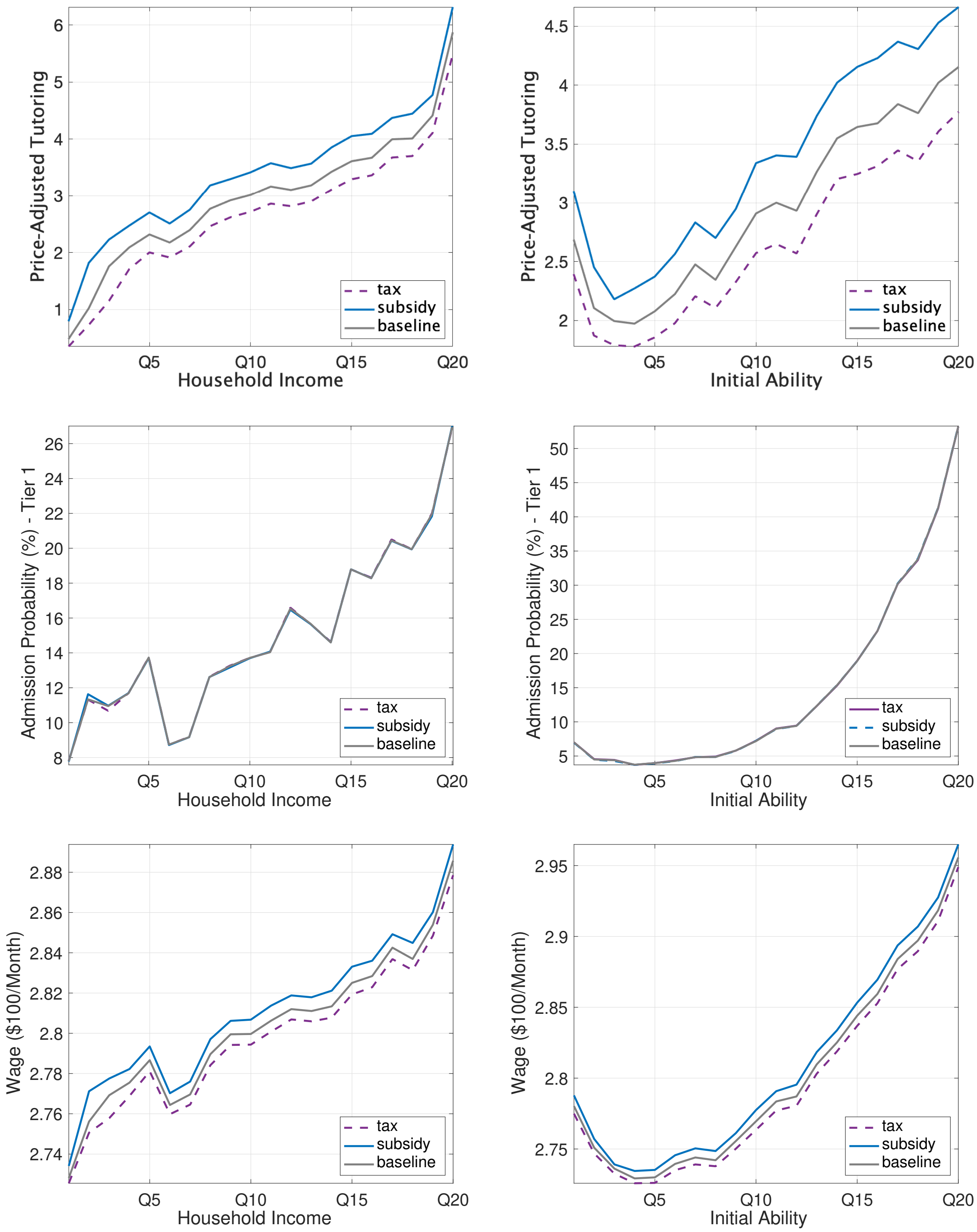
Future work can include further counterfactual analysis, for example, expansion of the selective universities or restriction on the quantity of tutoring service (e.g., limiting the hours of tutoring schools). Another possibility is to decompose the two channels of tutoring, especially their roles in explaining the contractual experiment results. Therefore, we can gain a better understanding of the value of incorporating the two channels in a unified study.

3.7 Appendix

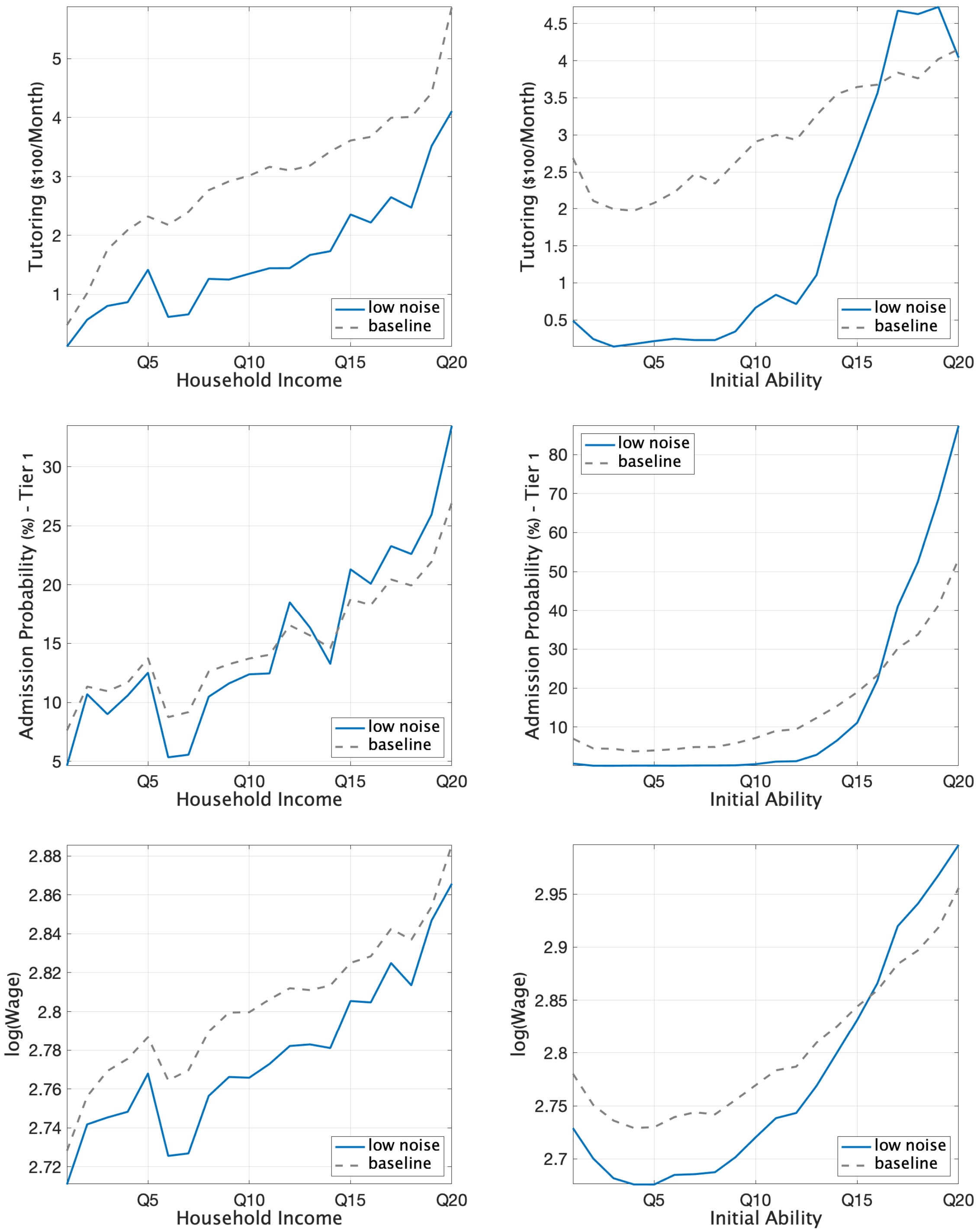
Appendix 1. Model Fit



Appendix 2. Counterfactual: Tax and Subsidy



Appendix 3. Counterfactual: Admission Noise



Acemoglu, D., U. Akcigit, H. Alp, N. Bloom, and W. Kerr (2018): "Innovation, reallocation, and growth," *American Economic Review*, 108(11), 3450–3491.

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll (2022): "Income and wealth distribution in macroeconomics: A continuous-time approach," *The review of economic studies*, 89(1), 45–86.

Agarwal, R., M. Ganco, and J. Raffiee (2021): "Immigrant entrepreneurship: The effect of early career immigration constraints on new venture formation," *Organization Science*, 33(4), 1372–1395.

Agarwal, R., M. Ganco, and R. H. Ziedonis (2009): "Reputations for toughness in patent enforcement: Implications for knowledge spillovers via inventor mobility," *Strategic Management Journal*, 30(13), 1349–1374.

Aghion, P., and P. Howitt (1990): "A model of growth through creative destruction," .

Akcigit, U., and S. T. Ates (2021): "Ten facts on declining business dynamism and lessons from endogenous growth theory," *American Economic Journal: Macroeconomics*, 13(1), 257–98.

Akcigit, U., S. Caicedo, E. Miguelez, S. Stantcheva, and V. Sterzi (2018): "Dancing with the stars: Innovation through interactions," Discussion paper, National Bureau of Economic Research.

Akcigit, U., and N. Goldschlag (2023): "Measuring the characteristics and employment dynamics of US inventors," Discussion paper, National Bureau of Economic Research.

Albino, V., A. C. Garavelli, and G. Schiuma (1998): "Knowledge transfer and inter-firm relationships in industrial districts: the role of the leader firm," *Technovation*, 19(1), 53–63.

Alchian, A. A., and H. Demsetz (1973): "The property right paradigm," *The Journal of Economic History*, 33(1), 16–27.

Almeida, P., and B. Kogut (1999): "Localization of knowledge and the mobility of engineers in regional networks," *Management science*, 45(7), 905–917.

Arcidiacono, P., P. Bayer, and A. Hizmo (2010): "Beyond signaling and human capital: Education and the revelation of ability," *American Economic Journal: Applied Economics*, 2(4), 76–104.

Arora, A., S. Belenzon, and L. Sheer (2021): "Knowledge spillovers and corporate investment in scientific research," *American Economic Review*, 111(3), 871–898.

Arrow, K. (1962): "Economic Welfare and the Allocation of Resources for Invention," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, pp. 609–626. Princeton University Press.

Arrow, K. J. (1972): *Economic welfare and the allocation of resources for invention*. Springer.

Arslan, H. A. (2018): "An empirical analysis of college admissions with endogenous entrance exam scores," .

Balasubramanian, N., J. W. Chang, M. Sakakibara, J. Sivadasan, and E. Starr (2022): "Locked in? The enforceability of covenants not to compete and the careers of high-tech workers," *Journal of Human Resources*, 57(S), S349–S396.

Balasubramanian, N., E. Starr, and S. Yamaguchi (2023): "Employment Restrictions on Resource Transferability and Value Appropriation from Employees," Available at SSRN: https://ssrn.com/abstract=3814403 or http://dx.doi.org/10.2139/ssrn.3814403.

Barles, G., and P. E. Souganidis (1991): "Convergence of approximation schemes for fully nonlinear second order equations," *Asymptotic analysis*, 4(3), 271–283.

Barney, J., M. Wright, and D. J. Ketchen Jr (2001): "The resource-based view of the firm: Ten years after 1991," *Journal of Management*, 27(6), 625–641.

Baslandze, S. (2022): "Entrepreneurship through employee mobility, innovation, and growth," .

Becker, G. S. (1962): "Investment in human capital: A theoretical analysis," *Journal of political economy*, 70(5, Part 2), 9–49.

Benhabib, J., J. Perla, and C. Tonetti (2021): "Reconciling models of diffusion and innovation: A theory of the productivity distribution and technology frontier," *Econometrica*, 89(5), 2261–2301.

Berry, H. (2006): "Leaders, laggards, and the pursuit of foreign knowledge," *Strategic Management Journal*, 27(2), 151–168.

Bilal, A., N. Engbom, S. Mongey, and G. L. Violante (2022): "Firm and worker dynamics in a frictional labor market," *Econometrica*, 90(4), 1425–1462.

Bloom, N., E. Brynjolfsson, L. Foster, R. Jarmin, M. Patnaik, I. Saporta-Eksten, and J. Van Reenen (2019): "What drives differences in management practices?," *American Economic Review*, 109(5), 1648–1683.

Bodoh-Creed, A. L., and B. R. Hickman (2018): "College assignment as a large contest," *Journal of Economic Theory*, 175, 88–126.

Boschma, R. A., and J. G. Lambooy (2002): "Knowledge, market structure, and economic coordination: dynamics of industrial districts," *Growth and Change*, 33(3), 291–311.

Braunerhjelm, P., D. Ding, and P. Thulin (2020): "Labour market mobility, knowledge diffusion and innovation," *European Economic Review*, 123, 103386.

Bray, M., and C. Lykins (2012): *Shadow education: Private supplementary tutoring and its implications for policy makers in Asia*, no. 9. Asian Development Bank.

Buera, F. J., and E. Oberfield (2020): "The global diffusion of ideas," *Econometrica*, 88(1), 83–114.

Burdett, K., and D. T. Mortensen (1998): "Wage differentials, employer size, and unemployment," *International Economic Review*, pp. 257–273.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): "Wage bargaining with on-the-job search: Theory and evidence," *Econometrica*, 74(2), 323–364.

Campbell, B. A., R. Coff, and D. Kryscynski (2012a): "Re-thinking Competitive Advantage from Human Capital: How the Concept of Firm-Specificity Has Led Strategy Theorists Astray," *Academy of Management Review*, 37(3), 376–395.

Campbell, B. A., M. Ganco, A. M. Franco, and R. Agarwal (2012b): "Whole Leaves, Where to, and Why Worry? Employee mobility, entrepreneurship and effects on source firm performance," *Strategic Management Journal*, 33(1), 65–87.

Carnahan, S., R. Agarwal, and B. A. Campbell (2012): "Heterogeneity in turnover: The effect of relative compensation dispersion of firms on the mobility and entrepreneurship of extreme performers," *Strategic Management Journal*, 33(12), 1411–1430.

Chiu, J., C. Meh, and R. Wright (2017): "Innovation and growth with financial, and other, frictions," *International Economic Review*, 58(1), 95–125.

Choi, H., and Á. Choi (2016): "Regulating private tutoring consumption in Korea: Lessons from another failure," *International Journal of Educational Development*, 49, 144–156.

Choi, J., and R. M. Cho (2016): "Evaluating the effects of governmental regulations on South Korean private cram schools," *Asia Pacific Journal of Education*, 36(4), 599–621.

Cinelli, C., J. Ferwerda, and C. Hazlett (2020): "Sensemakr: Sensitivity analysis tools for ols in r and stata," Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract$_i$d = 3588978.

Cinelli, C., and C. Hazlett (2020): "Making sense of sensitivity: Extending omitted variable bias," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(1), 39–67.

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

COLVIN, A., AND H. SHIERHOLZ (2019): "Noncompete agreements: Ubiquitous, harmful to wages and to competition, and part of a growing trend of employers requiring workers to sign away their rights," Economic Policy Institute. Bulgaria. Retrieved from https://policycommons.net/artifacts/1409606/noncompete-agreements/2023868/.

CONTI, R. (2014): "Do non-competition agreements lead firms to pursue risky R&D projects?," *Strategic Management Journal*, 35(8), 1230–1248.

COOLEY, T. F. (1995): *Frontiers of business cycle research*. Princeton University Press.

COWGILL, B., B. FREIBERG, AND E. STARR (2024): "Clause and Effect: A Field Experiment on Noncompete Clauses, Knowledge Flows, Job Mobility, and Wages," Working paper.

DANG, H.-A. (2007): "The determinants and impact of private tutoring classes in Vietnam," *Economics of Education Review*, 26(6), 683–698.

DREISIGMEYER, D., N. GOLDSCHLAG, M. KRYLOVA, W. OUYANG, E. PERLMAN, ET AL. (2018): "Building a Better Bridge: Improving Patent Assignee-Firm Links," Discussion paper, Center for Economic Studies, US Census Bureau.

EESLEY, C. E., AND Y.-S. LEE (2020): "Do university entrepreneurship programs promote entrepreneurship?," *Strategic Management Journal*, 42(4), 833–861.

FALLICK, B., C. A. FLEISCHMAN, AND J. B. REBITZER (2006): "Job-hopping in Silicon Valley: some evidence concerning the microfoundations of a high-technology cluster," *The review of Economics and Statistics*, 88(3), 472–481.

FERNANDEZ, R., AND J. GALI (1999): "To each according to...? Markets, tournaments, and the matching problem with borrowing constraints," *The Review of Economic Studies*, 66(4), 799–824.

FOSS, N. J., T. PEDERSEN, M. REINHOLT FOSGAARD, AND D. STEA (2015): "Why complementary HRM practices impact performance: The case of rewards, job design, and work climate in a knowledge-sharing context," *Human Resource Management*, 54(6), 955–976.

GANCO, M., R. H. ZIEDONIS, AND R. AGARWAL (2015): "More stars stay, but the brightest ones still leave: Job hopping in the shadow of patent enforcement," *Strategic Management Journal*, 36(5), 659–685.

GARMAISE, M. J. (2011): "Ties that truly bind: Noncompetition agreements, executive compensation, and firm investment," *The Journal of Law, Economics, and Organization*, 27(2), 376–425.

Gilson, R. J. (1999): "The legal infrastructure of high technology industrial districts: Silicon Valley, Route 128, and covenants not to compete," *NYUl Rev*, 74, 575.

Goldfarb, B., and A. A. King (2016): "Scientific apophenia in strategic management research: Significance tests & mistaken inference," *Strategic Management Journal*, 37(1), 167–176.

Gottfries, A., and G. Jarosch (2023): "Dynamic monopsony with large firms and an application to non-competes," .

Grau, N. (2018): "The impact of college admissions policies on the academic effort of high school students," *Economics of Education Review*.

Grossman, G. M., and E. Helpman (1991): "Trade, knowledge spillovers, and growth," *European economic review*, 35(2-3), 517–526.

Heckman, J. J., and B. Singer (2017): "Abducting Economics," *American Economic Review*, 107(5), 298–302.

Heggedal, T.-R., E. R. Moen, and E. Preugschat (2017): "Productivity spillovers through labor mobility in search equilibrium," *Journal of Economic Theory*, 169, 551–602.

Herkenhoff, K., J. Lise, G. Menzio, and G. M. Phillips (2018): "Production and learning in teams," Discussion paper, National Bureau of Economic Research.

Hickman, B. R. (2013): "Pre-college human capital investment and affirmative action: a structural policy analysis of US college admissions," *University of Chicago. Unpublished*.

Hopenhayn, H., and L. Shi (2020): "Knowledge creation and diffusion with limited appropriation," Discussion paper, working paper.

Hopkins, E. (2012): "Job market signaling of relative position, or Becker married to Spence," *Journal of the European Economic Association*, 10(2), 290–322.

Hosios, A. J. (1990): "On the efficiency of matching and related models of search and unemployment," *The Review of Economic Studies*, 57(2), 279–298.

Huselid, M. A. (1995): "The impact of human resource management practices on turnover, productivity, and corporate financial performance," *Academy of Management Journal*, 38(3), 635–672.

Ito, K. (1997): "Domestic competitive position and export strategy of Japanese manufacturing firms: 1971–1985," *Management Science*, 43(5), 610–622.

Ito, K., and V. Pucik (1993): "R&D spending, domestic competition, and export performance of Japanese manufacturing firms," *Strategic Management Journal*, 14(1), 61–75.

JAFFE, A. B., M. TRAJTENBERG, AND R. HENDERSON (1993): "Geographic localization of knowledge spillovers as evidenced by patent citations," *the Quarterly journal of Economics*, 108(3), 577–598.

JEFFERS, J. (2023): "The impact of restricting labor mobility on corporate investment and entrepreneurship," *Available at SSRN 3040393*.

JOHNSON, M. S., AND M. LIPSITZ (2022): "Why are low-wage workers signing noncompete agreements?," *Journal of Human Resources*, 57(3), 689–724.

JOHNSON, M. S., M. LIPSITZ, AND A. PEI (2023): "Innovation and the Enforceability of Noncompete Agreements," Discussion paper, National Bureau of Economic Research.

JONES, C. I. (2005): "Growth and ideas," in *Handbook of economic growth*, vol. 1, pp. 1063–1111. Elsevier.

KAISER, U., H. C. KONGSTED, AND T. RØNDE (2015): "Does the mobility of R&D labor increase innovation?," *Journal of Economic Behavior & Organization*, 110, 91–105.

KALTON, G., AND I. FLORES-CERVANTES (2003): "Weighting methods," *Journal of Official Statistics*, 19(2), 81–97.

KENNAN, J. (1986): "The economics of strikes," *Handbook of labor economics*, 2, 1091–1137.

KESAN, J. P., AND G. G. BALL (2006): "How are Patent Cases Resolved-An Empirical Examination of the Adjudication and Settlement of Patent Disputes," *Wash U. Law Review*, 84, 237–310.

KIM, J., AND G. MARSCHKE (2005): "Labor mobility of scientists, technological diffusion, and the firm's patenting decision," *RAND Journal of Economics*, 36(2), 298–317.

KIM, J. Y., AND H. K. STEENSMA (2017): "Employee mobility, spin-outs, and knowledge spill-in: How incumbent firms can learn from new ventures," *Strategic Management Journal*, 38(8), 1626–1645.

KINI, O., R. WILLIAMS, AND S. YIN (2021): "CEO noncompete agreements, job risk, and compensation," *The Review of Financial Studies*, 34(10), 4701–4744.

KLEPPER, S. (2001): "Employee startups in high-tech industries," *Industrial and Corporate Change*, 10(3), 639–674.

KÖNIG, M. D., J. LORENZ, AND F. ZILIBOTTI (2016): "Innovation vs. imitation and the evolution of productivity distributions," *Theoretical Economics*, 11(3), 1053–1102.

KORTUM, S. S. (1997): "Research, patenting, and technological change," *Econometrica: Journal of the Econometric Society*, pp. 1389–1419.

Krugman, P. (1991): "Increasing returns and economic geography," *Journal of political economy*, 99(3), 483–499.

Lavetti, K., K. Simon, and W. D. White (2020): "The impacts of restricting mobility of skilled service workers evidence from physicians," *Journal of Human Resources*, 55(3), 1025–1067.

Lazear, E. P., and S. Rosen (1981): "Rank-order tournaments as optimum labor contracts," *Journal of political Economy*, 89(5), 841–864.

Lee, S. (2007): "The timing of signaling: To study in high school or in college?," *International Economic Review*, 48(3), 785–807.

Lentz, R. (2014): "Optimal employment contracts with hidden search," Discussion paper, National Bureau of Economic Research.

Lentz, R., and N. Roys (2015): "Training and search on the job," Discussion paper, National Bureau of Economic Research.

Lipsitz, M., and E. Starr (2022): "Low-wage workers and the enforceability of noncompete agreements," *Management Science*, 68(1), 143–170.

Lucas Jr, R. E. (2009): "Ideas and growth," *Economica*, 76(301), 1–19.

Lucas Jr, R. E., and B. Moll (2014): "Knowledge growth and the allocation of time," *Journal of Political Economy*, 122(1), 1–51.

Luttmer, E. G. (2007): "Selection, growth, and the size distribution of firms," *The Quarterly Journal of Economics*, 122(3), 1103–1144.

Marx, M. (2011): "The firm strikes back: non-compete agreements and the mobility of technical professionals," *American Sociological Review*, 76(5), 695–712.

Marx, M., J. Singh, and L. Fleming (2015): "Regional disadvantage? Employee non-compete agreements and brain drain," *Research Policy*, 44(2), 394–404.

Marx, M., D. Strumsky, and L. Fleming (2009): "Mobility, skills, and the Michigan non-compete experiment," *Management Science*, 55(6), 875–889.

Mascarenhas, B. (1986): "International strategies of non-dominant firms," *Journal of International Business Studies*, 17(1), 1–25.

McElheran, K. (2015): "Do market leaders lead in business process innovation? The case (s) of e-business adoption," *Management Science*, 61(6), 1197–1216.

MECCHERI, N. (2009): "A note on noncompetes, bargaining and training by firms," *Economics Letters*, 102(3), 198–200.

MINCER, J. (1974): "Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2.," .

MØEN, J. (2005): "Is mobility of technical personnel a source of R&D spillovers?," *Journal of labor economics*, 23(1), 81–114.

MYONG, S. (2016): "Need based aid from selective universities and the achievement gap between rich and poor," .

OSTROM, E., AND C. HESS (????): "Private and Common Property Rights," in *Encyclopedia of Law & Economics*.

PANICCIA, I. (1998): "One, a hundred, thousands of industrial districts. Organizational variety in local networks of small and medium-sized enterprises," *Organization Studies*, 19(4), 667–699.

PERLA, J., AND C. TONETTI (2014): "Equilibrium imitation and growth," *Journal of Political Economy*, 122(1), 52–76.

PETRONGOLO, B., AND C. A. PISSARIDES (2001): "Looking into the black box: A survey of the matching function," *Journal of Economic literature*, 39(2), 390–431.

PHILLIPS, D. J. (2005): "Organizational genealogies and the persistence of gender inequality: The case of Silicon Valley law firms," *Administrative Science Quarterly*, 50(3), 440–472.

PORTER, M. E. (1985): *Competitive advantage: Creating and sustaining competitive advantage*. Free Press, New York.

POSTEL-VINAY, F., AND J.-M. ROBIN (2002): "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 70(6), 2295–2350.

PRATO, M. (2022): "The Global Race for Talent: Brain Drain, Knowledge Transfer, and Growth," *Knowledge Transfer, and Growth (November 28, 2022)*.

PRESCOTT, J., AND E. STARR (2022): "Subjective beliefs about contract enforceability," Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract$_i$d = 3873638#.

RAFFIEE, J. (2017): "Employee mobility and interfirm relationship transfer: Evidence from the mobility and client attachments of United States federal lobbyists, 1998–2014," *Strategic Management Journal*, 38(10), 2019–2040.

Rosenkopf, L., and P. Almeida (2003): "Overcoming local search through alliances and mobility," *Management Science*, 49(6), 751–766.

Rothstein, D. S., and E. Starr (2021): "Mobility Restrictions, Bargaining, and Wages: Evidence from the National Longitudinal Survey of Youth 1997," Available at: https://ssrn.com/abstract=3974897 or http://dx.doi.org/10.2139/ssrn.3974897.

Rubin, P. H., and P. Shedd (1981): "Human capital and covenants not to compete," *The Journal of Legal Studies*, 10(1), 93–110.

Ryu, D., and C. Kang (2013): "Do Private Tutoring Expenditures Raise Academic Performance? Evidence from Middle School Students in South Korea," *Asian Economic Journal*, 27(1), 59–83.

Samila, S., and O. Sorenson (2011): "Noncompete covenants: Incentives to innovate or impediments to growth," *Management Science*, 57(3), 425–438.

Schneider, B., M. G. Ehrhart, and W. H. Macey (2011): "Perspectives on organizational climate and culture," in *APA handbook of industrial and organizational psychology*, ed. by S. Zedeck, vol. 1, pp. 373–414. American Psychological Association.

Seo, E., and D. Somaya (2021): "Living It Up at the Hotel California: Employee Mobility Barriers and Collaborativeness in Firms' Innovation," *Organization Science*, 33(2), 766–784.

Sevcenko, V., and S. Ethiraj (2018): "How do firms appropriate value from employees with transferable skills? A study of the appropriation puzzle in actively managed mutual funds," *Organization Science*, 29(5), 775–795.

Shi, L. (2023): "Optimal regulation of noncompete contracts," *Econometrica*, 91(2), 425–463.

Shi, L., et al. (2021): "The macro impact of noncompete contracts," Discussion paper, Einaudi Institute for Economics and Finance (EIEF).

Singh, J., and A. Agrawal (2011): "Recruiting for ideas: How firms exploit the prior inventions of new hires," *Management science*, 57(1), 129–150.

Spence, M. (1973): "Job Market Signaling," *Quarterly Journal of Economics*, 87(3), 355–374.

Starr, E. (2019): "Consider this: Training, wages, and the enforceability of covenants not to compete," *ILR Review*, 72(4), 783–817.

Starr, E., N. Balasubramanian, and M. Sakakibara (2018): "Screening spinouts? How noncompete enforceability affects the creation, growth, and survival of new firms," *Management Science*, 64(2), 552–572.

STARR, E., M. GANCO, AND B. A. CAMPBELL (2018): “Strategic Human Capital Management in the Context of Cross-Industry and Within-Industry Mobility Frictions,” *Strategic Management Journal*, 39(8), 2226–2254.

STARR, E., J. PRESCOTT, AND N. D. BISHARA (2021): “Noncompete agreements in the US labor force,” *The Journal of Law and Economics*, 64(1), 53–84.

STOYANOV, A., AND N. ZUBANOV (2012): “Productivity spillovers across firms through worker mobility,” *American Economic Journal: Applied Economics*, 4(2), 168–198.

TZABBAR, D. (2009): “When does scientist recruitment affect technological repositioning?,” *Academy of Management Journal*, 52(5), 873–896.

TZABBAR, D., AND B. CIRILLO (2020): “Introduction: An Integrated Perspective of Employee Intra- and Inter-firm Mobility,” in *Employee Inter- and Intra-Firm Mobility*, ed. by D. Tzabbar, and B. Cirillo, vol. 41 of *Advances in Strategic Management*, pp. 1–11. Emerald Publishing Limited, Bingley.

YOUNGE, K. A., AND M. MARX (2016): “The value of employee retention: Evidence from a natural experiment,” *Journal of Economics & Management Strategy*, 25(3), 652–677.