Worker Mobility, Knowledge Diffusion, and Non-Compete Contracts

Jingnan Liu *†‡

December 18, 2023

Please click here for the latest draft.

Abstract

This paper 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.

*University of Wisconsin-Madison. Email: jingnan.liu@wisc.edu. Website: jingnanliu.com.

†I am grateful to my advisors, John Kennan, Rasmus Lentz, and Dean Corbae, for their invaluable guidance and support. This paper has benefited from discussions with Ufuk Akcigit, Manuel Amador, Carter Braxton, Santiago Caicedo, Julieta Caunedo, Jason Choi, Sharada Dharmasankar, Benjamin Friedrich, Martin Ganco, Chao He, Long Hong, Karam Jo, Matthew Johnson, Chad Jones, Seula Kim, Philipp Kircher, Rishabh Kirpalani, Pete Klenow, Thibaut Lamadon, Simone Lenzu, Jeremy Lise, Paolo Martellini, Ellen McGrattan, Espen R. Moen, Ezra Oberfield, Michael Peters, Tommaso Porzio, Xincheng Qiu, Mark Rempel, Shihan Shen, Kjetil Storesletten, Christopher Taber, Christopher Tonetti, Chenzi Xu, Stephen Yeaple, Anson Zhou, as well as numerous seminar participants.

‡Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2296. (CBDRB-FY24-P2296-R10970) This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation.

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.

[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."

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 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. When workers move, they bring knowledge from their origin to their 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 by allocating a greater share of surplus towards origin firms. 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 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 2 presents the theory, predictions, and efficiency properties. Section 3 describes the data and empirical evidence. The section 4 quantifies the impor-

[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.

tance of the worker-mobility channel of diffusion in aggregate growth and characterizes the optimal regulation of non-compete contracts.

2 Model

I study the worker-mobility channel and growth implications of knowledge diffusion. Section 2.1 proposes a framework that introduces on-the-job search to an endogenous growth model of innovation and diffusion. Section 2.2 characterizes the balanced growth path equilibrium and derives a growth decomposition. Section 2.3 solves the constrained planner's problem to characterize the inefficiency of decentralized equilibrium. Section 2.4 proposes a theory of non-compete contracts, and illustrates how the enforceability of non-compete contracts can alleviate the inefficiency.

2.1 Baseline Model

2.1.1 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.

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

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 $\kappa_{\text{r\&d}}(\mu_t)Z_t$.[6] The cost function $\kappa_{\text{r\&d}}(\cdot)$ is strictly increasing, continuously differentiable, and convex. Innovation intensity governs the speed of knowledge growth:

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

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.

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

[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.

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 condition:

$$\int_{\underline{Z}_t}^{\infty} \mathcal{V}(Z,t)\, d\mathcal{F}_e(Z,t) = \kappa_{\text{entry}}(t). \tag{1}$$

2.1.2 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]

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

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.

2.1.3 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.

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:

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

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:

$$\max_{\mathcal{C}} \quad E_{T,\{s_t\}_{t\geq t_0}} \left[\int_{t_0}^{T} e^{-rt} \Big(Z_t - \kappa_{\text{r\&d}}(\mu_t) Z_t - w_t \Big) dt + e^{-rT} \mathcal{V}(Z_T, T) \right]$$

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

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

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.

2.1.4 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:

$$r_t\,\mathcal{M}(Z,t) = \max_\mu \underbrace{Z}_{\text{Production}} - \underbrace{\kappa_{\text{r\&d}}(\mu)\,Z}_{\text{Innovation Cost}} + \underbrace{\mu Z\,\boldsymbol{\partial}_z\mathcal{M}(Z,t)}_{\text{Gains from Innovation}} + \boldsymbol{\partial}_t\mathcal{M}(Z,t) \tag{2}$$

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:

$$\boldsymbol{\partial}_\mu\kappa_{\text{r\&d}}(\mu) = \boldsymbol{\partial}_z\mathcal{M}(Z,t). \tag{3}$$

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

$$r_t \, \mathcal{V}(Z,t) = \underbrace{Z}_{\text{Production}} + \underbrace{\lambda(\theta_t)}_{\text{Meeting Rate}} \int \underbrace{\left[\mathcal{S}(Z,Z',t)\right]^+}_{\text{Surplus from Trade}} d\mathcal{F}_m(Z',t) + \boldsymbol{\partial}_t \mathcal{V}(Z,t). \tag{4}$$

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

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

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 promises 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.

2.1.5 Knowledge distributions

The distributions of knowledge are endogenous and determined in equilibrium. They shape firms' innovation and hiring decisions and evolve as innovation and worker mobility dictate. 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:

$$\begin{aligned}
\partial_t \mathcal{F}_m(Z,t) = \quad & -\underbrace{\mu(Z,t)Z\,\partial_Z \mathcal{F}_m(Z,t)}_{\text{Innovation}} && (5)\\
\textit{Workers move up}: \quad & -\lambda(\theta_t)\theta_t \int_{\underline{Z}_t}^{Z}\int_{Z}^{\infty} \mathbb{1}\left\{\mathcal{S}(Z',Z'',t)>0\right\} d\mathcal{F}_v(Z',t)d\mathcal{F}_m(Z'',t)\\
\textit{Workers move down}: \quad & +(1-p)\lambda(\theta_t)\theta_t \int_{Z}^{\infty}\int_{\underline{Z}_t}^{Z} \mathbb{1}\left\{\mathcal{S}(Z',Z'',t)>0\right\} d\mathcal{F}_v(Z',t)d\mathcal{F}_m(Z'',t)
\end{aligned}$$

The left-hand side describes the instantaneous change in CDF evaluated at knowledge Z at time t. The first term 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 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:

$$\begin{aligned}
\partial_t \mathcal{F}_v(Z,t) = \quad & \underbrace{\frac{N_e(t)}{\theta_t}\mathcal{F}_e(Z,t)}_{\text{Entry}} - \underbrace{g\underline{Z}_t\,\partial_Z \mathcal{F}_v(\underline{Z}_t,t)}_{\text{Endogenous Exit}} - \underbrace{\delta_v \mathcal{F}_v(\underline{Z}_t,t)}_{\text{Exogenous Exit}} && (6)\\
\textit{Workers move up}: \quad & +\lambda(\theta_t) \int_{Z}^{\infty}\int_{\underline{Z}_t}^{Z} \mathbb{1}\left\{\mathcal{S}(Z',Z'',t)>0\right\} d\mathcal{F}_m(Z'',t)d\mathcal{F}_v(Z',t)\\
\textit{Workers move down}: \quad & -\lambda(\theta_t) \int_{\underline{Z}_t}^{Z}\int_{Z}^{\infty} \mathbb{1}\left\{\mathcal{S}(Z',Z'',t)>0\right\} d\mathcal{F}_m(Z'',t)d\mathcal{F}_v(Z',t)
\end{aligned}$$

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

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

2.2 Balanced Growth Path

2.2.1 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 (2) (4), with $\mu(Z,t)$ the associated decision rules (3);
2. Given $\{\mu(Z,t), \theta_t\}$, $\{\mathcal{F}_m(Z,t), \mathcal{F}_v(Z,t)\}$ evolve according to KF equations (5) (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) holds;
4. r_t is consistent with the representative household's intertemporal 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:

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

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 (2) (4) simplify to

$$\begin{aligned}(r-g)\boldsymbol{M}(z) &= \max_{\mu} \underbrace{e^z}_{\text{Production}} - \underbrace{\kappa_{\text{r\&d}}(\mu)\, e^z}_{\text{Innovation Cost}} + \underbrace{(\mu - g)\boldsymbol{\partial}_z \boldsymbol{M}(z)}_{\text{Gains from Innovation}} \\ (r-g)\boldsymbol{V}(z) &= \underbrace{e^z}_{\text{Production}} + \underbrace{\lambda(\theta)}_{\text{Meeting Rate}} \int \underbrace{\left[\boldsymbol{S}(z,z')\right]^+}_{\text{Surplus from Trade}} d\boldsymbol{F}_m(z') - g\,\boldsymbol{\partial}_z \boldsymbol{V}(z),\end{aligned}$$

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

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

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:

$$\boldsymbol{\partial}_\mu \kappa_{\text{r\&d}}(\mu) = e^{-z}\boldsymbol{\partial}_z \boldsymbol{M}(z)$$

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 (5) (6) take the form

$$\begin{aligned}
0 &= -\boldsymbol{\partial}_z \left[(\boldsymbol{\mu}(z))\, \boldsymbol{f}_m(z)\right] + g\, \boldsymbol{\partial}_z \boldsymbol{f}_m(z) \\
&\quad + \lambda(\theta)\theta\, \boldsymbol{f}_v(z) \int_{\underline{z}}^{z} \mathbb{1}\left\{\boldsymbol{S}(z,z') > 0\right\} d\boldsymbol{F}_m(z') + (1-p)\lambda(\theta)\theta\, \boldsymbol{f}_v(z) \int_{z}^{\infty} \mathbb{1}\left\{\boldsymbol{S}(z,z') > 0\right\} d\boldsymbol{F}_m(z') \\
&\quad - (1-p)\lambda(\theta)\theta\, \boldsymbol{f}_m(z) \int_{\underline{z}}^{z} \mathbb{1}\left\{\boldsymbol{S}(z',z) > 0\right\} d\boldsymbol{F}_v(z') - \lambda(\theta)\theta\, \boldsymbol{f}_m(z) \int_{z}^{\infty} \mathbb{1}\left\{\boldsymbol{S}(z',z) > 0\right\} d\boldsymbol{F}_v(z') \\
0 &= g\boldsymbol{\partial}_z \boldsymbol{f}_v(z) + \frac{n_e}{\theta} \boldsymbol{f}_e(z) - \delta_v \boldsymbol{f}_v(z) \\
&\quad + \lambda(\theta) \boldsymbol{f}_m(z) \int_{\underline{z}}^{\infty} \mathbb{1}\left\{\boldsymbol{S}(z',z) > 0\right\} d\boldsymbol{F}_v(z') - \lambda(\theta) \boldsymbol{f}_v(z) \int_{\underline{z}}^{\infty} \mathbb{1}\left\{\boldsymbol{S}(z,z') > 0\right\} d\boldsymbol{F}_m(z')
\end{aligned}$$

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

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

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, $\theta \boldsymbol{f}_v(\underline{z})$, 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.

2.2.2 Growth decomposition

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

$$\boldsymbol{\mathcal{Z}} = \underbrace{\int_{\underline{z}}^{\infty} e^z\, d\boldsymbol{F}_m(z)}_{\text{Matches}} + \underbrace{\theta \int_{\underline{z}}^{\infty} e^z\, d\boldsymbol{F}_v(z)}_{\text{Vacant Firms}}.$$

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:

$$g = \frac{1}{\mathcal{Z}} \left\{ \begin{array}{l} \underbrace{\int_{\underline{z}}^{\infty} \boldsymbol{\mu}(z) e^{z} \, d\boldsymbol{F}_m(z)}_{\text{Innovation}} + \underbrace{\lambda(\theta)\theta p \int_{\underline{z}}^{\infty} \int_{z}^{\infty} \left(e^{z'} - e^{z}\right) \mathbb{1}\left\{\boldsymbol{S}(z,z') > 0\right\} d\boldsymbol{F}_m(z') d\boldsymbol{F}_v(z)}_{\text{Diffusion through worker mobility}} \\ \underbrace{+ n_e \int_{\underline{z}}^{\infty} e^{z} dF_e(z)}_{\text{Entry}} - \underbrace{\theta \, g f_v(\underline{z})}_{\text{Endogenous Exit}} \underbrace{- \delta_v \, \theta \int_{\underline{z}}^{\infty} e^{z} dF_v(z)}_{\text{Exogenous Exit}} \end{array} \right\}.$$

2.2.3 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 2.3 and quantitative analysis in section 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 $\mathcal{F}_m(Z'|Z' \geq Z, t)$.

To maintain tractability, I make the following functional form assumptions. The innovation cost function is $\kappa_{\text{r\&d}}(\mu) = \tilde{\kappa}_{\text{r\&d}} \, \mu^{\gamma}$. 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. *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*

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

2.2.4 Growth Decomposition.

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

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

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

2.3 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:

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

Social welfare is denoted by $\mathbf{\Omega}(f_m(\cdot,t),\phi_v(\cdot,t))$, 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:

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

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

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

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

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:

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

The social values are defined along the optimal trajectory of the densities, $\phi_v(\cdot,t)$ and $f_m(\cdot,t)$. Here $\frac{\delta}{\delta \phi_v(Z,t)}$ 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:

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

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

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

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

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

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:

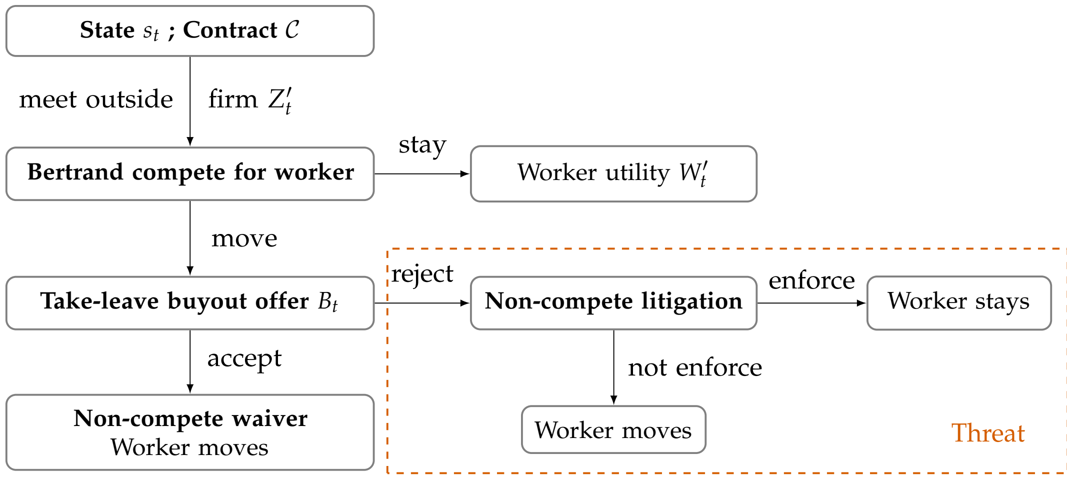


Figure 1: Game tree in search-and-matching stage

2.4 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:

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

Non-compete clause $\mathbb{1}_{\{\text{NCA}\}_{t_0}} \in \{0, 1\}$ 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 describes the dynamics of the contract.

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:

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

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

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

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':

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

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 3.1 describes the primary datasets employed and variable constructions. Section 3.2 presents the empirical strategy and findings.

3.1 Data

3.1.1 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.

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

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

[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.

observe earnings. Following the standard practice in the literature, I keep the primary job (job 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 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: 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

[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).

3.1.2 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.

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.

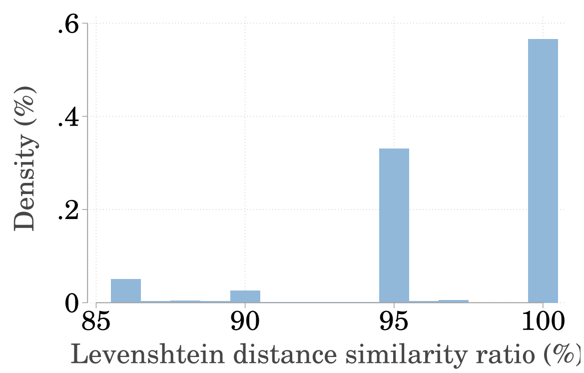
The compiled dataset contains the different firms an executive worked for, job positions at each firm, the dates those positions began and ended, and detailed compensation. The primary job is kept if an executive works for multiple firms in a fiscal year, often as a board member. A primary job is a managerial 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 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 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)).

3.2 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.

3.2.1 Worker mobility and knowledge diffusion

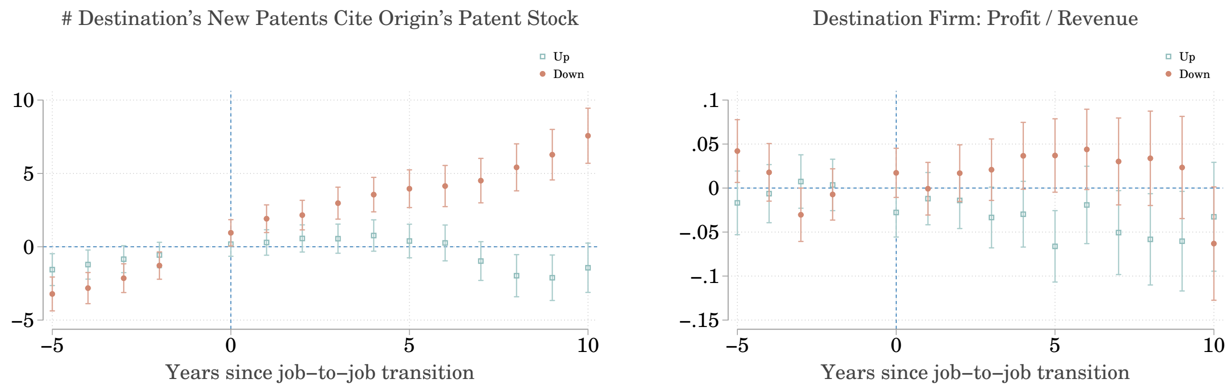
Event study. I use an event study framework to document how knowledge diffuses when workers change jobs. I define an event as job-to-job mobility: a worker moves from an origin firm (o) to a destination firm (d) in the same or the subsequent quarter. The event study takes the form:

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

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 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 3: Inter-Firm Patent Citations



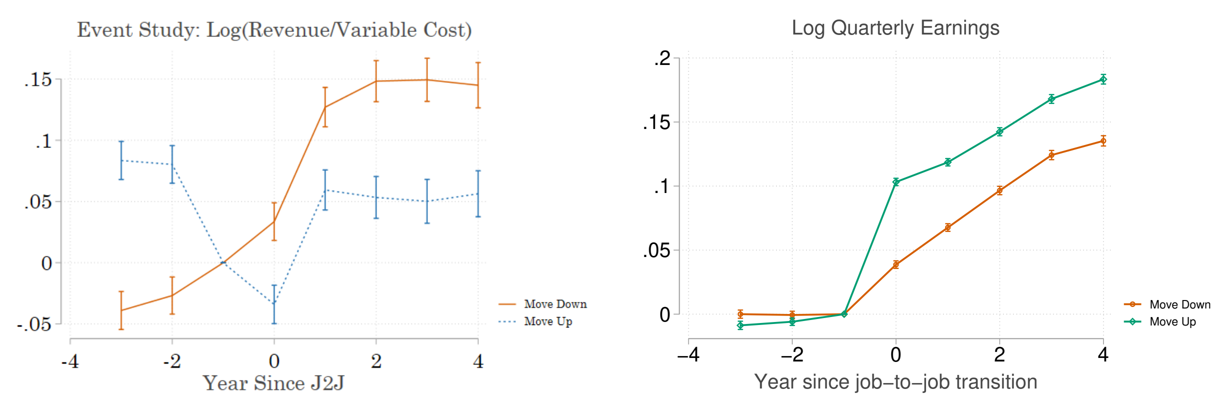
(a) Executive Mobility

(b) Profit Rate

Inventors diffuse knowledge and are compensated by knowledge diffusion.

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

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

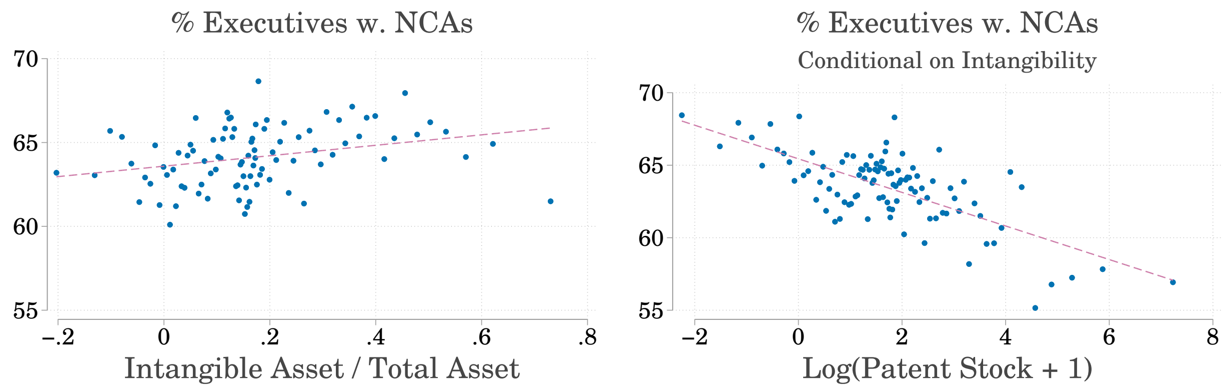


(a) Patent Citation

(b) Profit Rate

3.2.2 Innovation-diffusion tradeoff of non-compete contracts

Non-compete contracts protect intangibles, substitute for patents.



(a) Intangibility

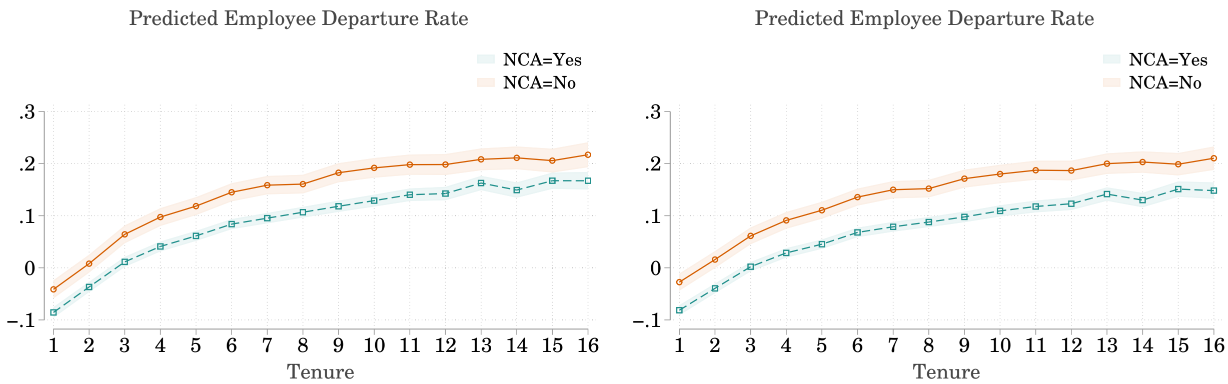
(b) Patent Stock

Non-compete contracts encourage R&D.

	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

4 Quantitative Analysis

Having shown that the model's key predictions align qualitatively with the data, we now proceed to quantify the model. Section 4.1 calibrates the model along a BGP equilibrium to match moments from the matched firm-inventor data at the U.S. Census Bureau. Section ... quantifies the importance of the worker-mobility channel of knowledge diffusion in aggregate growth. Section ... studies the optimal regulation of non-compete contracts.

4.1 Calibration

Parameterization. I specify the following functional forms. The innovation cost function is $\kappa_{\text{r\&d}}(\mu) = \tilde{\kappa}_{\text{r\&d}}\mu^{\gamma}$. 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

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

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.

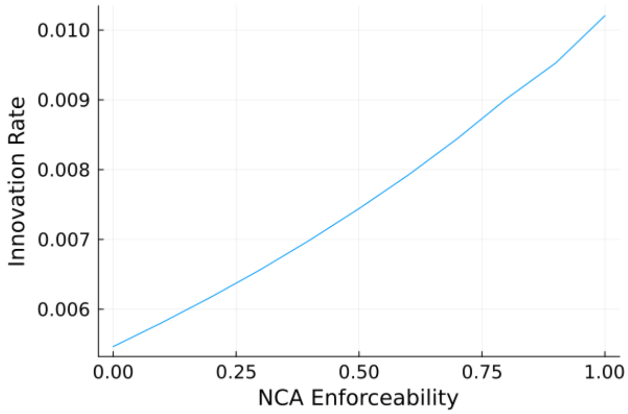
While the parameters are jointly calibrated, I provide a heuristic discussion of the most relevant moments for each parameter. The The learning probability is set to match the 3-year average change in $\log(TFPR)$ since hiring an inventor from more productive firms, relative to the $\log(TFPR)$ before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

Parameter	Definition	Value	Moment	Model	Data
p	Learning probability	0.11	$E[\Delta \log(TFPR_t)]$ 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: $\kappa_{\text{r\&d}}(\mu) = \tilde{\kappa}_{\text{r\&d}}\mu^{\gamma}$		
γ	Elasticity	2.00	Acemoglu et. al (2018)	–	–
$\tilde{\kappa}_{\text{r\&d}}$	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 $\log(TFPR_t)$ of entrant	0.88	0.88
δ_v	Exogenous exit rate	0.02	Size adjusted entry rate	1.87%	1.96%
κ_{scrap}	Scrap value (voluntary exit)	18.80	Smooth pasting condition	-	-

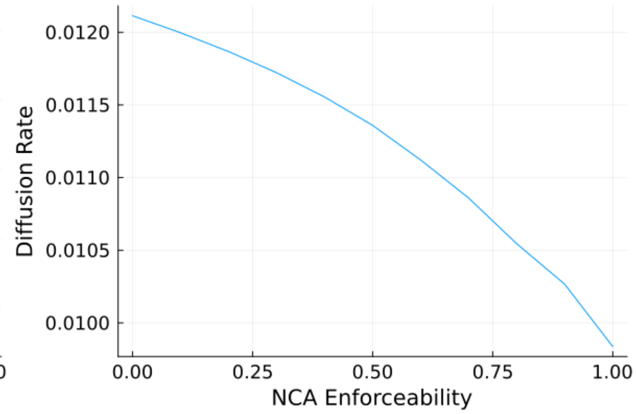
Table 2: Externally-calibrated parameters

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

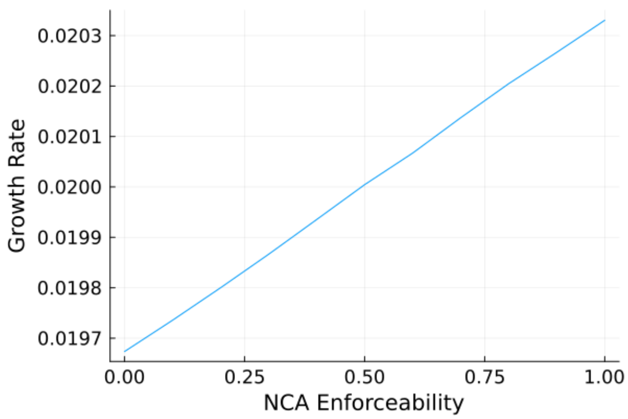
Comparative Statics. Higher NCA enforcement probability is associated with higher aggregate growth rate.



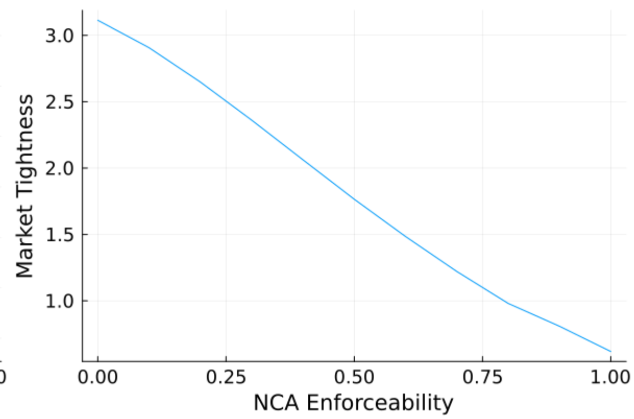
(a) Innovation ↑



(b) Diffusion ↓



(a) Growth Rate



(b) Market Tightness

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.

References

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.

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

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

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.

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.

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

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.

BRAUNERHJELM, P., D. DING, AND P. THULIN (2020): "Labour market mobility, knowledge diffusion and innovation," *European Economic Review*, 123, 103386.

BUERA, F. J., AND E. OBERFIELD (2020): "The global diffusion of ideas," *Econometrica*, 88(1), 83–114.

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

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.

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

HOPENHAYN, H., AND L. SHI (2020): "Knowledge creation and diffusion with limited appropriation," Discussion paper, working paper.

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.

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

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.

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.

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.

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.

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

SHI, L. (2023): “Optimal regulation of noncompete contracts,” *Econometrica*, 91(2), 425–463.

SINGH, J., AND A. AGRAWAL (2011): “Recruiting for ideas: How firms exploit the prior inventions of new hires,” *Management science*, 57(1), 129–150.

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

APPENDIX

A 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 (7) (**??**) 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}

A.1 SOLVING HJB EQUATIONS

A.1.1 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 (7) (**??**) can be jointly discretized as:

$$(r-g)\,R^{\tau+1} = \pi^{\tau} + A^{\tau} R^{\tau+1} \tag{7}$$

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

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

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 A.1.4.

A.1.2 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 (7) in terms of an HJB variational inequality (HJBVI):

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

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

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

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

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

$$\begin{aligned} X^T \left([(r-g)\, I - A^{\tau}]\, BX + [(r-g)\, I - A^{\tau}]\, B\underline{R} - \pi^{\tau}\right) &= 0 \\ [(r-g)\, I - A^{\tau}]\, BX + [(r-g)\, I - A^{\tau}]\, B\underline{R} - \pi^{\tau} &\geq 0 \\ X &\geq 0 \end{aligned} \tag{8}$$

A.1.3 Discretization with Upwind Finite-Differences

To compute the HJB equations (7) (**??**), 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

$$\begin{aligned} \textit{Forward difference}: \quad & \partial_{z,F} M_i := \frac{M_{i+1} - M_i}{\Delta_z} \\ \textit{Backward difference}: \quad & \partial_{z,B} M_i := \frac{M_i - M_{i-1}}{\Delta_z} \end{aligned}$$

The drift $\mu - g$ of state variable is derived from the dynamic decision rule (7) and approximated with

$$\begin{aligned} \textit{Forward}: \quad & \mu_{F,i} - g = (c')^{-1} \left(e^{-z_i}\, \partial_{z,F} M_i\right) - g \\ \textit{Backward}: \quad & \mu_{B,i} - g = (c')^{-1} \left(e^{-z_i}\, \partial_{z,B} M_i\right) - g \end{aligned}$$

Define the forward and backward Hamiltonians:

$$\begin{aligned} \textit{Forward}: \quad & H_{F,i} = -c(\mu_{F,i}) + e^{-z_i}\, (\mu_{F,i} - g)\, \partial_{z,F} M_i \\ \textit{Backward}: \quad & H_{B,i} = -c(\mu_{B,i}) + e^{-z_i}\, (\mu_{B,i} - g)\, \partial_{z,B} M_i \end{aligned}$$

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

$$\partial_z M_i \approx \quad \partial_{z,F} M_i\, \mathbb{1}_{F,i} + \partial_{z,B} M_i\, \mathbb{1}_{B,i} + e^{z_i} c'(g)\, \mathbb{1}_{C,i} \tag{9}$$

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

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

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

The optimal innovation decision is

$$\mu_i^\tau = (c')^{-1}\left(e^{-z_i}\,\partial_z M_i\right) \tag{10}$$

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 (**??**), the drifts $(-g)$ are negative. I calculate the approximate $\partial_z V_i$ with backward difference operator:

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

A.1.4 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).

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

Linear operator: innovation.

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

$$A^\tau_{1mm} = \begin{pmatrix} a_{B,1} + a_{C,1} & a_{F,1} & 0 & \dots & 0 \\ a_{B,2} & a_{C,2} & a_{F,2} & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & a_{B,n-1} & a_{C,n-1} & a_{F,n-1} \\ 0 & \dots & 0 & a_{B,n} & a_{C,n} + a_{F,n} \end{pmatrix}, \; A^\tau_{1vv} = \begin{pmatrix} \frac{-\sigma_v^2}{2\Delta_z^2} & \frac{\sigma_v^2}{2\Delta_z^2} & 0 & \dots & 0 \\ \frac{g}{\Delta_z} + \frac{\sigma_v^2}{2\Delta_z^2} & \frac{-g}{\Delta_z} - \frac{\sigma_v^2}{\Delta_z^2} & \frac{\sigma_v^2}{2\Delta_z^2} & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \frac{g}{\Delta_z} + \frac{\sigma_v^2}{2\Delta_z^2} & \frac{-g}{\Delta_z} - \frac{\sigma_v^2}{\Delta_z^2} & \frac{\sigma_v^2}{2\Delta_z^2} \\ 0 & \dots & 0 & \frac{g}{\Delta_z} + \frac{\sigma_v^2}{2\Delta_z^2} & \frac{-g}{\Delta_z} - \frac{\sigma_v^2}{2\Delta_z^2} \end{pmatrix}$$

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

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

Linear operator: worker mobility.

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

$$
A_{2mm}^\tau = \lambda(\theta)\theta \begin{pmatrix} -\sum_{j=2}^n \widetilde{d}_{1,j}\, f_{v,j} & \widetilde{d}_{1,2}\, f_{v,2} & \dots & \widetilde{d}_{1,n}\, f_{v,n} \\ (1-p)\widetilde{d}_{2,1}\, f_{v,1} & -(1-p)\sum_{j=1}^1 \widetilde{d}_{2,j}\, f_{v,j} - \sum_{j=3}^n \widetilde{d}_{2,j}\, f_{v,j} & \dots & \widetilde{d}_{2,n}\, f_{v,n} \\ \vdots & \vdots & \ddots & \vdots \\ (1-p)\widetilde{d}_{n,1}\, f_{v,1} & (1-p)\widetilde{d}_{n,2}\, f_{v,2} & \dots & -(1-p)\sum_{j=1}^{n-1} \widetilde{d}_{2,j}\, f_{v,j} \end{pmatrix},
$$

$$
A_{2mv}^\tau = \lambda(\theta)\theta \begin{pmatrix} (1-\tau)\sum_{j=1}^n \widetilde{d}_{1,j}\, f_{v,j} & -\widetilde{d}_{1,2}\, f_{v,2} & -\widetilde{d}_{1,3}\, f_{v,3} & \dots & -\widetilde{d}_{1,n}\, f_{v,n} \\ -\widetilde{d}_{2,1}\, f_{v,1} & (1-\tau)\sum_{j=1}^n \widetilde{d}_{2,j}\, f_{v,j} & -\widetilde{d}_{2,3}\, f_{v,3} & \dots & -\widetilde{d}_{2,n}\, f_{v,n} \\ -\widetilde{d}_{3,1}\, f_{v,1} & -\widetilde{d}_{3,2}\, f_{v,2} & (1-\tau)\sum_{j=1}^n \widetilde{d}_{3,j}\, f_{v,j} & \dots & -\widetilde{d}_{3,n}\, f_{v,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\widetilde{d}_{n,1}\, f_{v,1} & -\widetilde{d}_{n,2}\, f_{v,2} & -\widetilde{d}_{n,3}\, f_{v,3} & \dots & (1-\tau)\sum_{j=1}^n \widetilde{d}_{n,j}\, f_{v,j} \end{pmatrix},
$$

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

$$
A_{2vm}^\tau = \lambda(\theta) \begin{pmatrix} (1-p)\sum_{j=2}^n \widehat{d}_{1,j}\, f_{m,j} & -(1-p)\widehat{d}_{1,2}\, f_{m,2} & \dots & -(1-p)\widehat{d}_{1,n}\, f_{m,n} \\ -\widehat{d}_{2,1}\, f_{m,1} & \sum_{j=1}^1 \widehat{d}_{2,j}\, f_{m,j} + (1-p)\sum_{j=3}^n \widehat{d}_{2,j}\, f_{m,j} & \dots & -(1-p)\widehat{d}_{2,n}\, f_{m,n} \\ \vdots & \vdots & \ddots & \vdots \\ -\widehat{d}_{n,1}\, f_{m,1} & -\widehat{d}_{n,2}\, f_{m,2} & \dots & \sum_{j=1}^{n-1} \widehat{d}_{2,j}\, f_{m,j} \end{pmatrix},
$$

$$A^{\tau}_{2vv} = \lambda(\theta) \begin{pmatrix} -\sum_{j=1}^{n} \widehat{d}_{1,j} f_{m,j} & (1-\tau)\widehat{d}_{1,2} f_{m,2} & (1-\tau)\widehat{d}_{1,3} f_{m,3} & \dots & (1-\tau)\widehat{d}_{1,n} f_{m,n} \\ (1-\tau)\widehat{d}_{2,1} f_{m,1} & -\sum_{j=1}^{n} \widehat{d}_{2,j} f_{m,j} & (1-\tau)\widehat{d}_{2,3} f_{m,3} & \dots & (1-\tau)\widehat{d}_{2,n} f_{m,n} \\ (1-\tau)\widehat{d}_{3,1} f_{m,1} & (1-\tau)\widehat{d}_{3,2} f_{m,2} & -\sum_{j=1}^{n} \widetilde{d}_{3,j} f_{m,j} & \dots & (1-\tau)\widehat{d}_{3,n} f_{m,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (1-\tau)\widehat{d}_{n,1} f_{m,1} & (1-\tau)\widehat{d}_{n,2} f_{m,2} & (1-\tau)\widehat{d}_{n,3} f_{m,3} & \dots & -\sum_{j=1}^{n} \widehat{d}_{n,j} f_{m,j} \end{pmatrix},$$

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.

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

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

A.1.5 Summary of the algorithm.

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

1. Compute $\partial_z M_i$ using (9)
2. Compute μ^{τ} from (10)
3. Find $R^{\tau+1}$ from (8)
4. If $R^{\tau+1}$ is close enough to R^{τ}: stop. Otherwise, go to step 1.

A.2 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,

$$\begin{aligned} D^\tau f^{\tau+1} &= 0 \\ \widehat{f}^{\tau+1} &= x f^{\tau+1} + (1-x) f^\tau \\ D^{\tau+1} &= D(\widehat{f}^{\tau+1}) \end{aligned} \tag{11}$$

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).

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

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

A.3 SOLVING MARKET TIGHTNESS

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

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