Pre-Match Investment in Knowledge-Based Hierarchies*

Martín Grunwald†

October 30, 2026

[Click here for the latest version]

Abstract

This paper studies how individuals' education choices interact with the organization of firms. Agents that differ in ability decide whether to undertake a costly, risky investment in education before sorting into occupations as workers or managers. Managers match with workers to form firms, helping them produce. Crucially, education choices depend on expected earnings, which, in turn, are shaped by aggregate education decisions through the occupational sorting and matching they induce. We show the equilibrium is unique and derive comparative statics. Although the effects are a priori ambiguous, we provide a condition for signing the effects of lower costs of managerial help, and show when this can lead to less investment. Lower education costs, however, unambiguously increase investment, raising wages, lowering rents, and reallocating employment towards larger firms. On the other hand, improvements in the distribution of education outcomes can lead to a decrease in equilibrium investment. Even when fewer agents invest, wages rise for a positive mass of workers.

*I am very grateful to my advisor Hector Chade and my committee Mark Whitmeyer and Galina Vereshchagina. Additional thanks are extended to Ludovic Renou, Christopher Almacen, and seminar participants at Arizona State University, Stony Brook International Conference on Game Theory, Jornadas Anuales de Economía del BCU, and Universidad ORT for very helpful comments and suggestions. This project has benefited from large language models. All errors are my own.

†Arizona State University, Department of Economics (email: mgrunwald@asu.edu).

1 Introduction

In the U.S. firms with at least ten employees account for around 90% of employment and 93% of payroll (United States Census Bureau, 2025). On the one hand, how firms organize, including how they use the knowledge and time of their employees, shapes the earnings of these workers and the distribution thereof. On the other hand, the knowledge these employees hold depends on their education choices, which are typically made before knowing the firm or position in which they will use said expertise. Because education precedes employment, individuals must weigh education costs against the expected increase in earnings, which depends on the organization within firms. Collectively, these education decisions determine the knowledge available in the economy, thus affecting the way production is organized. The question then is: *How does the organization of knowledge inside firms interact with individuals' incentives to invest in education?*

This paper studies that interaction within the knowledge-based hierarchies framework of Garicano and Rossi-Hansberg (2004), in which production involves solving problems that require knowledge. Agents sort into occupations according to their knowledge: *workers* attempt to solve problems and ask *managers* for help when they cannot. We introduce a prior education stage in which agents choose whether to invest in their education at a cost. As in Chade and Lindenlaub (2022), investing improves the distribution from which agents draw their knowledge, rather than selecting a specific knowledge level. We establish existence and uniqueness of the equilibrium. Our main contribution is to derive analytical comparative statics for changes in the technologies through which knowledge is acquired and communicated. In particular, lower education costs and an improvement in the distribution of its outcomes can have opposing effects on the share of agents that undertake the investment. While the first one unambiguously increases the share of agents that invest, the second can lead to a decrease in that share due to the adjustments in occupational sorting and earnings.

The model has an *investment* stage followed by an *organizational* stage. In the investment stage, risk-neutral agents of heterogeneous ability choose whether to invest in their education. Investing improves the distribution from which knowledge is drawn in the likelihood-ratio sense: higher knowledge realizations are relatively more likely among investors. Higher ability only reduces education costs; conditional on the investment decision, all individuals face the same distribution of knowledge. An agent will invest if their expected improvement in earnings outweighs their cost of investing. The distribution of knowledge in the economy is then determined by the investment decisions of all agents. Before the next stage, each agent draws their knowledge from the distribution

determined by their investment choice. Stochastic knowledge acquisition separates two channels of improvements in *knowledge-acquisition technologies*: those that affect the costs of acquiring education, and those that improve the outcomes of the investment which in this paper implies a "better" knowledge distribution conditional on investing.

The organizational stage incorporates the framework of Garicano and Rossi-Hansberg (2004), in which agents choose their occupation and match to form teams. Production involves drawing problems of uncertain difficulty, which takes one unit of time. Agents can solve that problem if their knowledge is higher than the problem's difficulty. Importantly, knowledge can be communicated: if an agent lacks the knowledge to solve a problem, another agent can help them. It costs the latter $h \in (0, 1)$ units of time to help a single individual. We refer to h as the communication cost. Working in production precludes helping others, so agents specialize in production or helping. Individuals can choose to produce independently —also referred to as being *self-employed*— or for firms/teams, as well as who to team with. In this paper, firms are restricted to have one layer of production workers and up to a single layer of helpers. In equilibrium, firms take a hierarchical structure: one agent specialized in helping —called a *manager*— and a team of *workers* focused on drawing problems. The equilibrium determines occupations, matching, and team sizes.

To illustrate the model, consider the example of individuals choosing whether to study accounting before becoming a tax preparer. Clients bring problems of varying difficulty. Every preparer can work on their own, solving only the problems for which they have enough knowledge. Nevertheless, it can be beneficial for some to form teams/firms: a knowledgeable preparer can specialize in the difficult cases that less knowledgeable teammates cannot solve. Before knowing what position they will occupy and how much they will earn, each individual must decide whether to invest in a college degree. Going to college can, in expectation, deliver higher knowledge than not going, but its stochastic nature does not guarantee a particular role. Preparers with the same education can enter different occupations if their knowledge realizations differ. Agents compare the expected improvement in earnings, which depends on the knowledge other preparers acquire, with how costly the investment is. In turn, the aggregate decisions of all agents will determine how common each knowledge level is, and hence, how the market rewards it.

We show that there is a unique equilibrium. An ability threshold determines the investment decisions: individuals with ability above it invest, and those below it do not. Conditional on the knowledge distribution resulting from investment choices, occupational sorting and matching follow Garicano and Rossi-Hansberg (2004) and Carmona, Laohakunakorn, and Tsujiyama (2026). There are two cutoffs such that agents with low

knowledge become workers in firms, those with intermediate knowledge become self-employed (a region that might be empty in equilibrium), and those with high knowledge become managers. Since more knowledgeable workers require less help, they enable higher-knowledge managers to supervise larger teams. Thus matching is positively assortative: more knowledgeable workers are teamed with more knowledgeable managers. Our investment stage determines that distribution through individuals' education decisions.

Uniqueness of the investment threshold follows from the marginal benefit of investing being increasing in this threshold (i.e., in the mass of non-investors). When fewer agents invest, "low knowledge" is relatively more abundant in the economy. Since workers concentrate in the lower end of the knowledge spectrum, this drives more competition among them and therefore decreases wages. On the other hand, higher knowledge is relatively scarcer, increasing managerial rents. Summarizing, earnings for lower knowledge levels decrease while those for higher knowledge levels increase. Consequently, the benefit of stochastically drawing higher knowledge —of improving the chances of becoming a manager— increases. A higher investment threshold therefore raises the expected earnings gain from education while lowering the education cost of the agent at the threshold. The gain and cost can coincide at most once.

We then study the effects of improvements in *knowledge-acquisition* technologies, the costs of investments and the distribution of investment outcomes; and in *knowledge-communication* technologies. Communication costs affect the number of workers each manager can supervise and the relative returns to each occupation, thus altering the returns to education. To understand the forces on the equilibrium, we first study the organizational response to a decrease in h holding the knowledge distribution fixed, before allowing education decisions to adjust. The effects are, in principle, ambiguous: Fuchs, Garicano, and Rayo (2015) show that the mass of workers increases, which pushes towards more managers, but it is also cheaper to help each worker, which pushes towards less managers. We provide a condition for signing the effects of lower costs of managerial help. Specifically, if the most knowledgeable agents can solve every problem and the distribution of knowledge has an increasing density, then a decrease in h lowers the managerial cutoff (i.e., the mass of managers increases).

To understand why, suppose instead that the managerial cutoff remained unchanged. Lower communication costs would allow the marginal manager to supervise larger teams, so the wage paid to the least knowledgeable workers (those who match with the the lowest-knowledge manager) must rise to keep that marginal manager indifferent between management and self-employment. Since teams can be larger some previously self-

employed agents have to become workers. Since these entrants are more knowledgeable than incumbent workers, they would push incumbents toward less knowledgeable managers. This would reduce output per worker and put downward pressure on wages. An increasing density limits this productivity loss: a given mass of managers occupies a narrower knowledge interval at higher knowledge levels. The resulting downward pressure on wages would be too weak resulting in more workers wanting to join beyond what the managers could supervise. Restoring equilibrium therefore requires some self-employed agents to enter management.

Under this density condition and holding the knowledge distribution fixed, lower helping costs have opposing effects on the incentive to invest. On the one hand, they raise wages but compress wage differences across worker knowledge levels, weakening the return to education. On the other hand, larger teams allow managers to apply their knowledge to more problems, increasing rents and widening rent differences across managerial knowledge levels, thereby strengthening the return to education. We establish a separate sufficient condition under which the first effect dominates and a decrease in h lowers equilibrium investment. This condition has two requirements. First, not investing must yield high average knowledge. Second, the poor outcomes that education helps avoid must correspond to sufficiently low knowledge realizations. High average knowledge limits the demand for managerial help —because individuals can already solve many problems themselves— keeping the economy's average gain from cheaper help small. In contrast, workers with very low knowledge need help with almost every problem. Cheaper help therefore saves considerable managerial time when employing these workers, raising their wages. Education makes these poor outcomes less likely, so people who do not invest are more exposed to those wage increases. Under the condition, their expected earnings rise more than investors', reducing the incentive to invest.

While lower communication costs can discourage investment by changing the returns to education, reducing education costs directly makes investment more attractive. The effort cost of education may fall because, for example, information is easier to access. Intuitively, a decrease in costs of investment leads more agents to invest. This makes high-knowledge relatively more abundant. Consequently, competition among managers raises wages and decreases managerial rents. The ratio of highest to lowest earnings falls both in the economy as a whole and among managers. While rents fall, the average firm size increases. In particular, comparing team sizes at a given level of managerial knowledge, initially large teams expand and initially small firms contract. Some individuals who initially managed a team go into self-employment. A "better" distribution of ability (a first-order stochastically dominating distribution) produces the same consequences.

The other side of the coin of "knowledge-acquisition technologies" is an improvement in the consequences of education. In particular, we study an improvement in the likelihood-ratio sense of the distribution of knowledge conditional on investing. This improvement has, a priori, ambiguous effects since different forces push in opposite directions. On the one hand, for a fixed earnings schedule, an improvement in the distribution of knowledge from education means stochastically better draws, which incentivizes investment. On the other hand, for a fixed share of investors, better educational outcomes, by increasing the "supply" of higher knowledge, yield higher wages and lower rents. Because education makes higher knowledge more likely, such a change in relative earnings discourages investment.

In the case of an initial equilibrium without self-employment, we provide a sufficient condition for a marginal improvement in educational outcomes to decrease the mass of investors. In particular, this condition requires that education substantially increase the chance of becoming a manager and that few individuals have knowledge close to the occupational cutoff. Holding investment and the occupational cutoff fixed, better educational outcomes increase the number of managers and reduce the help workers need. This creates excess managerial capacity. Restoring the balance between the volume of questions asked by workers and the time managers use to help them requires the cutoff to increase. That is, some individuals who would otherwise manage have to become workers. If few individuals have knowledge near the cutoff, a small increase in it does not move enough agents into production. Hence, the individuals further above the original cutoff must become workers. Inducing them to become workers requires a larger improvement in worker pay relative to managerial earnings. Since investing makes managerial positions more likely, that adjustment reduces the expected improvement in earnings from education. Under our sufficient condition, this reduction outweighs the direct benefit of better learning outcomes. Thus, fewer people invest even when education becomes more effective.

The decline in participation counteracts the initial improvement in wages and decrease in rents (see the intuition for uniqueness). Comparing earnings at the same realized knowledge across equilibria, some workers can earn less and some managers can earn more: the wage decrease from fewer investors outweighs the increase from a better knowledge distribution conditional on investment, and the opposite is true for rents. Nevertheless, a positive mass of workers earn more and a positive mass of managers earn less than in the original equilibrium.

1.1 Related literature

This paper contributes to two strands of the literature: knowledge-based hierarchies, initiated by Garicano (2000) (see Garicano and Rossi-Hansberg (2015) for a survey), and investment before matching.

Within the first strand, we build on the competitive equilibrium formulation of Garicano and Rossi-Hansberg (2004), where they study how agents that differ in their exogenously given knowledge choose occupation and organize in firms.[1] Carmona, Laohakunakorn, and Tsujiyama (2026) develop a general framework for knowledge economies with occupational choice and many-to-one matching. They establish the equivalence of stability, competitive equilibrium, and surplus maximization. These papers take individuals' knowledge as an attribute of each agent, exogenously given. We contribute by accounting for the equilibrium effects between knowledge acquisition and organization. Additionally, we complement the result in Fuchs, Garicano, and Rayo (2015) that lower communication costs expand the worker region. When the most knowledgeable agents can solve every problem, we provide a sufficient condition for cheaper communication to lower the managerial cutoff when knowledge is fixed and self-employment remains active: the density of knowledge is increasing.

Other papers treat knowledge acquisition as endogenous. Chief among them is Garicano and Rossi-Hansberg (2006). They jointly study knowledge acquisition, occupational choice, organization, and earnings. They characterize the equilibrium and provide numerical comparative statics for reductions in costs of knowledge acquisition and in costs of communication.[2] Knowledge acquisition is deterministic: learning maps exactly to knowledge levels, which in turn map to earnings. Caicedo, Lucas Jr, and Rossi-Hansberg (2019) treat learning as a stochastic process of meeting other agents and show how it generates career progression and experience-earnings profiles. While stochastic in nature,

[1] Applications and extensions of the framework study how international teams affect matching and inequality (Antràs, Garicano, and Rossi-Hansberg, 2006); how informational frictions shape the organization and contracting of expert advice (Fuchs and Garicano, 2010; Fuchs, Garicano, and Rayo, 2015); how the introduction of artificial intelligence affects occupational choices and earnings (Ide and Talamas, 2025); how size-dependent payroll taxes distort occupational allocation and worker–manager sorting (López and Torres, 2020); and how progressive taxation changes managerial time, the organization of production, and the wage distribution (Kapička and Slavík, 2022).

[2] Knowledge acquisition is also endogenous in Garicano and Hubbard (2005), where they show how on-the-job learning strengthens skill complementarities; in Caliendo and Rossi-Hansberg (2012), where they study how trade induces reorganization and changes knowledge choices (Caliendo, Monte, and Rossi-Hansberg (2015) and Caliendo, Mion, et al. (2020) study empirically the effects on firm growth and productivity, respectively); in Garicano and Rossi-Hansberg (2012) they connect organizational development to the incentives to exploit existing technologies and invent new ones; Wu (2015) shows that wider spans of control are complementary to high-value products; and Pieri and Vatiero (2025) predict that thicker knowledge markets support wider spans of control and fewer layers.

learning is not a choice in their setup, and it involves no costs. In this regard, the closest paper is that of Garicano and Rossi-Hansberg (2006). In our framework, individuals choose whether to improve their distribution of acquired knowledge before occupations and earnings are determined. Education costs and conditional learning outcomes enter separately, allowing us to compare their equilibrium effects. Nevertheless, we restrict attention to firms having one managerial layer, while Garicano and Rossi-Hansberg (2006) allow the number of layers to be endogenous.

A second strand of the literature studies investment before matching. Investment incentives depend on how matching markets allocate surplus. The literature on investment before matching examines how coordination and pricing arrangements affect private returns and the efficiency of investment (Cole, Mailath, and Postlewaite, 2001; Mailath, Postlewaite, and Samuelson, 2013; Nöldeke and Samuelson, 2015). When investment is risky, competition for subsequent matches also shapes how much individuals invest and which risks they are willing to take (Bhaskar and Hopkins, 2016; Zhang, 2020). Our closest connection to this literature is Chade and Lindenlaub (2022). They study risky skill acquisition by risk-averse workers before they match with heterogeneous firms. They characterize how risk, heterogeneity, technology, and risk-sharing opportunities affect investment and earnings. In particular, they also analyze improvements in educational outcomes and the opposing direct and equilibrium effects on investment. As opposed to their baseline worker-firm model, our framework does not have two pre-specified sides of a matching market. We have endogenous occupational roles and many-to-one matching: workers and managers emerge from the same population after knowledge is realized, and each manager is matched to many workers. Education therefore changes the knowledge available in these occupations and the selection of agents into each side. Even in the absence of risk aversion, our setting captures rich equilibrium behavior, and we provide a sufficient condition under which better educational outcomes reduce participation.

The rest of the paper is structured as follows. Section 2 introduces the model. Section 3 studies the equilibrium and establishes existence and uniqueness. We explore the economic implications of changes in communication costs, education costs, and educational outcomes in Section 4, and conclude in Section 5.

2 Model

We endogenize the distribution of knowledge in Garicano and Rossi-Hansberg (2004) by introducing a costly and risky pre-match investment decision in the style of Chade and Lindenlaub (2022). The model has two stages. In the first stage, agents with heteroge-

neous ability choose whether to invest in education before their knowledge is realized and hence before knowing their future occupation. Ability affects the cost of education, while the investment decision determines the distribution of knowledge. In the second stage, knowledge becomes publicly observable, agents choose occupations (worker, self-employed, or manager), and workers and managers match to form firms. Investment choices determine the distribution of knowledge entering production, and the resulting earnings schedule determines the incentive to invest.

Education and Knowledge. There is a unit mass of risk-neutral agents in the economy, indexed by their ability $\theta \in \Theta = [0,1]$. The cumulative distribution function (cdf) of ability is $\Gamma(\theta)$. Γ is continuously differentiable and admits a strictly positive density γ on $[0,1]$. Each agent has one unit of time, which can be used for production or for helping other agents.

In the *investment stage*, each agent has a binary investment choice $i \in \{0,1\}$, where $i = 1$ denotes the agent investing in education. Investment determines the distribution from which agents draw their knowledge, z, which will be publicly observable. If an agent decides to invest ($i = 1$), they draw from a cdf $G_1(z)$; otherwise, they draw from $G_0(z)$. Both distributions have support $[0, \bar{z}]$. Investment makes higher knowledge realizations relatively more likely, but investors and non-investors can attain the same knowledge levels. Both admit strictly positive, continuously differentiable densities g_1 and g_0, respectively. In addition, the distributions are ordered in the (strict) likelihood-ratio order: $G_1 \succ_{LR} G_0$. This means that the ratio $\ell(z) := \frac{g_1(z)}{g_0(z)}$ is strictly increasing on $[0, \bar{z}]$.[3] Likelihood-ratio dominance implies first-order stochastic dominance.

Investment is costly, with its cost depending on the ability of each agent. Agents that do not invest, $i = 0$, incur no cost. Investing, $i = 1$, has cost $c(\theta) \geq 0$. In particular, the cost c is differentiable and strictly decreasing in ability, $c'(\theta) < 0$. We assume that $c(1) = 0$ and impose a lower bound on $c(0)$ to rule out universal investment. The restriction on $c(0)$ is stated below after the communication costs are defined. The cost is incurred before production. We interpret $c(\theta)$ as the *effort* cost of investing in education: for example, putting in the work to go to classes, study, seek help, etc. We also refer to $c(\theta)$ as the cost of acquiring information.

The education technology imposes two restrictions. First, ability affects the cost of

[3] At a given knowledge level, the likelihood ratio compares the density of investors with the density of non-investors. This is a standard assumption, for example, in the literature on moral hazard (MH). The usual interpretation is the following. If investment were unobservable (as effort is in the case of MH), the posterior "belief" that the agent invested (exerted higher effort in MH) is increasing in the observed knowledge (output in MH).

investment but not the distribution of knowledge conditional on the investment decision. Second, investment is binary and captures the extensive margin of education: whether to undertake it, rather than how much to invest.

Production and Organizations. In the *organizational* stage, knowledge (which is already realized) and time (both investors and non-investors have one unit of time) are used to produce output. In particular, output is produced by solving problems x distributed according to the cdf $P(x)$. This distribution is usually assumed to have a decreasing density: harder problems are less frequent. Nevertheless, we can normalize $P(x)$ to be uniform on $[0, 1]$.[4] Knowledge is then measured as the fraction of problems an agent can solve. Consequently, $\bar{z} \leq 1$. We may also refer to problems as *productive opportunities*. In the tax preparer example, we can think of each problem as the request of a client. It takes a full unit of time to originate a problem (draw from $P(x)$) and attempt a solution. Solving a problem x requires an agent with knowledge $z \geq x$. Under the uniform normalization, such problems occur with probability z. If the problem is solved, output is 1; otherwise output is 0.

Importantly, knowledge can be communicated between agents. It is possible for an agent to help another solve the latter's problem. In particular, it takes $h \in (0, 1)$ units of time to help another agent. The time cost h is incurred by the agent being asked for help, regardless of whether the problem is ultimately solved or not. Notice that originating a problem takes the full unit of time, so it precludes helping another agent. In this way, it may then be beneficial for agents to form teams: some agents generate the problems while others help with unsolved ones. Denote as *producers* the agents that originate problems, and as *managers* those that help others. We assume that teams can have at most two layers: a layer of producers and up to one of managers. We denote as *workers* those producers that are part of teams, and as *self-employed* (or independent producers) those producers that choose to work on their own. Workers can tell whether they can solve a problem but cannot assess the difficulty of an unsolved one before requesting help. Every unsolved problem is therefore referred to the manager, who spends h units of time assessing it.

Teams (henceforth *firms*) form before problems are drawn. In equilibrium, they take the form of a *knowledge hierarchy*. Under the two-layer restriction, an active firm is composed of one manager and n workers. All n workers will be dedicated to production (i.e. drawing problems and attempting solutions). They solve those within their knowledge

[4]Without loss of generality, we index problems by their quantiles under P, so that problems are uniformly distributed on $[0, 1]$ and knowledge z measures the fraction of problems an agent can solve. This transformation is increasing so it preserves likelihood-ratio dominance of G_1 over G_0 (Shaked and Shanthikumar, 2007, Theorem 1.C.8). See Appendix B.1 for more details.

and ask the manager for help with the rest. Whenever help is required, the manager incurs the time-cost of h.

Consider a manager with workers of knowledge $z < 1$. Each of the n workers requires help with probability $1 - z$ (problems are uniform), and each time help is needed the manager spends h units of time. Hence, for the manager to, in expectation, use his one unit of time efficiently, it must be the case that

$$n(1-z)h = 1 \quad \Longrightarrow \quad n(z) = \frac{1}{h(1-z)}. \tag{1}$$

Notice that more knowledgeable workers require help less often because they can solve a larger fraction of problems. Each sends fewer requests, allowing the manager to supervise more workers. We refer to $n(z)$ as the span of control of the manager. While, technically, the size of a team/firm is $1 + n$ (one manager and n workers), with a slight abuse of notation, $n(z)$ is also referred to as the *size* of the firm.[5]

Earnings and Investment. We start by describing the *interim* payoffs, that is, the payoffs for a known knowledge level but still uncertain problem draw. We then use them to define the *ex-ante* (expected) payoffs. A self-employed (or unmatched) agent with knowledge z gets output 1 with probability z, and output 0 with probability $1 - z$ (recall problems x are uniformly distributed). Hence, their expected payoff is z. This is the outside option of every agent, as they can always choose to remain unmatched.

A matched worker with knowledge z will earn a *wage* $w(z)$. The manager is the residual recipient of output. First, note that a manager will have more knowledge than their workers. Otherwise, the manager is unable to help with the problems the workers need help on. A manager with knowledge $m > z$ working with n workers of knowledge z earns a *rent* $R(m, z)$.[6][7] Each of the workers solves the problems drawn with probability $\mathbb{P}(x \leq z) = z$, in which case the output is one. With complementary probability the problem is unsolved by the worker, so help is required from the manager. Output is one if the manager has the knowledge to solve the problem (which happens with probability $\mathbb{P}(x \leq m | x > z)$) and is zero otherwise. Regardless of the outcome, each worker earns

[5] While $n(z)$ need not be an integer, it represents the limiting ratio of worker to manager masses between vanishing intervals of knowledge. The reader is referred to Fuchs, Garicano, and Rayo (2015, Section 2.3) for a more thorough discussion about $n(z)$ being non-integer.

[6] Note that a manager with knowledge $m \leq z$ does not contribute to production, as problems that cannot be solved by workers cannot be solved by the manager either.

[7] Appendix B.2 shows that a manager will never strictly prefer to hire workers of different knowledge. This, paired with the fact that in equilibrium matching is positively assortative, means that each manager hires workers of a single knowledge level.

their wage $w(z)$. Consequently, the rent of said manager is:

$$R(m,z) = n(z)\big(m - w(z)\big) = \frac{m - w(z)}{h(1-z)}. \tag{2}$$

The team's output is $n(z)m$: the number of problems drawn (one per worker) times the share of them that can be solved (the manager's knowledge). Notice that it is supermodular in manager knowledge m and worker knowledge z. Worker and manager knowledge are complementary: a manager can apply greater knowledge to more production opportunities when supervising more knowledgeable workers (they require less help).

Let $\mathcal{W}$, $\mathcal{S}$, and $\mathcal{M}$ denote the sets of knowledge levels of agents who choose to be workers, self-employed, and managers, respectively. It will be useful to define the following earnings function for an agent with knowledge t,

$$Y(t) = \begin{cases} w(t) & \text{if } t \in \mathcal{W} \\ t & \text{if } t \in \mathcal{S} \\ R(t) & \text{if } t \in \mathcal{M}.^{8} \end{cases} \tag{3}$$

For a given aggregate distribution of knowledge, agents with the same realized knowledge face the same occupational opportunities and earnings, regardless of their ability or education choice.

These payoffs are interim: knowledge is known but problems have not yet been drawn. Recall that knowledge levels are drawn from a chosen distribution. Define

$$U_1 := \int_0^{\bar{z}} Y(t) dG_1(t) \qquad \text{and} \qquad U_0 := \int_0^{\bar{z}} Y(t) dG_0(t)$$

as the expected gross earnings from investing and from not investing, respectively. Each agent takes the earnings schedule as given when comparing investment with non-investment. All agents face the same expected earnings gain $U_1 - U_0$. Ability affects investment through the cost $c(\theta)$. Hence, the expected payoff for an agent of ability θ is $U_1 - c(\theta)$ if they invest and U_0 if they do not. Then, an agent of type θ finds it optimal to invest if and only if

$$U_1 - U_0 \geq c(\theta). \tag{4}$$

[8] With a slight abuse of notation, let $R(t) = \max_{z<t} R(t,z)$ be the highest rent obtainable by an active manager with knowledge t, resulting from the optimal choice of which workers to hire.

That is, if the expected increase in earnings from investment is at least as large as the cost of investing. Henceforth, we refer to $U_1 - U_0$ as the marginal benefit of investing. We assume that $c(0) > 1/h$ for the communication costs considered, given that, as shown later, equilibrium earnings are bounded above by $\frac{1}{h}$ (see footnote 13)

When making an investment decision, each agent takes the earnings schedule (as a function of knowledge) as given. Nevertheless, the aggregate education choices of all agents determine the distribution of knowledge in the economy, and through it, the earnings at each knowledge level. An equilibrium requires that the anticipated earnings schedule coincide with the earnings schedule induced by the agents' choice of investment.

Competitive Equilibrium. For a measurable investment profile $\iota : \Theta \to \{0, 1\}$, the resulting aggregate distribution of knowledge is

$$G(z; \iota) = G_1(z) \int_{\{\theta \in \Theta:\, \iota(\theta)=1\}} d\Gamma(\theta) + G_0(z) \int_{\{\theta \in \Theta:\, \iota(\theta)=0\}} d\Gamma(\theta).$$

An important implication of ability only affecting investment costs is visible here: the organizational stage depends on the fraction of agents who invest, since ability does not otherwise affect the knowledge distribution or production technology.

Definition 1 (Competitive equilibrium)**.** A competitive equilibrium is an investment profile ι, occupational sets $\mathcal{W}, \mathcal{S}, \mathcal{M}$, a matching function M, and an earnings schedule Y such that, taking Y as given, agents optimally choose investment and occupations, managers hire rent-maximizing teams, and labor markets clear under the knowledge distribution $G(\cdot; \iota)$ induced by investment.[9]

3 Equilibrium

Modeling knowledge acquisition as stochastic allows us to solve the equilibrium backwards. For a tentative investment profile, the interim equilibrium determines occupations, matching, and earnings. Agents take those earnings as given when choosing whether to invest. Under the maintained assumptions, fewer investors weakly increase the expected earnings gain from investment. Together with investment costs that decrease with ability, this feedback yields a unique equilibrium investment threshold.

[9]The formal equilibrium definition is provided in Appendix B.3.

3.1 Interim Equilibrium

Fix a tentative investment profile $\iota : \Theta \to \{0, 1\}$. The investment profile determines the mass drawing knowledge from each distribution: $\iota(\theta) = 1$ indicates that type θ invests, and $\iota(\theta') = 0$ denotes no investment for type θ'. Given this investment profile, the realized (cumulative) distribution of knowledge in the economy is

$$G(z; \iota(\theta)) = G_1(z) \int_{t:\iota(t)=1} d\Gamma(t) + G_0(z) \int_{t:\iota(t)=0} d\Gamma(t)$$

with probability density function (pdf) $g(z)$.

Conditional on the resulting knowledge distribution, the organization stage does not depend on which ability types invested. For a fixed distribution of knowledge, the equilibrium (with a maximum of two layers) is first described in Garicano and Rossi-Hansberg (2004). As stated in Carmona, Laohakunakorn, and Tsujiyama (2026), the unique (interim) equilibrium is defined by a tuple $(z_1, z_2, M(z), w(z))$.

Cutoffs. The interval of knowledge levels $[0, \bar{z}]$ (i.e. the support of $G(z; \iota(\theta))$) is partitioned into three intervals. In particular, there exist cutoffs $0 < z_1 \leq z_2 < \bar{z}$, such that: the set of workers is $\mathcal{W} = [0, z_1]$, the set of self-employed is $\mathcal{S} = (z_1, z_2)$, and that of managers is $\mathcal{M} = [z_2, \bar{z}]$.[10] Agents with little knowledge benefit from access to a manager, while the most knowledgeable agents apply their expertise to many workers' problems. Intermediate knowledge types are too knowledgeable to be hired as workers (they demand higher wages) while being not knowledgeable enough to find it profitable to manage a team. Note that z_1 could potentially equal z_2, meaning that $\mathcal{S}$ could be empty. Throughout this section, we restrict attention to values of h for which every tentative investment profile generates an interior interim equilibrium. Thus $0 < z_1 < z_2 < \bar{z}$, and the self-employment region is nonempty.

For all but one result, we focus on interim equilibria with $z_1 < z_2$, that is, with a positive mass of self-employed agents. To do so we assume that the communication cost is high enough that the equilibrium contains self-employed agents at every investment profile under consideration. In the case of $\bar{z} = 1$, we can use Lemma 2 in Fuchs, Garicano, and Rayo (2015). It states that, for a fixed knowledge distribution F, there exists an $h_0(F) \in (0, 1)$ such that the equilibrium has self-employment if and only if $h > h_0(F)$.

[10] To streamline the exposition, throughout the paper we refer to z_1 and z_2 as the *(occupational) cutoffs*, and to θ^*, defined below, as the *(investment) threshold*.

Assumption 1. Helping costs are high enough:

$$h > \sup_{\alpha \in [0,1]} h_0\big(\alpha G_0 + (1-\alpha) G_1\big).$$

Here α is the share of agents who do not invest, so the supremum ranges over all investment shares. It can be shown that when $\bar{z} < 1$, $h \in (\bar{z}, 1)$ is sufficient for the presence of self-employment at any realized distribution of knowledge.

This assumption has two reasons. The first is that when the equilibrium exhibits a positive mass of self-employment, the restriction of organizations having at most one layer of managers is non-binding.[11] Second, self-employment introduces an additional margin of occupational adjustment. As we show in the next section, without self-employment, balancing the time required to assist workers with managers' available time uniquely determines the equilibrium division between workers and managers, given the knowledge distribution and communication cost. With self-employment, multiple occupational allocations satisfy this balance, so agents' earnings comparisons across occupations are also needed to determine the equilibrium allocation.

Matching. Each manager wants to utilize their time efficiently. Let $M : \mathcal{W} \to \mathcal{M}$ denote the matching function. In particular $M(z)$ denotes the knowledge of the manager matched to a worker of knowledge z. The fact that output is supermodular yields that matching is positive assortative (PAM): better workers are matched to better managers (see Lemma 1 in Fuchs, Garicano, and Rayo (2015)). For all managers to use their time efficiently, in the economy it must be the case that

$$h \int_0^z (1-t) dG(t) = \int_{z_2}^{M(z)} dG(t) \qquad \forall z \in [0, z_1] \tag{5}$$

That is, the total volume of problems passed up by workers, times the amount of time it takes to help with each problem (the left-hand side) must equal the total mass of available manager time (the right-hand side). Differentiating (5), one reaches the ordinary differential equation (ODE) that implicitly defines the matching between workers and managers:[12]

$$M_z(z) = h(1-z)\frac{g(z)}{g(M(z))} > 0 \qquad \forall z \in [0, z_1] \tag{6}$$

[11] For the analogous result that only adjacent hierarchy depths coexist, see Garicano and Rossi-Hansberg (2006, Proposition 4 and Section III.C).

[12] M_z denotes the derivative of $M(z)$ with respect to z.

It, together with $M(0) = z_2$ (the lowest worker matches with the lowest manager) and $M(z_1) = \bar{z}$ (the highest worker matches with the highest manager), characterizes the matching function. Since $M(\cdot)$ is strictly increasing, it has an inverse, which we denote $Z(m)$. This indicates the knowledge of the worker matched to a manager with knowledge m.

Earnings. A manager with knowledge m chooses worker knowledge z to maximize

$$R(m, z) = \frac{m - w(z)}{h(1 - z)}$$

At an interior optimum, the first-order condition evaluated at the equilibrium match yields the ODE that describes equilibrium wages:

$$w_z(z) = \frac{M(z) - w(z)}{1 - z}. \tag{7}$$

More knowledgeable workers require less help and permit larger teams, but they command higher wages. The manager's optimal choice balances these effects. The marginal worker must be indifferent between being self-employed or matched to a manager, $w(z_1) = z_1$. Similarly, the marginal manager must be indifferent between being a manager or being self-employed, so $R(z_2) = z_2$. Because $M(0) = z_2$, this last condition implies that $w(0) = (1 - h)z_2$.[13]

Finally, for a tentative investment profile $\iota(\theta)$ the interim equilibrium is given by the tuple

$$\Big\{ z_1\big(\iota(\theta)\big),\ z_2\big(\iota(\theta)\big),\ M\big(z; \iota(\theta)\big),\ w\big(z; \iota(\theta)\big) \Big\}.$$

3.2 Single-Agent Problem

A single agent has zero population mass, so an individual investment decision does not change the aggregate knowledge distribution or the resulting interim earnings. The interim earnings profile generated by $\iota(\theta)$ is (as a function of knowledge t)[14]

$$Y(t; \iota) = \begin{cases} w(t; \iota) & \text{for } t \in \big[0, z_1(\iota)\big] \\ t & \text{for } t \in \big(z_1(\iota), z_2(\iota)\big) \\ R(t; \iota) & \text{for } t \in \big[z_2(\iota), \bar{z}\big] \end{cases}$$

[13] We can also see that the highest possible earnings are $R(\bar{z}) = \frac{\bar{z} - z_1}{h(1 - z_1)} \leq \frac{1}{h}$.

[14] To ease notation we omit the dependence of ι on θ.

Investment changes the distribution of knowledge over which the agent evaluates the same interim earnings schedule. Then, expected earnings under investment choice i are

$$U_i(\iota) := \int_0^{\bar{z}} Y(t;\iota) dG_i(t) \quad \text{for } i = 0, 1$$

Since the earnings schedule $Y(t;\iota)$ is strictly increasing, and investment strictly improves the distribution of knowledge realization, $G_1 \succ_{FOSD} G_0$ (likelihood-ratio order implies first-order stochastic dominance), we have that $U_1(\iota) - U_0(\iota) > 0$.

Consequently, for a given tentative investment profile $\iota(\theta)$, an agent of ability θ finds it optimal to invest if and only if $U_1(\iota) - c(\theta) \geq U_0(\iota)$. The lowest ability agents would not invest even if they were guaranteed to get the highest possible earnings: $c(0) > \frac{1}{h}$; and the highest ability agents find it costless to invest: $c(1) = 0$. Thus, we have $c(1) < U_1(\iota) - U_0(\iota) < \frac{1}{h} < c(0)$, and $c(\theta)$ is strictly decreasing. Therefore, there exists a $\theta^*(\iota)$ such that $U_1(\iota) - U_0(\iota) = c\big(\theta^*(\iota)\big)$. Optimal investment therefore follows a threshold rule: all those agents with ability higher than θ^* (and lower cost) invest, while those with ability lower than the threshold do not. More concretely,

$$i^*(\theta;\iota) = \begin{cases} 1 & \text{if } \theta \geq \theta^*(\iota) \\ 0 & \text{if } \theta < \theta^*(\iota) \end{cases}$$

The left panel of Figure 1 varies ability while holding interim earnings fixed. The horizontal blue line represents $U_1 - U_0$, and the red curve represents $c(\theta)$. Their intersection gives the optimal investment threshold for those earnings.

Since in equilibrium $\iota(\theta) = i^*(\theta;\iota)$ for almost every θ, we can summarize any possible equilibrium investment profile ι by a threshold θ^*. An equilibrium is a threshold θ^* such that the marginal investor must be indifferent under the earnings generated by the same threshold. Namely, it is a θ^* such that

$$U_1(\theta^*) - U_0(\theta^*) = c(\theta^*) \tag{8}$$

3.3 Full Equilibrium

To find the fixed point θ^* that solves (8) we need to understand how the expected earnings gain from investment $U_1(\theta^*) - U_0(\theta^*)$ behaves as a function of the investment threshold θ^*. It is shown that $U_1(\theta^*) - U_0(\theta^*)$ is increasing in θ^*. This, alongside $c(\theta)$ being decreasing, implies that the equilibrium θ^* is unique.

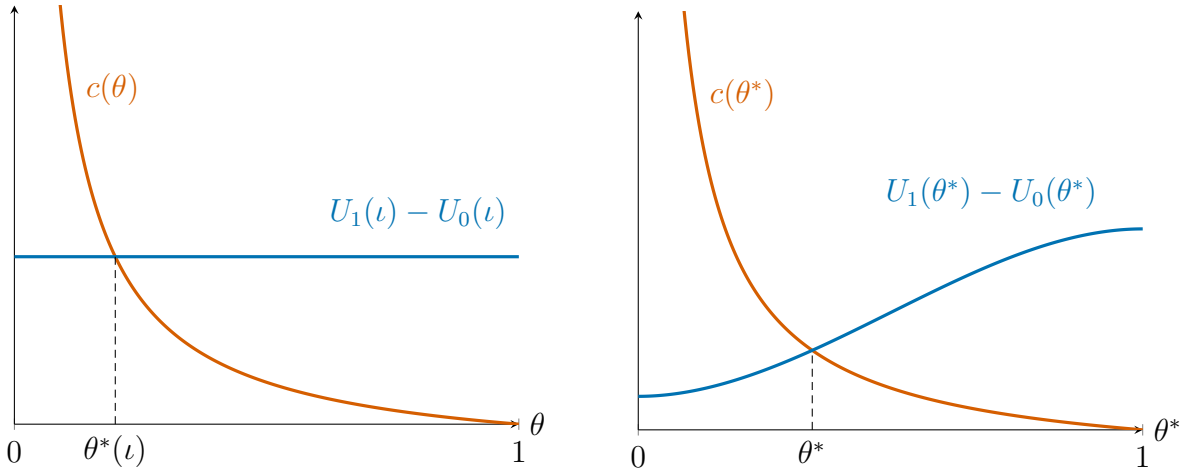


Figure 1: Individual investment decisions (left) and equilibrium determination (right).

Under the threshold rule for investment, a fraction $\Gamma(\theta^*)$ does not invest. Consequently, the investor share is $1 - \Gamma(\theta^*)$.[15] The realized distribution of knowledge for a tentative threshold θ^* will have cdf

$$G(z;\theta^*) = \Gamma(\theta^*)G_0(z) + \big(1 - \Gamma(\theta^*)\big)G_1(z)$$

and pdf

$$g(z;\theta^*) = \Gamma(\theta^*)g_0(z) + \big(1 - \Gamma(\theta^*)\big)g_1(z).$$

Raising the investment threshold increases the fraction drawing knowledge from G_0 and reduces the fraction drawing from G_1. This leads to a worse distribution of knowledge in the likelihood-ratio order.[16] Thus, $G(z;\theta_0^*) \succ_{LR} G(z;\theta_1^*)$: a higher investment threshold reduces the relative abundance of higher knowledge levels.

Holding h, G_0, and G_1 fixed, a higher investment threshold lowers occupational cutoffs and wages at knowledge levels that remain workers, while raising rents at knowledge levels that remain managers.

Proposition 1. *Let $G_1 \succ_{LR} G_0$, and $\theta_1^* > \theta_0^*$. Then:*

1. *The occupational cutoffs satisfy $z_1(\theta_1^*) \leq z_1(\theta_0^*)$ and $z_2(\theta_1^*) \leq z_2(\theta_0^*)$;*
2. *On $[0, z_1(\theta_1^*)]$, the matching function under θ_1^*, $M(z;\theta_1^*)$, single-crosses the matching function under θ_0^*, $M(z;\theta_0^*)$, from below;*[17]

[15] The ability distribution has no atoms.

[16] To see this, for $\theta_1^* > \theta_0^*$, write the ratio of densities as $\dfrac{g(z;\theta_0^*)}{g(z;\theta_1^*)} = \dfrac{\Gamma(\theta_0^*) + (1 - \Gamma(\theta_0^*))\ell(z)}{\Gamma(\theta_1^*) + (1 - \Gamma(\theta_1^*))\ell(z)}$, where $\ell(z) = \frac{g_1(z)}{g_0(z)}$. When the investor shares differ, the displayed ratio is increasing in $\ell(z)$. Since $\ell(z)$ is increasing in z, the density ratio is increasing in z.

[17] Namely, if $M(z;\theta_1^*) \geq M(z;\theta_0^*)$, then $M(z';\theta_1^*) \geq M(z';\theta_0^*)$ for all $z < z' \leq z_1(\theta_1^*)$.

3. *The wage functions satisfy* $w(z;\theta_1^*) \leq w(z;\theta_0^*)$ *for all* $z \in [0, z_1(\theta_1^*)]$*;*
4. *The rent functions satisfy* $R(m;\theta_1^*) \geq R(m;\theta_0^*)$ *for all* $m \in [z_2(\theta_0^*), \bar{z}]$.

Proof. See Appendix A.1. □

Holding occupational cutoffs fixed, a worse knowledge distribution (a higher θ^*) increases total demand for managerial time and reduces the mass of managers, as equation (5) shows at $z = z_1$. Low-knowledge is relatively more abundant, applying downward pressure on wages. Lower wages make self-employment relatively more attractive. Thus, the knowledge of the marginal worker, z_1, decreases. In turn, this reduces demand for help. Lower wages also translate to higher rents for managers. Higher managerial earnings make management relatively more attractive than self-employment, pushing z_2 (the knowledge of the marginal manager) down, thus increasing available managerial time. For part 2, the least knowledgeable worker is matched to a less knowledgeable manager. The worker with knowledge $z_1(\theta_1^*)$ is matched to $\bar{z}$ under θ_1^*, while the manager assigned to that same knowledge level under θ_0^* is less knowledgeable.

It should be noted that while the cutoffs decrease, since g is also changing, the mass of agents in each occupation is not determined by the cutoff comparisons alone. For example, consider the mass of agents that are workers, $G\big(z_1(\theta^*), \theta^*\big)$. There are two opposing forces at play. On the one hand, an increase in θ^* leads to a decrease in $z_1(\theta^*)$, pushing the mass of workers down. On the other hand, an increase in θ^* means that the new distribution is first-order stochastically worse than the previous. This pushes the mass of workers up.

The next result compares expected investment gains using the optimality conditions of the interim equilibria and Proposition 1.

Corollary 2. $U_1(\theta^*) - U_0(\theta^*)$ *is increasing in* θ^*.

Proof. See Appendix A.2. □

A higher θ^* means that fewer people invest in education. Proposition 1 then tells us that wages of those agents that remain workers go down, while rents of those that remain managers go up. Investing increases the chances of drawing higher knowledge compared to not investing. These wage and rent changes favor investment, which raises the probability of drawing knowledge levels associated with management.

Theorem 3. *There exists a unique equilibrium investment threshold* θ^*.

Proof. When everyone invests, $U_1(0) - U_0(0) - c(0) < 0$. When no one invests, $U_1(1) - U_0(1) - 0 > 0$. $U_1(\theta^*) - U_0(\theta^*) - c(\theta^*)$ is continuous. Therefore, by the intermediate value

theorem, there exists a θ^* such that $U_1(\theta^*) - U_0(\theta^*) = c(\theta^*)$. Furthermore, by Corollary 2, the left-hand side is increasing while the right-hand side is strictly decreasing. Thus, the crossing is unique. Finally, for any realized distribution of knowledge, the interim equilibrium is unique. □

The right panel of Figure 1 incorporates the feedback from aggregate investment to individual incentives. Unlike the left panel, the horizontal axis varies the tentative investment threshold θ^*. Interim earnings are recomputed at each threshold, yielding the curve $U_1(\theta^*) - U_0(\theta^*)$. At the intersection of $U_1(\theta^*) - U_0(\theta^*)$ and $c(\theta^*)$, the implied investment decisions reproduce the threshold used to generate those earnings.

4 Economic Implications

Having established that the equilibrium investment threshold is unique, we can now evaluate the economic implications of changes in the environment. We first discuss a decrease in communication costs. Our novel sufficient condition allows us to evaluate what happens for a fixed distribution of knowledge. We then provide a separate condition under which a reduction in this cost reduces investment. Next we study a decrease in the costs of investing and an improvement in the distribution of ability. We conclude by exploring what happens under better investment outcomes. Each comparison varies one primitive and distinguishes changes in investment incentives from the resulting reorganization of production. Unless stated otherwise, investment remains interior and the interim equilibrium contains self-employment, with $z_1 < z_2$. Wage and rent comparisons hold realized knowledge fixed and concern occupations common to both equilibria. Following Garicano and Rossi-Hansberg (2004), we denote as earnings inequality the ratio of the highest to the lowest (gross) earnings.

Cheaper communication can discourage investment because it raises the value of poor knowledge realizations that education helps individuals avoid. Lower education costs raise investment, as does a better distribution of ability. Conversely, better educational outcomes can reduce investment. The difference comes from equilibrium earnings. A reduction in costs makes investment more attractive at any given earnings schedule. Better educational outcomes also change the supply of knowledge and the organization of production, which can reduce the expected earnings advantage of investing.

4.1 Lower Communication Costs, h

4.1.1 At a Fixed Investment Profile

As noted in Garicano and Rossi-Hansberg (2004), even when the distribution of knowledge is fixed comparative statics with respect to h are hard without parametric assumptions.[18] Fuchs, Garicano, and Rayo (2015) establish that, when $\bar{z} = 1$, a decrease in h increases the cutoff between workers and self-employed. The effect on the marginal manager is ambiguous because there are two opposing forces. On one side, as stated above, the change increases the mass of workers. Those workers require managerial help, pushing the marginal manager's knowledge down. On the other hand, helping is now less costly, so for the same mass of workers fewer managers are required. This pushes the marginal managerial knowledge up. Our novel condition, a non-decreasing knowledge density, signs the response of the management cutoff.

Lemma 4. *If $\bar{z} = 1$ and $G(z;\theta^*)$ is convex (the density $g(z;\theta^*)$ is increasing), then $\frac{dz_2}{dh} > 0$.*

Proof. See Appendix A.5. □

To understand the mechanism, consider a reduction in h while holding z_2 fixed. The existing mass of managers can supervise more workers, so z_1 rises. Positive assortative matching then assigns incumbent workers to less knowledgeable managers: the new workers have higher knowledge, so they are hired by the most knowledgeable managers. To keep z_2 indifferent between management and self-employment, the wage of the lowest workers must increase. This propagates up to all workers. When the density of knowledge is increasing, reallocating a given mass of managers requires small reductions in their knowledge: more mass is at higher knowledge levels. This bounds the reduction in output per worker, limiting its downward force on wages enough that z_1 strictly prefers to earn a wage than to be self-employed. The equilibrium z_2 must fall.

Note that convexity of G_0 (increasing g_0) is sufficient for $G(z;\theta^*)$ to be convex for any θ^*. Likelihood-ratio dominance of G_1 over G_0 implies that g_1 is also increasing, so their mixture is increasing.

Proposition 5. *Suppose that $\bar{z} = 1$ and G_0 is convex. Holding investment fixed, a decrease in h*

- *increases z_1 and decreases z_2;*
- *raises wages and rents;*
- *lowers the matching function on the common worker domain.*

[18] They solve for the case of uniformly distributed problems and knowledge.

Proof. See Appendix A.6. □

Self-employment contracts at both boundaries. Individuals near the lower boundary join teams as workers, and individuals near the upper boundary become managers. For continuing workers, the saving in managerial time more than offsets the loss from matching with a less knowledgeable manager, so wages rise. A continuing manager matches with more knowledgeable workers and can supervise a larger team. Both the lower time requirement per question and the change in matching increase managerial earnings.

4.1.2 The Investment Response

The reduction in the helping cost has two effects on investment. On the one hand, because asking for help is less costly the earnings of workers increase. This force reduces the value of investing: the lowest realizations of knowledge are now more valuable than before. On the other hand, each manager can now help a larger team of workers (conditional on worker knowledge). The rents of managers increase. This increases the incentive to invest: the largest knowledge realizations (those that investing stochastically "produces more" of) increase in value. To determine the effect of the decrease in h on the investment threshold θ^*, it is necessary to tame these forces. To do so we make use of the following decomposition of the no-investment density:

$$g_0(z) = \frac{1}{\ell(1)} g_1(z) + \left(1 - \frac{1}{\ell(1)}\right) \frac{\ell(1) g_0(z) - g_1(z)}{\ell(1) - 1}.$$

Likelihood-ratio dominance ensures that the final fraction is a non-negative density that is likelihood-ratio dominated by G_0. Denote it by

$$f(z) := \frac{\ell(1) g_0(z) - g_1(z)}{\ell(1) - 1}.$$

The non-investment density can be decomposed as the mixture of g_1 and a residual density f.[19] This residual places relatively more weight on low knowledge realizations. In this representation, education eliminates exposure to the residual downside component.

The following condition ensures that the gains associated to low knowledge realizations dominate, reducing investment.

[19] The largest weight α for which $g_0 - \alpha g_1$ is always non-negative is $\frac{1}{\ell(1)}$. To express g_0 as a mixture of g_1 and another density, the contribution of g_1 cannot exceed g_0 (otherwise the other density would have to be negative). At $z = 1$, this means that the weight on g_1 cannot exceed $\frac{1}{\ell(1)}$. This weight also satisfies the constraint at every lower knowledge level because $\ell(z)$ is increasing.

Proposition 6. *Suppose that $\bar{z} = 1$ and g_0 is increasing. If*

$$\int_0^1 t f(t)dt < (1-h)\left[\int_0^1 t g_0(t)dt - \frac{1}{1+h}\right], \tag{9}$$

then a marginal decrease in h raises θ^* and reduces equilibrium investment.

Proof. See Appendix A.7. □

The left-hand side is the mean knowledge of $f(z)$, the component of g_0 that education removes. Since the mean is positive, the condition requires that $\int_0^1 t g_0(t)dt > \frac{1}{1+h}$ and that the mean of $f(z)$ be sufficiently low. In economic terms, agents must be relatively knowledgeable on average even without education, while education protects them against particularly poor realizations.

These restrictions determine which group benefits more from cheaper communication. High average knowledge without education limits the economy's potential demand for managerial time and hence its aggregate gain from a lower h. By contrast, the poor realizations against which education protects are especially dependent on managerial help. Under condition (9) the expected earnings gain associated with these realizations exceeds the average gain in the economy and the gain under the investment distribution. Non-investors therefore benefit more than investors, $U_1 - U_0$ falls at the original investment profile, and marginal investors no longer find education worth its cost.

The investment response modifies the organizational effects of cheaper help. The decline in participation shifts knowledge toward G_0, working against the initial increase in wages and reinforcing the increase in managerial rents. The combined effect lowers the management cutoff, so the final equilibrium has a lower z_2 and higher rents for common managers. The final wage response and the movement in z_1 need not have the signs obtained at a fixed investment share. Nevertheless, a positive mass of knowledge levels that remain in the worker interval receive higher wages. The fall in investment cannot erase the wage gain for every incumbent worker: otherwise cheaper help would favor high-knowledge managerial outcomes relative to worker outcomes, increasing the return to education—contradicting the decline in investment.[20]

4.2 Lower Investment Costs, $c(\theta)$

Lower investment costs increase participation in education and improve the equilibrium distribution of knowledge. Suppose the cost function is indexed by a parameter τ that

[20] See Appendix B.6

decreases costs, $c(\theta;\tau)$, point-wise with increases in τ: for $\tau' > \tau$ $c(\theta;\tau') < c(\theta;\tau)$ for all $\theta \in [0,1)$. Under the interpretation of $c(\theta)$ as an effort cost, an increase in τ represents an improvement in technologies that make accessing education easier. For example, a transition from having to go to the library to get information to being able to search for information on the internet.

Since τ only affects costs, at a fixed θ^* it does not change the realized distribution of knowledge nor the organization of production. Thus, marginal benefit of investing, $U_1 - U_0$, remains unchanged. Nevertheless, at the original equilibrium, the marginal agent now strictly prefers to invest (their costs decreased). Then, an increase in τ, by decreasing $c(\theta;\tau)$, leads to a lower equilibrium investment threshold. Since the mass of investors is $1 - \Gamma(\theta^*)$, it increases.

Proposition 7. *Suppose investment costs decrease point-wise. Then, for the economy as a whole:*

- *more agents invest (i.e. θ^* decreases);*
- *average earnings increase;*
- *overall earnings inequality decreases;*
- *average firm size increases;*

For workers:

- *wages increase;*
- *average worker earnings increase;*

For managers:

- *rents decrease;*
- *within-manager inequality decreases;*

Proof. The investment response follows from (8). The occupational and earnings comparisons follow from Proposition 1; the remaining implications are explained in the discussion below. □

The reduction in investment costs lowers the equilibrium investment threshold, as described above. Given that the mass of investors is $1 - \Gamma(\theta^*)$, this mass increases. Given a decrease in θ^*, the realized distribution of knowledge in the economy improves in the likelihood-ratio sense. By Proposition 1, this leads to an increase in both occupational cutoffs, z_1 and z_2, an increase in wages and a decrease in rents. The intuition is the same as the one in Proposition 1. Because the investment distribution itself is unchanged, the higher management cutoff also lowers the probability of becoming a manager conditional on investing. Education becomes more widespread even as a managerial outcome becomes less likely for each investor.

Recall that we use the earnings ratio of highest to lowest earnings as a measure of inequality. The overall inequality is then, $\frac{R(\bar{z};\theta^*)}{w(0;\theta^*)}$. It can easily be verified that the numerator is $\frac{\bar{z}-z_1}{h(1-z_1)}$ (if $\bar{z}=1$ it then equals $\frac{1}{h}$). The denominator is $w(0;\theta^*)=(1-h)z_2(\theta^*)$. Since z_1 and z_2 are decreasing in θ^* (see Proposition 1) a reduction in costs, by reducing θ^*, induces an increase in both. The numerator is decreasing in z_1, while the denominator increases, hence, overall inequality decreases. Similarly, the inequality among managers decreases: the denominator is now the earnings of the lowest manager. Since said manager has to be indifferent to being self-employed, he earns z_2. The numerator decreases while the denominator increases, so within-manager inequality decreases. On the other hand, within-worker inequality is ambiguous: it is $\frac{z_1}{(1-h)z_2}$. Since both numerator and denominator increase, the direction is ambiguous.

Average earnings increase in the economy because the better knowledge distribution improves the productive capabilities in the economy: the relative increase in higher realizations of knowledge means that managers and self-employed can solve more problems, and workers require less help. Thus output, which equals average earnings, increases. Average worker wages rise for three reinforcing reasons: the wage schedule shifts upward, the interval of workers increases, and the conditional distribution of knowledge (conditional on being a worker) improves in the LR-sense (see Theorem 1.C.6 in Shaked and Shanthikumar (2007)). While rents decrease, average rents for managers may rise. Lower rents at a given knowledge level push the average down, but the exit of the least knowledgeable managers and the improvement in managerial knowledge push it up.

The induced reorganization changes the size of firms. Average firm size increases because the improved conditional knowledge distribution of workers reduces their average demand for help. A manager can therefore supervise more workers. Individual firms need not all expand. Identifying the firms by the knowledge of their manager, some initially small firms disappear when their managers become self-employed. Managers near the new management cutoff hire less knowledgeable workers and operate smaller teams, while managers near the top hire more knowledgeable workers and expand.

4.3 Better Distribution of Ability, $\Gamma(\theta)$

We take a slight detour to evaluate the effects of an improvement in the distribution of ability in the economy. Suppose that for $\tau' > \tau$, $\Gamma(\theta;\tau') \succ_{FOSD} \Gamma(\theta;\tau)$ and that Γ is differentiable in τ. Then, for a given equilibrium θ^*, an increase in τ yields an increase in the threshold θ^*. Nevertheless, the mass of investors increases.

Proposition 8. *If the distribution of ability θ becomes FOSD-better, then the threshold for*

investment θ^* increases, but the mass of agents that invest increases. The effects on equilibrium earnings, inequality, and firm size match those of Proposition 7.

Proof. See Appendix A.3 □

The improvement in Γ places (weakly) more individuals above the initial threshold, so the direct effect is an increase in the share of investors. The resulting improvement in the aggregate knowledge distribution lowers the improvement in expected earnings from investing, so the equilibrium threshold rises. This adjustment offsets only part of the initial increase in participation. If it eliminated that increase, the knowledge distribution and the difference in expected earnings would return to their original values. But the higher marginal ability would face a lower investment cost, contradicting indifference.

4.4 Improvement in the Investment Distribution, $G_1(z)$

We now proceed to analyze the other side of the possible improvements in *knowledge-acquisition technologies*, namely an improvement in the outcomes of education investment. Better educational outcomes raise the value of investment at unchanged earnings, but can reduce investment after wages and rents adjust.

Let the investment distribution, $G_1(z;\tau)$, be indexed by a parameter τ such that, for $\tau' > \tau$, $G_1(z;\tau') \succ_{LR} G_1(z;\tau)$. This means that $\frac{g_1(z;\tau')}{g_1(z;\tau)}$ is increasing in z. For unchanged earnings, a better distribution in the LR-sense means better knowledge realizations. Since earnings are increasing in knowledge, this raises the expected payoff from investing.

The same improvement changes earnings even before participation adjusts. To compare aggregate knowledge distributions after an increase in τ we make the following assumption on the likelihood ratio between the distribution of investment and that of non-investment, $\ell(z,\tau) := \frac{g_1(z;\tau)}{g_0(z)}$.

Assumption 2. For $\tau' > \tau$,

$$\frac{\partial \ell(z;\tau')}{\partial z} > \frac{\partial \ell(z,\tau)}{\partial z}.$$

This is a sufficient condition that guarantees that $G(z;\theta^*,\tau') \succ_{LR} G(z;\theta^*,\tau)$. It states that the likelihood ratio increases faster as τ' increases.

Under this assumption, at a fixed investment threshold, the realized distribution of knowledge is better in the likelihood-ratio order. By Proposition 1 both occupational cutoffs increase, wages go up and rents decrease. Investors are more likely than non-investors to obtain the high knowledge realizations, so the adjustment in earnings works

against the direct benefit of better draws. Which force dominates determines the investment response. It is worth mentioning that examples exist for a movement in equilibrium investment in any direction.

The strength of this adjustment depends on the initial mass of investors. At the hypothetical $\theta^* = 1$, nobody invests and the realized distribution of knowledge is G_0. Improving G_1 then has no effect on earnings: only the direct benefit of better draws is present. Thus, for θ^* close to one, $U_1(\theta^*) - U_0(\theta^*)$ increases. A condition under which better educational outcomes reduce investment must therefore account for the initial participation level.

4.4.1 When Better Outcomes Reduce Investment

We focus the remainder of this subsection on the case in which an improvement in outcomes reduces investment. In order to explain the intuition behind the result we impose some additional assumptions. First, we consider likelihood-ratio improvements that are "proportional". Namely,

$$\frac{\partial g_1(z;\tau)}{\partial \tau} = g_1(z;\tau) - g_0(z). \tag{10}$$

This perturbation increases the difference between the investor and non-investor distributions. The improvement removes probability from knowledge levels at which investors are less represented than non-investors and adds it at levels at which they are more represented. We assume these changes keep $g_1(z;\tau)$ strictly positive. For every knowledge threshold z, it reduces investors' probability of falling below that threshold by

$$-\frac{\partial G_1(z;\tau)}{\partial \tau} = G_0(z) - G_1(z;\tau).$$

This perturbation improves both the investment distribution, G_1, and, at a fixed investment threshold, the aggregate knowledge distribution, G, in the likelihood-ratio order.

For the next proposition we assume that the helping cost satisfies

$$h < h_0(G), \tag{11}$$

such that there is no self-employment in the original equilibrium. That is, $z_1 = z_2 = \hat{z}$, where $\hat{z}$ is defined from the equation determining the efficient use of manager time:

$$h \int_0^{\hat{z}} (1-t) dG(t;\theta^*,\tau) = 1 - G(\hat{z};\theta^*,\tau). \tag{12}$$

As briefly discussed in the model section, conditional on G and h this equation uniquely determines the cutoff $\hat{z}$.

Proposition 9. *Assume that $\bar{z} = 1$, that there is no self-employment at the original equilibrium θ^*, and that the LR-improvement in g_1 is proportional. If*

$$g(\hat{z}; \theta^*, \tau) < \big(1 - \Gamma(\theta^*)\big)(1-h)\big[G_0(\hat{z}) - G_1(\hat{z}; \tau)\big]^2, \tag{13}$$

then a marginal increase in τ raises the equilibrium investment threshold $\frac{d\theta^*}{d\tau} > 0$, reducing the mass of investors.

Proof. See Appendix A.4 □

This condition guarantees that the effect of the adjustment in earnings due to reorganization outweighs the improvement that comes from higher realizations when investing. To see this it is useful to rewrite the condition as

$$\frac{1}{h} < (1 - \Gamma(\theta^*))\frac{(1-h)}{h}\frac{G_0(\hat{z}) - G_1(\hat{z}; \tau)}{g(\hat{z}; \theta^*, \tau)}\big[G_0(\hat{z}) - G_1(\hat{z}; \tau)\big]$$

The left-hand side is an upper bound on the earnings achievable by investing: $\frac{1}{h}$ is the maximum payoff achievable, it is the rent of $\bar{z}$ when $\bar{z} = 1$ and $z_1 < z_2$. Hence, investment yields earnings below this value (the investment distribution has full support).

Fix the investment and $\hat{z}$. The proportional improvement in G_1 means that the "additional mass above the cutoff" increases by the mass of investors times the increase:

$$(1 - \Gamma(\theta^*))(G_0(\hat{z}) - G_1(\hat{z}; \tau)).$$

Since the distribution is fixed, a small increase in the cutoff removes population mass approximately equal to the density $g(\hat{z}; \theta^*, \tau)$ times the increase in the cutoff. At a fixed cutoff, the improvement both increases the mass of potential managers and reduces workers' demand for managerial help. Restoring the balance of managerial time therefore requires a cutoff increase at least as large as

$$\frac{(1 - \Gamma(\theta^*))[G_0(\hat{z}) - G_1(\hat{z}; \tau)]}{g(\hat{z}; \theta^*, \tau)}.$$

If the density at the cutoff is low, $\hat{z}$ must move substantially.

At the original cutoff, working and managing pay the same, but managerial earnings rise faster than worker earnings because managers leverage their knowledge across a team. Consequently, at the original earnings schedule, the advantage of managing

slightly above the cutoff is approximately $R'(\hat{z}) - w'(\hat{z})$ times the additional knowledge. That is the improvement in income that must be overcome to make those people willing to become workers. To raise the occupational cutoff, worker pay must improve relative to managerial pay by enough to offset that advantage. Thus, the required change in pay is approximately $R'(\hat{z}) - w'(\hat{z})$ times the increase in the cutoff. Since $R'(\hat{z}) = \frac{1}{h}$ and $w'(\hat{z}) \leq 1$, we can get a lower bound for the required change in pay as

$$\frac{1-h}{h} \frac{(1-\Gamma(\theta^*))[G_0(\hat{z}) - G_1(\hat{z};\tau)]}{g(\hat{z};\theta^*,\tau)}.$$

That improvement in pay, times the increase in the probability of drawing above $\hat{z}$, $G_0(\hat{z}) - G_1(\hat{z};\tau)$, means that improvement in earnings due to investing then falls by at least

$$\frac{1-h}{h} \frac{\big(1-\Gamma(\theta^*)\big)\big[G_0(\hat{z}) - G_1(\hat{z};\tau)\big]}{g(\hat{z};\theta^*,\tau)} \big[G_0(\hat{z}) - G_1(\hat{z};\tau)\big].$$

Condition 13 guarantees that the resulting reduction in the earnings advantage of investing exceeds the direct benefit of better draws (which is at most $\frac{1}{h}$). The premium for investing consequently falls at the initial investment profile, and the marginal investors no longer find education worth its unchanged cost.

Economically, when the density at the occupational cutoff, $g(\hat{z};\theta^*,\tau)$, is low, a small adjustment of it moves little mass between occupations. Hence, in order to equalize the demand and supply for manager time, $\hat{z}$ must change considerably. A large increase in $\hat{z}$ requires a large increase in worker wages. After all, the cutoff must earn at least their outside option. The condition guarantees that this increase in wages is large enough.

Some comments are in order. The condition on Proposition 9 is a *sufficient* condition for the improvement in G_1 to reduce investment. We can find distributions that do not satisfy the condition, but for which a likelihood-ratio improvement in G_1 still increases θ^*, even when self-employment is present in the original equilibrium.

4.4.2 Equilibrium Earnings

Return to interim equilibria with self-employment. If better educational outcomes reduce investment, the participation response works against the initial improvement in the aggregate knowledge distribution. At the original participation level, better education raises wages and weakly lowers rents. Lower participation then shifts the distribution toward G_0, pushing wages down and rents up. In general, the final knowledge distribution need not be ordered relative to the initial one in the likelihood-ratio sense. The directions of the initial earnings changes therefore do not by themselves determine the

final changes.

Nevertheless, there are some claims that can be established. First and foremost, it is not possible for the general equilibrium effects (decrease in investment) to completely erase the wage improvement resulting from the better investment technology. This means that there is a positive mass of common workers that strictly benefits.[21]

5 Concluding Remarks

This paper studies the interaction between investment in education and the organization of knowledge within firms. Individuals acquire knowledge before knowing their occupation or the earnings they will receive. Modeling education as a risky investment separates the acquisition of knowledge acquisition from how it is optimally organized while maintaining the equilibrium effects: education decisions determine the knowledge available in the economy, and the organization of that knowledge determines the returns to education.

We show that the equilibrium is unique and derive comparative statics. Holding investment fixed, we establish a sufficient condition to sign the effects of lower helping costs, and show when it can reduce equilibrium investment. Our main finding is that reducing education costs and improving the distribution of outcomes can have opposing effects on participation. Lower costs increase investment, raise worker wages, and reduce managerial rents. At a given managerial knowledge, large teams expand while small teams contract. A better distribution of outcomes can discourage investment due to the resulting changes in occupations, matching, and earnings. Thus, changes in educational participation cannot be understood independently of how firms use and reward the knowledge that education produces.

Regarding assumptions of the model: binary investment and ability-independent knowledge distributions simplify the analysis. Both assumptions can be relaxed, although preserving the equilibrium characterization and comparative statics requires additional restrictions.

Two extensions are particularly relevant. Endogenous organizational depth, as in Garicano and Rossi-Hansberg (2006), would allow education to affect the number of managerial layers as well as team sizes. Finally, it would be interesting to study how the introduction of AI, à la Ide and Talamas (2025), would affect the incentives to invest in education.

[21] See Appendix B.7.

A Proofs

A.1 Proof of Proposition 1

Statement. *Let $G_1 \succ_{LR} G_0$, and $\theta_1^* > \theta_0^*$. Then:*

1. *The occupational cutoffs satisfy $z_1(\theta_1^*) \leq z_1(\theta_0^*)$ and $z_2(\theta_1^*) \leq z_2(\theta_0^*)$;*
2. *The matching function under θ_1^*, $M(z;\theta_1^*)$, single-crosses the matching function under θ_0^*, $M(z;\theta_0^*)$, from below;*[22]
3. *The wage functions satisfy $w(z;\theta_1^*) \leq w(z;\theta_0^*)$ for all $z \in [0, z_1(\theta_1^*)]$;*
4. *The rent functions satisfy $R(m,\theta_1^*) \geq R(m;\theta_0^*)$ for all $m \in [z_2(\theta_0^*), \bar{z}]$.*

The proof of Proposition 1 consists of four lemmas. We first prove points 1 through 4 hold for any parameter τ, indexing the realized distribution of knowledge $g(z;\tau)$, such that the score is decreasing. Recall that for a distribution $g(z;\tau)$ parametrized by τ, the score (or informant) is $S(z;\tau) = \frac{\partial \log g(z;\tau)}{\partial \tau} = \frac{g_\tau(z;\tau)}{g(z;\tau)}$. The next lemma describes how the occupational cutoffs, $z_1(\tau)$ and $z_2(\tau)$, move with movements in a parameter τ.[23]

Lemma 10. *If the score $S(z;\tau)$ is decreasing (increasing) in z, then $z_1(\tau)$ and $z_2(\tau)$ are decreasing (increasing) in τ.*

Proof. We will show the result for $S(\cdot;\tau)$ decreasing.

Differentiating the Matching ODE. The matching is defined by the ODE (6) (where M_z denotes the derivative of M with respect to the first argument, z)

$$M_z(z;\tau) = h(1-z)\frac{g(z;\tau)}{g(M(z;\tau);\tau)} \tag{6}$$

with boundary conditions $M(0;\tau) = z_2(\tau)$ and $M(z_1(\tau);\tau) = \bar{z}$.

Differentiating (6) with respect to τ, we get:[24][25]

$$M_{\tau z}(z;\tau) = M_z(z;\tau)\left[\Big(S(z;\tau) - S(M(z;\tau);\tau)\Big) - \frac{g_z(M(z;\tau);\tau)}{g(M(z;\tau);\tau)}M_\tau(z;\tau)\right]$$

[22] Namely, if $M(z;\theta_1^*) \geq M(z;\theta_0^*)$, then $M(z';\theta_1^*) \geq M(z';\theta_0^*)$ for all $z' > z$.

[23] The cutoffs' dependence on τ is through the distribution $g(z;\tau)$.

[24] See Theorem 2.11 in Teschl (2012)

[25] To see this: $M_{\tau z}(z;\tau) = h(1-z)\left[\frac{g_\tau(z;\tau)g(M(z;\tau);\tau) - g(z;\tau)\Big(g_\tau(M(z;\tau);\tau) + g_z(M(z;\tau);\tau)M_\tau(z;\tau)\Big)}{g(M(z;\tau);\tau)^2}\right]$. Rearranging one gets $M_{\tau z}(z;\tau) = h(1-z)\left[\frac{g(z;\tau)}{g(M(z;\tau);\tau)}\Big(S(z) - S(M(z;\tau))\Big) - \frac{g(z;\tau)}{g(M(z;\tau);\tau)}\frac{g_z(M(z;\tau);\tau)M_\tau(z;\tau)}{g(M(z;\tau);\tau)}\right]$, from which we can apply (6).

Let

$$\begin{aligned} Y(z;\tau) &:= M_\tau(z;\tau) \\ a(z;\tau) &:= M_z(z;\tau)\frac{g_z(M(z;\tau);\tau)}{g(M(z;\tau);\tau)} \\ b(z;\tau) &:= M_z(z;\tau)\Big[S(z;\tau) - S(M(z;\tau);\tau)\Big] \end{aligned}$$

Monotonicity of $S(\cdot;\tau)$ will yield the sign of $b(z;\tau)$. Since we are assuming S is decreasing, and $M(z) > z$, we have that $b > 0$. We then get the following ODE:

$$Y_z(z;\tau) + a(z;\tau)Y(z;\tau) = b(z;\tau) \tag{14}$$

We get the boundary conditions from differentiating the boundaries of the original ODE:

$$M(0;\tau) = z_2(\tau) \quad \implies Y(0;\tau) = \frac{dz_2}{d\tau} \tag{15}$$

$$M(z_1(\tau);\tau) = \bar{z} \quad \implies M_z(z_1(\tau);\tau)\frac{dz_1}{d\tau} + Y(z_1(\tau);\tau) = 0. \tag{16}$$

Integrating factor. To solve the differential equation we make use of an integrating factor $\mu(z)$. For it to be an integrating factor it must be the case that $\mu'(z) = a(z)\mu(z)$.Hence $(\ln(\mu(z))' = a(z)$, resulting in $\mu(z;\tau) = k\exp\left(\int a(u;\tau)du\right)$. Where k comes from the constant of integration. Notice that[26]

$$\begin{aligned} \int a(u;\tau)du = C + \int_0^z a(u;\tau)du &= C + \int_0^z M_z(u;\tau)\frac{g_z(M(u;\tau);\tau)}{g(M(u;\tau);\tau)}du \\ = C + \int_{M(0;\tau)}^{M(z;\tau)} \frac{g_z(t;\tau)}{g(t;\tau)}dt &= C + \ln g(t;\tau)\Big|_{M(0;\tau)}^{M(z;\tau)} \\ = C + \ln g(M(z;\tau);\tau) - \ln g(z_2(\tau);\tau) &= C + \ln\frac{g(M(z;\tau);\tau)}{g(z_2(\tau);\tau)} \end{aligned}$$

for some constant C. Consequently,

$$\begin{aligned} \mu(z;\tau) &= \exp\left(\ln\frac{g(M(z;\tau);\tau)}{g(z_2(\tau);\tau)} + C\right) \\ &= k\frac{g(M(z;\tau);\tau)}{g(z_2(\tau);\tau)} \end{aligned}$$

where $k = \exp(C)$.

[26]Note for the weary reader: the antiderivative is $G(x) = \int g(x)dx$, the definite integral is $\int_0^x g(t)dt = G(x) - G(0)$. Hence $G(x) = G(0) + \int_0^x g(t)dt$.

Solving the ODE. Going back to (14):

$$\Big(\mu(z;\tau)Y(z;\tau)\Big)' = \mu(z;\tau)b(z;\tau) \tag{17}$$

$$\implies \mu(z;\tau)Y(z;\tau) = \int \mu(t;\tau)b(t;\tau)dt + C_1$$

$$\implies Y(z;\tau) = \frac{1}{\mu(z;\tau)}\left(\int \mu(t;\tau)b(t;\tau)dt + C_1\right) \tag{18}$$

Let's now rewrite

$$\int \mu(z;\tau)b(z;\tau)dz = C_0 + \int_0^z \mu(s;\tau)b(s;\tau)ds$$

Then,

$$Y(z;\tau) = \frac{C_0 + C_1 + \int_0^z \mu(t;\tau)b(t;\tau)dt}{\mu(z;\tau)}$$

Then, evaluating at zero we get (the second equality comes from the boundary condition (15))

$$Y(0;\tau) = \frac{C_0 + C_1}{\mu(0;\tau)} = \frac{dz_2}{d\tau}$$

Plugging back in:

$$Y(z;\tau) = \frac{\mu(0;\tau)\frac{dz_2}{d\tau} + \int_0^z \mu(t;\tau)b(t;\tau)dt}{\mu(z;\tau)} \tag{19}$$

Notice that the constants multiplying μ all cancel out, turning $\mu(0) = 1$.

Differentiating the Wage ODE. Consider (49) at z_1 together with the indifference condition at z_2 (i.e. $w(0) = (1-h)z_2$).

$$w(z_1) = (1 - z_1)\left[(1-h)z_2 + \int_0^{z_1} \frac{M(t)}{(1-t)^2}dt\right]$$

$$\frac{z_1}{1 - z_1} = (1-h)z_2 + \int_0^{z_1} \frac{M(t)}{(1-t)^2}dt.$$

Differentiating we get

$$\frac{\frac{dz_1}{d\tau}}{(1-z_1)^2} - (1-h)\frac{dz_2}{d\tau} = \frac{M(z_1;\tau)}{(1-z_1)^2}\frac{dz_1}{d\tau} + \int_0^{z_1} \frac{M_\tau(t;\tau)}{(1-t)^2}dt$$

Given that $M(z_1;\tau) = \bar{z}$:

$$\frac{1-\bar{z}}{(1-z_1)^2}\frac{dz_1}{d\tau} = (1-h)\frac{dz_2}{d\tau} + \int_0^{z_1} \frac{Y(t;\tau)}{(1-t)^2}dt \tag{20}$$

From (19), the integral becomes

$$\begin{aligned}\int_0^{z_1} \frac{Y(t;\tau)}{(1-t)^2}dt &= \frac{dz_2}{d\tau}\underbrace{\int_0^{z_1} \frac{\mu(0;\tau)}{(1-t)^2\mu(t;\tau)}dt}_{:=I_1} + \underbrace{\int_0^{z_1} \frac{1}{(1-t)^2}\frac{\int_0^t \mu(s;\tau)b(s;\tau)ds}{\mu(t;\tau)}dt}_{:=I_2} \\ &= I_1\frac{dz_2}{d\tau} + I_2.\end{aligned}$$

I stands for (definite) integral. Notice that $I_1 > 0$, and since $b(s;\tau) > 0$, so is $I_2 > 0$. Thus

$$\frac{1-\bar{z}}{(1-z_1)^2}\frac{dz_1}{d\tau} = (1-h+I_1)\frac{dz_2}{d\tau} + I_2 \tag{21}$$

From the boundary condition (16) we have

$$\frac{dz_1}{d\tau} = -\frac{Y(z_1;\tau)}{M_z(z_1;\tau)}.$$

Substituting 19 we reach

$$\frac{dz_1}{d\tau} = -\frac{\frac{dz_2}{d\tau} + \int_0^{z_1} \mu(t;\tau)b(t;\tau)dt}{M_z(z_1;\tau)\mu(z_1;\tau)}.$$

Denote by $I_3 := \int_0^{z_1} \mu(t;\tau)b(t;\tau)dt$. Again, since $b > 0$, $I_3 > 0$.

Then, (21) becomes

$$-\frac{1-\bar{z}}{(1-z_1)^2}\frac{\frac{dz_2}{d\tau} + I_3}{\mu(z_1;\tau)M_z(z_1;\tau)} = (1-h+I_1)\frac{dz_2}{d\tau} + I_2$$

To ease notation, denote $T := \frac{1-\bar{z}}{(1-z_1)^2\mu(z_1;\tau)M_z(z_1;\tau)}$. We have that $T \geq 0$ because $\bar{z} \leq 1$,

$z_1 < 1$, $M_z > 0$ (by PAM), and $\mu > 0$. Grouping terms we get

$$-(1-h+I_1+T)\frac{dz_2}{d\tau} = I_2 + TI_3$$
$$\frac{dz_2}{d\tau} = -\frac{I_2+TI_3}{1-h+I_1+T} < 0. \tag{22}$$

The last inequality follows because $I_2 > 0$ and $T \geq 0$, so the numerator is positive, and $h < 1$, $I_1 > 0$ and $T \geq 0$, so the denominator is also positive. Hence with the minus sing, the derivative is negative. If the score were increasing, then $b < 0$ and so the signs of I_2 and I_3 flip, turning the derivative positive.

Change in z_1. Return to the boundary condition (16), we have

$$\frac{dz_1}{d\tau} = -\frac{Y(z_1;\tau)}{M_z(z_1;\tau)}.$$

Since $M_z > 0$ (the matching is PAM), we need to show that $Y(z_1;\tau) > 0$ (if S is decreasing).

From (19),

$$\begin{aligned} Y(z_1;\tau) &= \frac{\frac{dz_2}{d\tau} + \int_0^{z_1} \mu(t;\tau)b(t;\tau)dt}{\mu(z_1;\tau)} \\ &= \frac{\frac{dz_2}{d\tau} + I_3}{\mu(z_1;\tau)}. \end{aligned}$$

Substituting (22):

$$\begin{aligned} Y(z_1;\tau) &= \frac{-\frac{I_2+TI_3}{1-h+I_1+T} + I_3}{\mu(z_1;\tau)} \\ &= \frac{-I_2 - TI_3 + (1-h+I_1+T)I_3}{(1-h+I_1+T)\mu(z_1;\tau)} \\ &= \frac{-I_2 + (1-h+I_1)I_3}{(1-h+I_1+T)\mu(z_1;\tau)} \end{aligned}$$

We now aim to show that $I_2 \leq I_1 I_3$ ($\geq$ when S is increasing) to show that the numerator is positive.

Recall:

$$I_1 := \int_0^{z_1} \frac{dt}{(1-t)^2\mu(t;\tau)}$$
$$I_2 := \int_0^{z_1} \frac{1}{(1-t)^2\mu(t;\tau)} \underbrace{\int_0^t \mu(s;\tau)b(s;\tau)ds}_{=:I(t)} dt$$

Notice that $I(t) \leq I_3$ for all t, as $I(z_1) = I_3$.[27]. Hence:

$$I_2 = \int_0^{z_1} \frac{1}{(1-t)^2\mu(t;\tau)} I(t)dt \leq I_3 \int_0^{z_1} \frac{dt}{(1-t)^2\mu(t;\tau)} = I_3 I_1.$$

Going back to $Y(z_1;\tau)$ we reach:

$$Y(z_1;\tau) = \frac{(1-h)I_3 + \overbrace{I_1 I_3 - I_2}^{\geq 0}}{(1-h+I_1+T)\mu(z_1;\tau)}.$$

Which is positive because both numerator and denominator are positive.

Since $Y(z_1) > 0$ we have that $\frac{dz_1}{d\tau} = -\frac{Y(z_1)}{M_z(z_1)} < 0$ (signs are reversed when the score is increasing). That is both cutoffs decrease as the distribution becomes worse in the MLRP sense.

□

The next lemma proves the statement about the matching function.

Lemma 11. *Let $\tau' > \tau$. If the score $S(z;\tau)$ is decreasing, then $M(z;\tau)$ single-crosses $M(z;\tau')$ from above. Namely, if $M(z;\tau') \geq M(z;\tau)$, then $M(z';\tau') > M(z';\tau)$ for all $z < z' \leq z_1(\tau')$.*

Proof. Define $\phi(z) = M(z;\tau') - M(z;\tau)$. First, lets look at the boundaries. By Lemma 10, $z_2(\tau') < z_2(\tau)$, so we have that $M(0;\tau') < M(0;\tau)$ and so $\phi(0) < 0$. Similarly, $z_1(\tau') < z_1(\tau)$, hence $M(z_1(\tau');\tau') = \bar{z} > M(z_1(\tau');\tau)$.[28] That means that $\phi(z_1(\tau')) > 0$. The claim is then that $\phi(\cdot)$ single-crosses zero from below.

Since M is continuous, $\phi(z)$ is continuous too. Then, by the intermediate value theorem, $\phi(z)$ crosses zero at least once. Let $\widetilde{z}$ be such a point. We have $\widetilde{m} := M(\widetilde{z};\tau) = M(\widetilde{z};\tau')$. Let's evaluate the derivative (both $M(\cdot)$ are differentiable) of ϕ at $\widetilde{z}$.

[27] $I(t) \geq I_3$ for all t when $b < 0$, that is, when S is increasing

[28] Note that we evaluate at $z_1(\tau')$, the upper endpoint of the common worker domain.

$$\begin{aligned}\phi'(\widetilde{z}) &= M'(\widetilde{z};\tau') - M'(\widetilde{z};\tau) \\ &= h(1-\widetilde{z})\left[\frac{g(\widetilde{z};\tau')}{g(\widetilde{m};\tau')} - \frac{g(\widetilde{z};\tau)}{g(\widetilde{m};\tau)}\right] > 0\end{aligned}$$

The term in brackets is positive because $\frac{g(z;\tau)}{g(z;\tau')}$ is increasing in z for $\tau' > \tau$, and $\widetilde{m} > \widetilde{z}$. Therefore, whenever ϕ crosses zero it does so with strictly positive slope. Hence, ϕ must single cross zero.

□

For the effect on wages we have the following lemma.

Lemma 12. *Let $\tau' > \tau$. If $S(z;\tau)$ is decreasing in z, then $w(z;\tau) \geq w(z;\tau')$ for all $z \in [0, z_1(\tau')]$.*[29]

Proof. First notice that by Lemma 10, $z_2(\tau') < z_2(\tau)$ and $z_1(\tau') < z_1(\tau)$. Since z_2 must be indifferent between being a manager or being self-employed, and considering that z_2 is matched to workers of type 0, we know that $w(0) = (1-h)z_2$. Hence $w(0;\tau') < w(0;\tau)$. Similarly, z_1 must be indifferent between being a worker or self-employed, so we must have $w(z_1) = z_1$. So we also have that $w(z_1(\tau');\tau') = z_1(\tau') < w(z_1(\tau');\tau)$, where the inequality comes from optimal occupational choice of $z_1(\tau')$ under τ. Define $\xi(z) := w(z;\tau) - w(z;\tau')$. We have that $\xi(0) > 0$ and $\xi(z_1(\tau')) > 0$.

Towards a contradiction suppose there exist y such that $\xi(y) < 0$. By continuity and IVT, there exist $a < y < b$ such that $w(a;\tau) = w(a;\tau')$ ($\xi(a) = 0$) and $w'(a;\tau) \leq w'(a;\tau')$, and $w(b;\tau) = w(b;\tau')$ ($\xi(b) = 0$) and $w'(b;\tau) \geq w'(b;\tau')$.

At a we then have

$$\begin{aligned}w'(a;\tau) = \frac{M(a;\tau) - w(a;\tau)}{1-a} &\leq \frac{M(a;\tau') - w(a;\tau')}{1-a} = w'(a;\tau') \\ \implies M(a;\tau) &\leq M(a;\tau').\end{aligned}$$

At b we have

$$\begin{aligned}w'(b;\tau') = \frac{M(b;\tau') - w(b;\tau')}{1-b} &\leq \frac{M(b;\tau) - w(b;\tau)}{1-b} = w'(b;\tau) \\ \implies M(b;\tau') &\leq M(b;\tau).\end{aligned}$$

This contradicts Lemma 11, that $M(\cdot;\tau)$ single-crosses $M(\cdot;\tau')$. □

[29] For those in $[z_1(\tau'), z_1(\tau)]$, for sufficiently small increase in τ, income decreases since they used to earn $w(z;\tau) > z$ and now they are self-employed earning z.

Finally, the effect on rents.

Lemma 13. *Let $\tau' > \tau$. If $S(z;\tau)$ is decreasing in z, then $R(m;\tau) \leq R(m;\tau')$ for all $m \in [z_2(\tau), \bar{z}]$.*

Proof. By Lemma 10, $z_2(\tau') < z_2(\tau)$. For sufficiently small increase in τ, for those $m \in [z_2(\tau'), z_2(\tau))$ the payoff increases, as $R(m) \geq m$. Next, notice that $R(z_2(\tau);\tau) = z_2 < R(z_2(\tau);\tau')$ Recall that $Z(m)$ is the inverse of $M(z)$, then

$$R(m;\tau) = \frac{m - w(Z(m;\tau);\tau)}{h(1 - Z(m;\tau))}$$

Taking the derivative with respect to τ:

$$\begin{aligned}\frac{\partial R(m;\tau)}{\partial \tau} &= \frac{\Big[-w'(Z(m;\tau);\tau)\frac{\partial Z}{\partial \tau} - w_\tau(Z(m;\tau);\tau)\Big]h(1 - Z(m;\tau)) - \Big[m - w(Z(m;\tau);\tau)\Big](-h\frac{\partial Z}{\partial \tau})}{h^2(1 - Z(m;\tau))^2} \\ &= \frac{-w_\tau(Z(m;\tau);\tau)}{h(1 - Z(m;\tau))} \geq 0\end{aligned}$$

Where the second equality makes use of (7), and the inequality comes from $w_\tau \leq 0$ (from Lemma 12). □

We are now ready to prove Proposition 1.

Proof of Proposition 1. Let $\tau = \Gamma(\theta^*)$, then $S(z;\tau)$ is decreasing (recall $\ell(z) = \frac{g_1(z)}{g_0(z)}$):

$$\begin{aligned}S(z;\tau) &= \frac{g_0(z) - g_1(z)}{\tau g_0(z) + (1-\tau)g_1(z)} = \frac{1 - \ell(z)}{\tau + (1-\tau)\ell(z)} \\ \frac{\partial S(z;\tau)}{\partial \ell} &= \frac{-1}{\big(\tau + (1-\tau)\ell)\big)^2} < 0\end{aligned}$$

Hence:

$$\frac{\partial S(z;\tau)}{\partial z} = \frac{\partial S(z;\tau)}{\partial \ell}\frac{\partial \ell(z)}{\partial z} \leq 0.$$

Since Γ is a strictly increasing cdf, by Lemma 10, z_1 and z_2 decrease with θ^*. Lemma 11 yields the single-crossing of the matching function, while Lemmas 12 and 13 give the results for wages and rent, respectively. □

A.2 Proof of Corollary 2.

Statement. *$U_1(\theta^*) - U_0(\theta^*)$ is increasing in θ^*.*

Proof. Let $\tau := \Gamma(\theta^*)$, so $g(z;\tau) = \tau g_0(z) + (1-\tau)g_1(z)$. We will prove that $\frac{\partial(U_1-U_0)}{\partial\Gamma} = \frac{\partial(U_1-U_0)}{\partial\tau} \geq 0$, from which the result follows given that (by being a cdf) $\frac{\partial\Gamma}{\partial\theta^*} > 0$. See the proof of Proposition 1 to see that the score is decreasing.

Define $\zeta(\tau) := U_1(\tau) - U_0(\tau)$. We will show that $\frac{\partial\zeta}{\partial\tau} \geq 0$. The expression for ζ is

$$\begin{aligned}\zeta(\tau) := &\int_0^{z_1(\tau)} w(t;\tau)\Big[g_1(t) - g_0(t)\Big]dt \\ &+ \int_{z_1(\tau)}^{z_2(\tau)} t\Big[g_1(t) - g_0(t)\Big]dt \\ &+ \int_{z_2(\tau)}^{\bar{z}} R(t;\tau)\Big[g_1(t) - g_0(t)\Big]dt\end{aligned}$$

Then,

$$\begin{aligned}\frac{\partial\zeta(\tau)}{\partial\tau} = &\, w(z_1(\tau);\tau)\Big[g_1(z_1(\tau)) - g_0(z_1(\tau))\Big]\frac{\partial z_1}{\partial\tau} + \int_0^{z_1(\tau)} \frac{\partial w(t;\tau)}{\partial\tau}\Big[g_1(t) - g_0(t)\Big]dt \\ &+ z_2(\tau)\Big[g_1(z_2(\tau)) - g_0(z_2(\tau))\Big]\frac{\partial z_2}{\partial\tau} - z_1(\tau)\Big[g_1(z_1(\tau)) - g_0(z_1(\tau))\Big]\frac{\partial z_1}{\partial\tau} \\ &- R(z_2(\tau);\tau)\Big[g_1(z_2(\tau)) - g_0(z_2(\tau))\Big]\frac{\partial z_2}{\partial\tau} + \int_{z_2(\tau)}^{\bar{z}} \frac{\partial R(t;\tau)}{\partial\tau}\Big[g_1(t) - g_0(t)\Big]dt\end{aligned}$$

Using the indifference conditions at the cutoffs, $w(z_1) = z_1$ and $R(z_2) = z_2$, we reach:

$$\frac{\partial\zeta(\tau)}{\partial\tau} = \int_0^{z_1(\tau)} \frac{\partial w(t;\tau)}{\partial\tau}\Big[g_1(t) - g_0(t)\Big]dt + \int_{z_2(\tau)}^{\bar{z}} \frac{\partial R(t;\tau)}{\partial\tau}\Big[g_1(t) - g_0(t)\Big]dt$$

From the proof of Lemma 13 $\frac{\partial R}{\partial\tau} = \frac{-\frac{\partial w}{\partial\tau}(Z(m;\tau);\tau)}{h(1-Z(m;\tau))}$. Now consider the change of variables (in the second integral) $t = M(z;\tau)$ (so $dt = M'(z;\tau)dz$). We then have $\frac{\partial R}{\partial\tau} = \frac{-\frac{\partial w}{\partial\tau}(z;\tau)}{h(1-z)}$. Notice that the limits of integration also are affected. Consequently:

$$\frac{\partial\zeta(\tau)}{\partial\tau} = \int_0^{z_1(\tau)} \frac{\partial w(t;\tau)}{\partial\tau}\Big[g_1(t) - g_0(t)\Big]dt + \int_0^{z_1(\tau)} \frac{-\frac{\partial w}{\partial\tau}(z;\tau)}{h(1-z)}\Big[g_1(M(z;\tau)) - g_0(M(z;\tau))\Big]M'(z;\tau)dz$$

Using the matching ODE (6)

$$= \int_0^{z_1(\tau)} \frac{\partial w(t;\tau)}{\partial\tau}\left\{\Big[g_1(t) - g_0(t)\Big] - \frac{g(t;\tau)}{g(M(t;\tau);\tau)}\Big[g_1(M(t;\tau)) - g_0(M(t;\tau))\Big]\right\}dt \tag{23}$$

Notice that $g_1(t) - g_0(t) = -g_\tau(t;\tau) = -S(t;\tau)g(t;\tau)$. Hence

$$\begin{aligned}\frac{\partial \zeta(\tau)}{\partial \tau} &= \int_0^{z_1(\tau)} \frac{\partial w(t;\tau)}{\partial \tau}\Bigg\{ - g(t;\tau)S(t;\tau) + \frac{g(t;\tau)}{g(M(t;\tau);\tau)} g(M(t;\tau);\tau)S(M(t;\tau);\tau)\Bigg\}dt \\ &= \int_0^{z_1(\tau)} \frac{\partial w(t;\tau)}{\partial \tau}\Big\{S(M(t;\tau);\tau) - S(t;\tau)\Big\}g(t;\tau)dt\end{aligned}$$

As established in the proof of Proposition 1, the score $S(t;\tau)$ is decreasing in t, so the term in brackets is strictly negative. As established in Lemma 12, since the score is decreasing, $\frac{\partial w}{\partial \tau} \leq 0$. Furthermore, $w_\tau(0;\tau) = (1-h)\frac{dz_2}{d\tau} < 0$, so by continuity, the integrand is strictly positive and so is $\frac{\partial \zeta}{\partial \tau}$. □

A.3 Proof of Proposition 8

Let us first determine the change in the equilibrium threshold, θ^*. It is determined by

$$U_1(\theta^*, \tau) - U_0(\theta^*, \tau) = c(\theta^*).$$

Where $U_i = \int_0^{\bar{z}} Y(t;\theta^*,\tau)g_i(t)dt$. Let $\Phi \equiv U_1(\theta^*,\tau) - U_0(\theta^*,\tau) - c(\theta^*) = 0$ Using the implicit function theorem:

$$\frac{\partial \theta^*}{\partial \tau} = -\frac{\frac{\partial \Phi}{\partial \tau}}{\frac{\partial \Phi}{\partial \theta^*}}$$

Evaluating each term:

$$\frac{\partial \Phi}{\partial \tau} = \frac{\partial\Big(U_1(\theta^*;\tau) - U_0(\theta^*,\tau)\Big)}{\partial \tau} = \frac{\partial\Big(U_1(\theta^*;\tau) - U_0(\theta^*,\tau)\Big)}{\partial \Gamma}\frac{\partial \Gamma}{\partial \tau}$$

In the proof of Corollary 2, it was shown that $\frac{\partial(U_1 - U_0)}{\partial \Gamma} \geq 0$. On the other hand $\frac{\partial \Gamma}{\partial \tau} \leq 0$ by FOSD. Hence the numerator is negative.

The denominator:

$$\frac{\partial \Phi}{\partial \theta^*} = \frac{\partial\Big(U_1(\theta^*;\tau) - U_0(\theta^*,\tau)\Big)}{\partial \Gamma}\frac{\partial \Gamma}{\partial \theta^*} - c'(\theta^*).$$

The first term is positive as established in Corollary 2. The second is negative by assumption on $c(\theta)$. Consequently, the denominator is positive. The numerator is negative, denominator is positive, and there is a minus sign. As a result, $\frac{\partial \theta^*}{\partial \tau} \geq 0$.

Now, we show that the mass of non-investors declines. Let $\mu := \Gamma(\theta^*;\tau)$ be the mass

of agents that do not invest. Taking the total derivative with respect to τ:

$$\frac{d\mu}{d\tau} = \Gamma_\theta(\theta^*;\tau)\frac{d\theta^*}{d\tau} + \Gamma_\tau(\theta^*;\tau)$$

Substituting $\frac{d\theta^*}{d\tau}$

$$\begin{aligned}
\frac{d\mu}{d\tau} &= \frac{-\Gamma_\theta \frac{\partial \Phi}{\partial \tau} + \Gamma_\tau \frac{\partial \Phi}{\partial \theta^*}}{\frac{\partial \Phi}{\partial \theta^*}} \\
&= \frac{-\Gamma_\theta \left(\frac{\partial \Big(U_1(\theta^*;\tau) - U_0(\theta^*,\tau) \Big)}{\partial \Gamma} \Gamma_\tau \right) + \Gamma_\tau \left(\frac{\partial \Big(U_1(\theta^*;\tau) - U_0(\theta^*,\tau) \Big)}{\partial \Gamma} \Gamma_\theta - c'(\theta^*) \right)}{\frac{\partial \Phi}{\partial \theta^*}} \\
&= \frac{-\Gamma_\tau c'(\theta^*)}{\frac{\partial \Phi}{\partial \theta^*}} \leq 0
\end{aligned}$$

The inequality comes from $\frac{\partial \Phi}{\partial \theta^*} > 0$, $c' < 0$, and $\Gamma_\tau \leq 0$ (by FOSD). Hence the mass of non-investors, μ, declines.

A.4 Proof of Proposition 9

Proof of Proposition 9. Partial derivatives with respect to μ vary the noninvestor share while holding G_0 and $G_1(\cdot;\tau)$ fixed. Partial derivatives with respect to τ hold μ fixed. In both cases the organizational allocation and earnings adjust. Partial derivatives of M, w, and R hold the knowledge argument fixed; at moving endpoints they are understood as the corresponding interior limits. Unless a total derivative is displayed, all expressions below are evaluated at the initial investment profile.

Strict positivity of the densities makes the left-hand side of (12) strictly increasing in $\hat{z}$. The cutoff is therefore differentiable. The integrated matching equation then gives differentiable matching through inversion of the population CDF. The linear wage equation, together with occupational indifference, determines the wage level with a strictly positive coefficient $1 + h(1 - \hat{z})$. Consequently, earnings are differentiable as well. The strict inequality in (11) keeps the occupational regime unchanged locally.

The response of the common cutoff. Differentiating (12) with respect to μ gives

$$\begin{aligned} 0 = [1 + h(1 - \hat{z})][\mu g_0(\hat{z}) + (1 - \mu) g_1(\hat{z}; \tau)] \frac{\partial \hat{z}}{\partial \mu} \\ + G_0(\hat{z}) - G_1(\hat{z}; \tau) + h \int_0^{\hat{z}} (1 - t)[g_0(t) - g_1(t; \tau)] \, dt. \end{aligned}$$

Integration by parts yields

$$\int_0^{\hat{z}} (1 - t)[g_0(t) - g_1(t; \tau)] \, dt = (1 - \hat{z})[G_0(\hat{z}) - G_1(\hat{z}; \tau)] + \int_0^{\hat{z}} [G_0(t) - G_1(t; \tau)] \, dt.$$

Hence,

$$\begin{aligned} -\frac{\partial \hat{z}}{\partial \mu} &= \frac{G_0(\hat{z}) - G_1(\hat{z}; \tau)}{\mu g_0(\hat{z}) + (1 - \mu) g_1(\hat{z}; \tau)} \\ &\quad + \frac{h \int_0^{\hat{z}} [G_0(t) - G_1(t; \tau)] \, dt}{[1 + h(1 - \hat{z})][\mu g_0(\hat{z}) + (1 - \mu) g_1(\hat{z}; \tau)]} \\ &\geq \frac{G_0(\hat{z}) - G_1(\hat{z}; \tau)}{\mu g_0(\hat{z}) + (1 - \mu) g_1(\hat{z}; \tau)}. \end{aligned} \tag{24}$$

The inequality follows from $G_1(\cdot; \tau) \succ_{LR} G_0$, which implies $G_0(t) - G_1(t; \tau) \geq 0$.

A lower bound on the earnings response. The rent envelope equation and the matching equation give

$$\frac{\partial R(M(t))}{\partial \mu} = -\frac{1}{h(1 - t)} \frac{\partial w(t)}{\partial \mu}, \qquad M'(t) = h(1 - t) \frac{\mu g_0(t) + (1 - \mu) g_1(t; \tau)}{\mu g_0(M(t)) + (1 - \mu) g_1(M(t); \tau)}. \tag{25}$$

In the first equality, the partial derivative of R is evaluated at managerial knowledge $M(t)$; it is not the total derivative of the composition. Differentiating the investment premium, the moving-boundary terms cancel because $w(\hat{z}) = R(\hat{z})$. Changing variables from manager knowledge to matched worker knowledge then yields

$$\begin{aligned} \frac{\partial (U_1 - U_0)}{\partial \mu} = -\int_0^{\hat{z}} \frac{\partial w(t)}{\partial \mu} \Bigg[& g_0(t) - g_1(t; \tau) \\ & - \frac{\mu g_0(t) + (1 - \mu) g_1(t; \tau)}{\mu g_0(M(t)) + (1 - \mu) g_1(M(t); \tau)} [g_0(M(t)) - g_1(M(t); \tau)] \Bigg] dt. \end{aligned} \tag{26}$$

Differentiating integrated matching first with respect to μ and then with respect to worker knowledge gives

$$\begin{aligned}
&\frac{\partial}{\partial t}\left([\mu g_0(M(t)) + (1-\mu)g_1(M(t);\tau)]\frac{\partial M(t)}{\partial \mu}\right) \\
&\quad = h(1-t)\Bigg[g_0(t) - g_1(t;\tau) \\
&\qquad - \frac{\mu g_0(t) + (1-\mu)g_1(t;\tau)}{\mu g_0(M(t)) + (1-\mu)g_1(M(t);\tau)}[g_0(M(t)) - g_1(M(t);\tau)]\Bigg].
\end{aligned} \tag{27}$$

This identity only requires continuity of the densities: after the first differentiation, the product on the left is differentiable in worker knowledge through the integrated matching equation. The differentiated wage equation is

$$\frac{\partial M(t)}{\partial \mu} = (1-t)\frac{\partial}{\partial t}\left(\frac{\partial w(t)}{\partial \mu}\right) + \frac{\partial w(t)}{\partial \mu}. \tag{28}$$

Substitution into (26) and integration by parts therefore give

$$\begin{aligned}
h\frac{\partial(U_1 - U_0)}{\partial \mu} = &-\left[\frac{\partial w(t)}{\partial \mu}\frac{\mu g_0(M(t)) + (1-\mu)g_1(M(t);\tau)}{1-t}\frac{\partial M(t)}{\partial \mu}\right]_0^{\hat{z}} \\
&+ \int_0^{\hat{z}} \frac{\mu g_0(M(t)) + (1-\mu)g_1(M(t);\tau)}{(1-t)^2}\left(\frac{\partial M(t)}{\partial \mu}\right)^2 dt.
\end{aligned} \tag{29}$$

The boundary term must use the common occupational cutoff rather than indifference with self-employment. Differentiating $M(0) = \hat{z}$ and $M(\hat{z}) = 1$ gives

$$\begin{aligned}
&\frac{\partial M(0)}{\partial \mu} = \frac{\partial \hat{z}}{\partial \mu}, \\
&\frac{\partial M(\hat{z})}{\partial \mu} = -h(1-\hat{z})\frac{\mu g_0(\hat{z}) + (1-\mu)g_1(\hat{z};\tau)}{\mu g_0(1) + (1-\mu)g_1(1;\tau)}\frac{\partial \hat{z}}{\partial \mu}.
\end{aligned}$$

Since the lowest manager employs workers of knowledge zero,

$$w(\hat{z}) = R(\hat{z}) = \frac{\hat{z} - w(0)}{h}.$$

Differentiating this equality along the moving cutoff yields

$$h\frac{\partial w(\hat{z})}{\partial \mu} + \frac{\partial w(0)}{\partial \mu} = [1 - hw'(\hat{z})]\frac{\partial \hat{z}}{\partial \mu}. \tag{30}$$

Consequently, the negative boundary term in (29) equals

$$[\mu g_0(\hat{z}) + (1-\mu) g_1(\hat{z};\tau)] \frac{\partial \hat{z}}{\partial \mu} \left[h \frac{\partial w(\hat{z})}{\partial \mu} + \frac{\partial w(0)}{\partial \mu} \right]$$
$$= [\mu g_0(\hat{z}) + (1-\mu) g_1(\hat{z};\tau)][1 - h w'(\hat{z})] \left(\frac{\partial \hat{z}}{\partial \mu} \right)^2.$$

Thus the exact response identity is

$$\begin{aligned} h \frac{\partial (U_1 - U_0)}{\partial \mu} &= [\mu g_0(\hat{z}) + (1-\mu) g_1(\hat{z};\tau)][1 - h w'(\hat{z})] \left(\frac{\partial \hat{z}}{\partial \mu} \right)^2 \\ &\quad + \int_0^{\hat{z}} \frac{\mu g_0(M(t)) + (1-\mu) g_1(M(t);\tau)}{(1-t)^2} \left(\frac{\partial M(t)}{\partial \mu} \right)^2 dt. \end{aligned} \tag{31}$$

Because $M(\hat{z}) = 1$ and $w(\hat{z}) > \hat{z}$, the wage equation implies

$$w'(\hat{z}) = \frac{1 - w(\hat{z})}{1 - \hat{z}} < 1, \qquad 1 - h w'(\hat{z}) > 1 - h > 0.$$

Dropping the nonnegative integral in (31) and applying (24) gives

$$\frac{\partial (U_1 - U_0)}{\partial \mu} \geq \frac{1-h}{h} \frac{[G_0(\hat{z}) - G_1(\hat{z};\tau)]^2}{\mu g_0(\hat{z}) + (1-\mu) g_1(\hat{z};\tau)}. \tag{32}$$

The direct gain and the investment response. At the initial parameter value,

$$\begin{aligned} &\frac{\partial}{\partial \tau} [\mu g_0(t) + (1-\mu) g_1(t;\tau)] \\ &\quad = -(1-\mu) \frac{\partial}{\partial \mu} [\mu g_0(t) + (1-\mu) g_1(t;\tau)]. \end{aligned}$$

The organizational allocation depends on the knowledge distributions only through their population mixture. Its differentiable response therefore satisfies, away from the occupational boundary,

$$\frac{\partial Y(t)}{\partial \tau} = -(1-\mu) \frac{\partial Y(t)}{\partial \mu}.$$

The boundary terms again cancel when expected earnings are differentiated. Using (10),

$$\begin{aligned} \frac{\partial (U_1 - U_0)}{\partial \tau} &= \int_0^1 Y(t) \frac{\partial g_1(t;\tau)}{\partial \tau} \, dt + \int_0^1 \frac{\partial Y(t)}{\partial \tau} [g_1(t;\tau) - g_0(t)] \, dt \\ &= (U_1 - U_0) - (1-\mu) \frac{\partial (U_1 - U_0)}{\partial \mu}. \end{aligned} \tag{33}$$

The highest manager employs workers of knowledge $\hat{z}$, so

$$R(1) = \frac{1 - w(\hat{z})}{h(1 - \hat{z})} \leq \frac{1}{h}.$$

Earnings are increasing in knowledge. Together with strictly positive noninvestment earnings, this implies $U_1 - U_0 < 1/h$. Combining (32) and (33),

$$\frac{\partial(U_1 - U_0)}{\partial \tau} < \frac{1}{h} - \frac{(1-\mu)(1-h)}{h} \frac{[G_0(\hat{z}) - G_1(\hat{z};\tau)]^2}{\mu g_0(\hat{z}) + (1-\mu) g_1(\hat{z};\tau)} < 0. \tag{34}$$

The final inequality follows from (13).

Finally, the investment equilibrium satisfies $U_1 - U_0 = c(\theta^*)$ and $\mu = \Gamma(\theta^*)$. Differentiating gives

$$\frac{d\theta^*}{d\tau} = -\frac{\dfrac{\partial(U_1 - U_0)}{\partial \tau}}{\gamma(\theta^*)\dfrac{\partial(U_1 - U_0)}{\partial \mu} - c'(\theta^*)} > 0.$$

The denominator is positive by (32), $\gamma(\theta^*) > 0$, and $c'(\theta^*) < 0$. It follows that

$$\frac{d[1 - \Gamma(\theta^*)]}{d\tau} = -\gamma(\theta^*)\frac{d\theta^*}{d\tau} < 0.$$

□

A.5 Proof of Lemma 4

The matching is defined by the ODE (where M_z denotes the derivative of M wrt to the first (only) argument z)

$$M_z(z;\tau) = h(1-z)\frac{g(z;\tau)}{g(M(z;\tau);\tau)} \tag{35}$$

with boundary conditions $M(0;\tau) = z_2(\tau)$ and $M(z_1(\tau);\tau) = 1$. Differentiating (35) with respect to h, we get:[30]

$$M_{hz}(z;h) = (1-z)g(z)\left[\frac{g\big(M(z;h)\big) - hg_z\big(M(z;h)\big)M_h(z;h)}{g\big(M(z;h)\big)^2}\right]$$

[30]Appendix B.5 shows that the cutoffs are differentiable with respect to a parameter. Since $h(1-z)g(z)/g(m)$ is continuously differentiable in (z,m,h), Teschl (2012, Theorem 2.11) justifies differentiating the matching ODE.

Let $f(z;h) := M_h(z;h)$. Then,

$$f_z(z;h) = \frac{(1-z)g(z)}{g\big(M(z;h)\big)} - \frac{h(1-z)g(z)}{g\big(M(z;h)\big)}\frac{g_z\big(M(z;h)\big)}{g\big(M(z;h)\big)}f(z;h)$$

In ODE form:

$$f'(z;h) + a(z;h)f(z;h) = b(z;h)$$

Where

$$a(z;h) = \frac{h(1-z)g(z)}{g\big(M(z;h)\big)}\frac{g_z\big(M(z;h)\big)}{g\big(M(z;h)\big)} = M'(z;h)\frac{g_z\big(M(z;h)\big)}{g\big(M(z;h)\big)}$$
$$b(z;h) = \frac{(1-z)g(z)}{g\big(M(z;h)\big)} = \frac{M'(z;h)}{h} > 0$$

Let $\mu(z;h) = e^{\int_0^z a(u;h)du} = \frac{g(M(z;h))}{g(z_2)}$. Solving the ODE, using the boundary condition $f(0;h) = \frac{dz_2}{dh}$ yields:[31]

$$f(z;h) = \frac{\frac{dz_2}{dh} + \int_0^z \mu(t;h)b(t;h)dt}{\mu(z;h)} \tag{36}$$

From (49) at z_1 together with the indifference condition $w(z_1) = z_1$ we get

$$\frac{z_1}{1-z_1} = (1-h)z_2 + \int_0^{z_1} \frac{M(t)}{(1-t)^2}dt.$$

Differentiating we get

$$\cancel{\frac{1}{(1-z_1)^2}\frac{dz_1}{dh}} = -z_2(h) + (1-h)\frac{dz_2}{dh} + \cancel{\frac{M(z_1;h)}{(1-z_1)^2}\frac{dz_1}{dh}} + \int_0^{z_1(h)} \frac{f(t;h)}{(1-t)^2}dt$$
$$z_2(h) = (1-h)\frac{dz_2}{dh} + \int_0^{z_1(h)} \frac{f(t;h)}{(1-t)^2}dt$$

Substituting back the solution to $f(z;h)$ from (36):

$$z_2(h) = (1-h)\frac{dz_2}{dh} + \int_0^{z_1(h)} \frac{\frac{dz_2}{dh} + \int_0^u \mu(t;h)b(t;h)dt}{\mu(u;h)(1-u)^2}du$$

Let $A := \int_0^{z_1(h)} \frac{1}{\mu(u;h)(1-u)^2}du = g(z_2(h))\int_0^{z_1(h)} \frac{1}{g(M(u;h))(1-u)^2}du$ and $B := \int_0^{z_1(h)} \frac{\int_0^u \mu(t;h)b(t;h)dt}{\mu(u;h)(1-u)^2}du$.

[31] See proof of Lemma 10. Constants on μ cancel, so $\mu(0) = 1$.

Then:

$$\frac{dz_2}{dh} = \frac{z_2 - B}{1 - h + A}. \tag{37}$$

Notice that the denominator is positive, so the sign of $\frac{dz_2}{dh}$ is dependent on the sign of the numerator. To help sign it, let us write B fully:

$$\begin{aligned}
B &= \int_0^{z_1(h)} \frac{\int_0^x \frac{g(M(t;h))}{g(z_2)} b(t;h)dt}{(1-x)^2 \frac{g(M(x;h))}{g(z_2)}} dx \\
&= \int_0^{z_1(h)} \frac{\int_0^x g(M(t;h))b(t;h)dt}{(1-x)^2 g(M(x;h))} dx \\
&= \int_0^{z_1(h)} \frac{\int_0^x g(M(t;h)) \frac{(1-t)g(t)}{g\left(M(t;h)\right)} dt}{(1-x)^2 g(M(x;h))} dx \\
&= \int_0^{z_1(h)} \frac{\int_0^x (1-t)g(t)dt}{(1-x)^2 g(M(x;h))} dx \\
&= \frac{1}{h} \int_0^{z_1(h)} \frac{G(M(x;h)) - G(z_2(h))}{(1-x)^2 g(M(x;h))} dx
\end{aligned}$$

Convexity of G implies, for every $0 \le x < y \le 1$,

$$G(y) - G(x) \le (y-x)g(y) \quad \Longrightarrow \quad \frac{G(M(z)) - G(z_2)}{g(M(z))} \le M(z) - z_2.$$

Therefore,

$$\frac{1}{h} \int_0^{z_1(h)} \frac{G(M(x;h)) - G(z_2(h))}{(1-x)^2 g(M(x;h))} dx \;\le\; \frac{1}{h} \int_0^{z_1} \frac{M(x;h) - z_2(h)}{(1-x)^2} dx. \tag{38}$$

To finalize, notice that

$$\frac{d}{dz}\left(\frac{M(z) - z_2}{1-z}\right) = \frac{M'(z)}{1-z} + \frac{M(z) - z_2}{(1-z)^2}.$$

Hence,

$$\begin{aligned}\frac{1}{h}\int_0^{z_1}\frac{M(z)-z_2}{(1-z)^2}\,dz &= \frac{1}{h}\left[\frac{M(z)-z_2}{1-z}\right]_0^{z_1} - \frac{1}{h}\int_0^{z_1}\underbrace{\frac{M'(z)}{1-z}}_{w''(z)}\,dz \\ &= \frac{1}{h}\left[\frac{1-z_2}{1-z_1} - (1-hz_2)\right] \\ &= z_2 - \frac{z_2-z_1}{h(1-z_1)} < z_2, \end{aligned} \tag{39}$$

because $z_1 < z_2$. Returning to the numerator of (37), using (38) and (39),

$$\begin{aligned} z_2 - B &= z_2 - \frac{1}{h}\int_0^{z_1(h)}\frac{G(M(x;h)) - G(z_2(h))}{(1-x)^2 g(M(x;h))}dx \\ &\geq z_2 - \frac{1}{h}\int_0^{z_1}\frac{M(x;h)-z_2(h)}{(1-x)^2}dx. \\ &= \cancel{z_2} - \cancel{z_2} + \frac{z_2-z_1}{h(1-z_1)} \\ &> 0 \end{aligned}$$

Consequently, convexity of G is sufficient for z_2 to increase with h.

A.6 Proof of Proposition 5

Proof. For this proof, recall that the knowledge distribution is held fixed at $G(z;\theta^*)$. To ease notation we omit the dependence on θ^*. By Lemma 2 in Fuchs, Garicano, and Rayo (2015), z_1 increases with lower h. By Lemma 4, z_2 decreases when communication costs go down.

To show that the matching function decreases consider $h' > h$, and let $\Psi(z) := M(z;h') - M(z;h)$. Since $z_2(h') > z_2(h)$, $\Psi(0) > 0$. Similarly, $\Psi(z_1(h')) = 1 - M(z_1(h');h) > 0$, because for h the worker matched to $\bar{z}$ is $z_1(h) > z_1(h')$. Suppose towards a contradiction that there exists and $t \in [0, z_1(h')]$ such that $\Psi(t) = 0$, then $M(t;h) = M(t;h')$. The slope at such t is

$$\begin{aligned}\Psi'(t) &= M_z(t;h') - M_z(t;h) \\ &= h'(1-t)\frac{g(t)}{g(M(t;h'))} - h(1-t)\frac{g(t)}{g(M(t;h))} \\ &= (1-t)\frac{g(t)}{g(M(t;h'))}\big(h'-h\big) > 0\end{aligned}$$

Where the last equality comes from the assumption that $\Psi(t) = 0$. Since both endpoints are positive and any crossing of zero must have strictly positive slope, $M(z;h) < M(z;h')$ for all $z \in [0, z_1(h')]$.

Similarly, let $\Phi(z) = w(z;h') - w(z;h)$. Since z_1 is decreasing in h, $\Phi(z_1(h')) = z_1(h') - w(z_1(h');h) < 0$. Let $t \in [0, z_1(h')]$ be such that $\Phi(t) = 0$, that is $w(t;h') = w(t;h)$. The slope of Φ at such point is

$$\begin{aligned}\Phi'(t) &= w'(t;h') - w'(t;h) \\ &= \frac{M(t;h') - w(t;h')}{1-t} - \frac{M(t;h) - w(t;h)}{1-t} \\ &= \frac{M(t;h') - M(t;h)}{1-t} > 0.\end{aligned}$$

Where the last equality comes form the assumption that $\Phi(t) = 0$ and the last inequality from the result above that the matching decreases. Thus proving that wages increase.

Finally, consider the rents of managers. Let $\chi(z) = R(z;h') - R(z;h)$. $\chi(z_2(h')) = z_2(h') - R(z_2(h');h) < 0$ because $z_2(h) < z_2(h')$ and rents are strictly increasing. Suppose there exist a $t \in [z_2(h'), \bar{z}]$ such that $\chi(t) = 0$. Similarly, the slope of χ at t is

$$\begin{aligned}\chi'(t) &= R'(t;h') - R'(t;h) \\ &= \frac{1}{h'(1 - Z(t;h'))} - \frac{1}{h(1 - Z(t;h))} < 0\end{aligned}$$

Where the inequality comes from $h' > h$ and, since $M(z;h') > M(z;h)$, $Z(m;h') < Z(m;h)$. □

A.7 Proof of Proposition 6

Proof. We first establish a lower bound on the earnings gain at low knowledge levels and an upper bound on the population's average gain.

The earnings gain at low knowledge levels. The matching and wage equations imply

$$w''(z;h) = h\frac{g(z)}{g(M(z;h))}, \qquad w'(0;h) = hz_2, \qquad w'(z_1;h) = 1.$$

Integrating this expression and the matching equation gives

$$\frac{1}{h} = z_2 + \int_0^{z_1} \frac{g(t)}{g(M(t;h))}\,dt, \qquad \frac{1 - z_2}{h} = \int_0^{z_1} (1-t)\frac{g(t)}{g(M(t;h))}\,dt. \tag{40}$$

Combining the two identities yields

$$z_2(1-z_2) = \int_0^{z_1} (z_2 - t)\frac{g(t)}{g(M(t;h))}\,dt.$$

Since g is nondecreasing and $z_1 < z_2$, it follows that

$$z_2(1-z_2) \le z_2 z_1 - \frac{z_1^2}{2} \le \frac{z_2^2}{2}.$$

Moreover,

$$\frac{1}{h} \le z_2 + z_1 \le 2z_2, \qquad \frac{1}{h} \le 1 + z_1.$$

Consequently,

$$z_2 \ge \frac{2}{3}, \qquad hz_2 \ge \frac{1}{2}, \qquad z_1 \ge \frac{1-h}{h}. \tag{41}$$

Differentiating time clearing at fixed investment gives

$$\frac{\partial M(z;h)}{\partial h} = \frac{g(z_2)\dfrac{\partial z_2}{\partial h} + \displaystyle\int_0^z (1-t)g(t)\,dt}{g(M(z;h))}. \tag{42}$$

The wage boundary conditions imply

$$\frac{z_1}{1-z_1} = (1-h)z_2 + \int_0^{z_1} \frac{M(t;h)}{(1-t)^2}\,dt.$$

Differentiating and using $M(z_1;h) = 1$ cancels the terms involving $\partial z_1/\partial h$:

$$z_2 = (1-h)\frac{\partial z_2}{\partial h} + \int_0^{z_1} \frac{\partial M(t;h)/\partial h}{(1-t)^2}\,dt.$$

Substitution from (42) therefore gives

$$0 < \frac{\partial z_2}{\partial h} \le \frac{z_2}{1-h+g(z_2)\displaystyle\int_0^{z_1} \frac{dt}{(1-t)^2 g(M(t;h))}}. \tag{43}$$

Monotonicity of g, Jensen's inequality, and (40) imply

$$\begin{aligned} g(z_2)\int_0^{z_1}\frac{dt}{(1-t)^2 g(M(t;h))} &\geq \int_0^{z_1}\frac{g(t)}{(1-t)^2 g(M(t;h))}\,dt \\ &\geq \frac{\left[\int_0^{z_1}\dfrac{g(t)}{g(M(t;h))}\,dt\right]^3}{\left[\int_0^{z_1}(1-t)\dfrac{g(t)}{g(M(t;h))}\,dt\right]^2} \\ &= \frac{(1-hz_2)^3}{h(1-z_2)^2}. \end{aligned}$$

The arithmetic-geometric mean inequality gives

$$\begin{aligned} (1-hz_2)^3 &= \big[(1-z_2)+(1-h)z_2\big]^3 \\ &\geq \frac{27}{4}(1-z_2)^2(1-h)z_2. \end{aligned}$$

In addition, (41) gives

$$\begin{aligned} \frac{z_2[(1+h)z_2-1]}{h} &= z_2^2 - \frac{z_2(1-z_2)}{h} \\ &\geq z_2^2(2z_2-1) \\ &\geq \frac{4}{27}. \end{aligned}$$

In particular, $(1+h)z_2 - 1 > 0$. Combining these inequalities yields

$$g(z_2)\int_0^{z_1}\frac{dt}{(1-t)^2 g(M(t;h))} \geq \frac{27z_2(1-h)}{4h} \geq \frac{1-h}{(1+h)z_2-1}.$$

Substitution into (43) gives

$$(1-h)\frac{\partial z_2}{\partial h} \leq z_2 - \frac{1}{1+h}.$$

Since $w(0;h) = (1-h)z_2$, we have established

$$-\frac{\partial w(0;h)}{\partial h} = z_2 - (1-h)\frac{\partial z_2}{\partial h} \geq \frac{1}{1+h}. \tag{44}$$

To extend this bound to other worker knowledge levels, monotonicity of g and (42) imply

$$0 < \frac{\partial M(t;h)}{\partial h} \leq \frac{\partial z_2}{\partial h} + 1, \qquad 0 \leq t \leq z_1.$$

Differentiating the wage formula with its boundary condition at zero then gives

$$\begin{aligned}-\frac{\partial w(z;h)}{\partial h} &= (1-z)\left[-\frac{\partial w(0;h)}{\partial h} - \int_0^z \frac{\partial M(t;h)/\partial h}{(1-t)^2}\,dt\right] \\ &\geq (1-z)\left[-\frac{\partial w(0;h)}{\partial h}\right] - z\left(\frac{\partial z_2}{\partial h}+1\right) \\ &= -\frac{\partial w(0;h)}{\partial h} - z\left(z_2 + h\frac{\partial z_2}{\partial h}+1\right).\end{aligned}$$

Equation (43) also implies $\partial z_2/\partial h \leq z_2/(1-h)$. Hence

$$z_2 + h\frac{\partial z_2}{\partial h} + 1 \leq \frac{1}{1-h} + 1 \leq \frac{1}{h(1-h)}.$$

Together with (44), this establishes

$$-\frac{\partial w(z;h)}{\partial h} \geq \frac{1}{1+h} - \frac{z}{h(1-h)}, \qquad 0 \leq z \leq z_1.$$

By (41),

$$z_1 \geq \frac{1-h}{h} > \frac{h(1-h)}{1+h}.$$

Thus, the affine expression on the right is negative for $z \geq z_1$. Self-employed earnings do not change with h, and managerial earnings increase when h decreases. Consequently,

$$-\frac{\partial Y(z;h)}{\partial h} \geq \frac{1}{1+h} - \frac{z}{h(1-h)} \tag{45}$$

for almost every $z \in [0,1]$.

The population's average earnings gain. The envelope condition for a manager gives

$$\frac{\partial R(m;h)}{\partial h} = -\frac{R(m;h)}{h} - \frac{1}{h[1-Z(m;h)]}\left.\frac{\partial w(z;h)}{\partial h}\right|_{z=Z(m;h)}.$$

Integrating across managers and changing variables through the matching function cancels the wage-adjustment term against the aggregate change in worker earnings. Since

$R(M(z;h);h) = w'(z;h)/h$, the resulting identity is

$$\begin{aligned} -\int_0^1 \frac{\partial Y(z;h)}{\partial h} g(z)\,dz &= \frac{1}{h}\int_{z_2}^1 R(m;h)g(m)\,dm \\ &= \frac{1}{h}\int_0^{z_1} (1-z)w'(z;h)g(z)\,dz. \end{aligned}$$

The wage function is convex, and $w'(z_1;h) = 1$, so $w'(z;h) \leq 1$ for workers. Moreover, G first-order stochastically dominates G_0. Therefore,

$$\begin{aligned} -\int_0^1 \frac{\partial Y(z;h)}{\partial h} g(z)\,dz &\leq \frac{1}{h}\int_0^1 (1-z)g(z)\,dz \\ &\leq \frac{1}{h}\int_0^1 (1-z)g_0(z)\,dz. \end{aligned} \tag{46}$$

The change in the education premium. Likelihood-ratio dominance and the fact that the distributions differ imply $\ell(1) > 1$. Moreover,

$$\frac{\ell(1)g_0(z) - g_1(z)}{\ell(1)-1} \geq 0, \qquad \int_0^1 \frac{\ell(1)g_0(z) - g_1(z)}{\ell(1)-1}\,dz = 1.$$

Integrating (45) against the downside component and applying (9) yields

$$\begin{aligned} &-\int_0^1 \frac{\partial Y(z;h)}{\partial h} \frac{\ell(1)g_0(z) - g_1(z)}{\ell(1)-1}\,dz \\ &\quad\geq \frac{1}{1+h} - \frac{\displaystyle\int_0^1 z[\ell(1)g_0(z) - g_1(z)]\,dz}{h(1-h)[\ell(1)-1]} \\ &\quad> \frac{1}{h}\int_0^1 (1-z)g_0(z)\,dz \\ &\quad\geq -\int_0^1 \frac{\partial Y(z;h)}{\partial h} g(z)\,dz. \end{aligned}$$

The last inequality is (46). Thus, the downside component benefits more from cheaper help than the population average.

Using the population mixture $g(z) = \Gamma(\theta^*)g_0(z) + [1-\Gamma(\theta^*)]g_1(z)$, the preceding strict

inequality becomes

$$\begin{aligned} 0 &< -\int_0^1 \frac{\partial Y(z;h)}{\partial h}\left[\frac{\ell(1)g_0(z)-g_1(z)}{\ell(1)-1}-g(z)\right]dz \\ &= \left[\frac{\ell(1)}{\ell(1)-1}-\Gamma(\theta^*)\right]\frac{\partial(U_1-U_0)}{\partial h}. \end{aligned}$$

The coefficient in brackets is strictly positive. Consequently,

$$\frac{\partial(U_1-U_0)}{\partial h} > 0.$$

Differentiation under the earnings integrals introduces no occupational boundary terms: earnings are continuous at both cutoffs.

The equilibrium investment response. The equilibrium condition is

$$U_1 - U_0 = c(\theta^*).$$

The education premium is nondecreasing in the investment threshold, whereas $c'(\theta^*) < 0$. Therefore,

$$\frac{d\theta^*}{dh} = -\frac{\dfrac{\partial(U_1-U_0)}{\partial h}}{\dfrac{\partial(U_1-U_0)}{\partial \theta^*} - c'(\theta^*)} < 0.$$

A marginal decrease in h therefore raises θ^* and reduces the mass of investors, $1-\Gamma(\theta^*)$. □

B Mathematical Appendix

B.1 Problems are uniformly distributed without loss of generality

Likelihood-ratio order is preserved under increasing transformations (see Theorem 1.C.8 in Shaked and Shanthikumar (2007)). Normalizing to uniform implies taking the inverse of the cdf, which is increasing.

For completeness, let $u \sim P(u)$ be the random variable for problems, where $P(u)$ is arbitrary (to make things easier continuous and strictly increasing), and let $Q_0(v)$ and $Q_1(v)$ be the non-investment and investment distributions, respectively.

Consider the random variable $x = P(u)$, then $x \sim U[0,1]$. An agent with knowledge v could solve a proportion $P(v)$ of problems. Then, for a problem $x \sim U[0,1]$ the mass (for each distribution) that cannot solve it is $\mathbb{P}(P(v) \leq x) = \mathbb{P}(v \leq P^{-1}(x)) = Q_i(P^{-1}(x))$.

Let $G_i(z) := Q_i(P^{-1}(z))$, then $g_i(z) = q_i(P^{-1}(z))(P^{-1})'(z)$ which means that $g_i(z) = \frac{q_i(P^{-1}(z))}{p(P^{-1}(z))}$.

If we take the ratio, $\frac{g_1(z)}{g_0(z)} = \frac{\frac{q_1(P^{-1}(z))}{p(P^{-1}(z))}}{\frac{q_0(P^{-1}(z))}{p(P^{-1}(z))}} = \frac{q_1(P^{-1}(z))}{q_0(P^{-1}(z))}$. Since P^{-1} is increasing, if $Q_1 \succ_{LR} Q_0$ so would $G_1 \succ_{LR} G_0$.

B.2 Managers do not prefer to mix workers

Suppose a manager with knowledge m were to hire n_a workers of knowledge z_a and n_b with knowledge z_b. The rent of the manager would be

$$R(m, z_a, z_b) = n_a(m - w(z_a)) + n_b(m - w(z_b))$$

and the efficient use of time would mean

$$hn_a(1 - z_a) + hn_b(1 - z_b) = 1.$$

Let $s_a = hn_a(1 - z_a)$ be the share of time allocated to problems from workers z_a (analogously s_b, which additionally $s_b = 1 - s_a$). Then $n_a = \frac{s_a}{h(1-z_a)}$ We can rewrite the rent as

$$\begin{aligned} R(m, z_a, z_b) &= s_a \frac{m - w(z_a)}{h(1 - z_a)} + (1 - s_a)\frac{m - w(z_b)}{h(1 - z_b)} \\ &= s_a R(m, z_a) + (1 - s_a)R(m, z_b) \end{aligned}$$

Hence, hiring two types of workers yields rents that are a convex combination of the

rents of hiring each knowledge-type. Consequently, $R(m, z_a, z_b) \leq \max\Big\{R(m, z_a); R(m, z_b)\Big\}$, and so, the manager never strictly prefers to mix workers.

B.3 Equilibrium definition

Definition 2 (Competitive equilibrium)**.** A competitive equilibrium consists of an investment profile $\iota : \Theta \longrightarrow \{0, 1\}$, occupational sets $\mathcal{W}, \mathcal{S}, \mathcal{M} \subseteq [0, \bar{z}]$, a matching function $M : \mathcal{W} \to \mathcal{M}$, and an earnings schedule $Y : [0, \bar{z}] \to \mathbb{R}$, such that[32]:

1. *Investment consistency.* Taking the earnings schedule $Y(\cdot)$ as given, investment maximizes expected earnings net of education costs. That is, for Γ-almost every $\theta \in \Theta$:

$$\iota(\theta) \in \underset{i \in \{0,1\}}{\arg\max} \left\{ \int_0^{\bar{z}} Y(t)\, dG_i(t) - i\, c(\theta) \right\};$$

2. *Optimal occupational choice.* Taking the earnings schedule $Y(\cdot)$ and the matching function $M(\cdot)$ as given, agents sort into occupations optimally.
   Agents with knowledge $z \in \mathcal{W}$ weakly prefer working to self-employment

$$Y(z) \geq z,$$

   and to managing a team

$$Y(z) \geq \frac{z - Y(z')}{h(1 - z')} \quad \text{for every } 0 \leq z' < z.$$

   Agents with knowledge $z \in \mathcal{S}$ weakly prefer self-employment to working for a manager

$$z \geq m - h(1 - z)Y(m) \quad \text{for every } z < m \leq \bar{z},$$

   and to managing a team

$$z \geq \frac{z - Y(z')}{h(1 - z')} \quad \text{for every } 0 \leq z' < z.$$

   Agents with knowledge $z \in \mathcal{M}$ weakly prefer managing to self-employment

$$Y(z) \geq z,$$

[32] All sets and functions appearing in the definition are measurable.

and to working for a manager

$$Y(z) \geq m - h(1-z)Y(m) \quad \text{for every } z < m \leq \bar{z}.$$

3. *Matching is optimal.* Taking the earnings as given, for each active manager, the matched worker knowledge attains maximal rent, and managerial earnings equal that rent.

$$Y(M(z)) = \frac{M(z) - Y(z)}{h(1-z)} \quad \text{for } G(\cdot;\iota)\text{-almost every } z \in \mathcal{W}.$$

4. *Market clearing.* The occupational regions cover the knowledge domain:

$$\mathcal{W} \cup \mathcal{S} \cup \mathcal{M} = [0, \bar{z}],$$

and the intersections of $\mathcal{W}$, $\mathcal{S}$, and $\mathcal{M}$ have $G(\cdot;\iota)$-measure zero.

5. *Managerial time clears.* For every measurable set of workers, the time demanded for assistance equals the time supplied by their matched managers. That is, for every measurable set $A \subseteq \mathcal{W}$,

$$\int_A h(1-z)\, dG(z;\iota) = \int_{M(A)} dG(m;\iota),$$

where $M(A) \subseteq \mathcal{M}$ denotes managers matched to workers in A, and $M(\mathcal{W}) = \mathcal{M}$.

B.4 Solving the Wage ODE

Take the equilibrium condition on the wage, $w'(z) = \frac{M(z)-w(z)}{1-z}$. Rearranging we get

$$\begin{aligned} w'(z) + \frac{w(z)}{1-z} &= \frac{M(z)}{1-z} \\ \frac{w'(z)}{1-z} + \frac{w(z)}{(1-z)^2} &= \frac{M(z)}{(1-z)^2} \\ \frac{\partial}{\partial z}\left(\frac{w(z)}{1-z}\right) &= \frac{M(z)}{(1-z)^2} \end{aligned}$$

Integrating

$$\frac{w(z)}{1-z} - \frac{w(0)}{1-0} = \int_0^z \frac{M(t)}{(1-t)^2} dt$$

Then, depending on the chosen boundary condition

$$w(0) = \frac{z_1}{1-z_1} - \int_0^{z_1} \frac{M(t)}{(1-t)^2} dt \tag{47}$$

$$w(z) = (1-z)\left[\frac{z_1}{1-z_1} - \int_z^{z_1} \frac{M(t)}{(1-t)^2} dt\right] \qquad \text{if } w(z_1) = z_1 \tag{48}$$

$$w(z) = (1-z)\left[(1-h)z_2 + \int_0^{z} \frac{M(t)}{(1-t)^2} dt\right] \qquad \text{if } w(0) = (1-h)z_2 \tag{49}$$

B.5 Differentiability of the occupational cutoffs

The differentiability of $z_1(\tau)$ and $z_2(\tau)$ can be established directly by reducing the equilibrium conditions to a single scalar equation and applying the implicit-function theorem.

Lemma 14. *Suppose that $g(z;\tau)$ is continuously differentiable jointly in z and τ, that*

$$g(z;\tau) > 0$$

on a neighborhood of $[0, \bar{z}]$, and that $G(\bar{z};\tau) = 1$ for every τ. Suppose also that, at the parameter value under consideration, the equilibrium is interior:

$$0 < z_1(\tau) < z_2(\tau) < \bar{z} \leq 1.$$

Then $z_1(\tau)$ and $z_2(\tau)$ are continuously differentiable with respect to τ.

Proof. We will apply the implicit-function theorem. For now consider z_1 as a number, and not the equilibrium function $z_1(\tau)$.

For any $t \in [0, z_1]$, time clearing between the workers in $[t, z_1]$ and the managers in $[M(t;\tau), \bar{z}]$ requires

$$h \int_t^{z_1} (1-s) g(s;\tau)\, ds = \int_{M(t;\tau)}^{\bar{z}} g(s;\tau)\, ds.$$

Since $G(\bar{z};\tau) = 1$, this can be written as

$$G(M(t;\tau);\tau) = 1 - h \int_t^{z_1} (1-s) g(s;\tau)\, ds.$$

Consequently,

$$M(t;\tau) = G^{-1}\left(1 - h \int_t^{z_1} (1-s) g(s;\tau)\, ds; \tau\right).$$

Because

$$\frac{\partial G(z;\tau)}{\partial z} = g(z;\tau) > 0,$$

the inverse-function theorem implies that $M(t;\tau)$ is continuously differentiable with respect to the free value of z_1 and with respect to τ.

Differentiating the integrated time-clearing condition with respect to z_1, holding t and τ fixed, gives

$$g(M(t;\tau);\tau)\frac{\partial M(t;\tau)}{\partial z_1} = -h(1-z_1)g(z_1;\tau).$$

Therefore,

$$\frac{\partial M(t;\tau)}{\partial z_1} = -\frac{h(1-z_1)g(z_1;\tau)}{g(M(t;\tau);\tau)} < 0. \tag{50}$$

The wage equation is

$$w'(z;\tau) = \frac{M(z;\tau) - w(z;\tau)}{1-z},$$

with boundary condition

$$w(0;\tau) = (1-h)z_2.$$

Solving this linear differential equation gives

$$w(z;\tau) = (1-z)\left[(1-h)z_2 + \int_0^z \frac{M(t;\tau)}{(1-t)^2}\,dt\right].$$

Since

$$z_2 = M(0;\tau)$$

and the marginal worker satisfies

$$w(z_1;\tau) = z_1,$$

the equilibrium value of z_1 must satisfy

$$\frac{z_1}{1-z_1} - (1-h)M(0;\tau) - \int_0^{z_1} \frac{M(t;\tau)}{(1-t)^2}\,dt = 0. \tag{51}$$

Here, $M(t;\tau)$ is the matching function generated by the free value of z_1.

The left-hand side of (51) is continuously differentiable with respect to (z_1,τ). Differentiating it with respect to z_1 gives

$$\frac{1}{(1-z_1)^2} - (1-h)\frac{\partial M(0;\tau)}{\partial z_1} - \frac{M(z_1;\tau)}{(1-z_1)^2} - \int_0^{z_1} \frac{1}{(1-t)^2}\frac{\partial M(t;\tau)}{\partial z_1}\,dt.$$

Using

$$M(z_1;\tau) = \bar{z},$$

this becomes

$$\frac{1-\bar{z}}{(1-z_1)^2} - (1-h)\frac{\partial M(0;\tau)}{\partial z_1} - \int_0^{z_1} \frac{1}{(1-t)^2}\frac{\partial M(t;\tau)}{\partial z_1}\,dt.$$

Substituting (50), and using $M(0;\tau) = z_2$, yields

$$\frac{1-\bar{z}}{(1-z_1)^2} + h(1-z_1)g(z_1;\tau)\left[\frac{1-h}{g(z_2;\tau)} + \int_0^{z_1} \frac{dt}{(1-t)^2 g(M(t;\tau);\tau)}\right] > 0. \tag{52}$$

The inequality is strict because

$$h \in (0,1), \qquad z_1 < 1, \qquad g(z;\tau) > 0.$$

In particular, when $\bar{z} = 1$, the first term in (52) is zero, but all the remaining terms are strictly positive.

Thus, the derivative with respect to z_1 of the left-hand side of (51) is nonzero. The scalar implicit-function theorem therefore implies that the equilibrium cutoff

$$z_1 = z_1(\tau)$$

is continuously differentiable with respect to τ.

Finally, at $t = 0$, the integrated time-clearing condition is

$$G(z_2(\tau);\tau) = 1 - h\int_0^{z_1(\tau)} (1-s)g(s;\tau)\,ds. \tag{53}$$

The derivative of the left-hand side of (53) with respect to z_2 is

$$g(z_2(\tau);\tau) > 0.$$

Since $z_1(\tau)$ is continuously differentiable, a second application of the implicit-function theorem to (53) implies that

$$z_2 = z_2(\tau)$$

is continuously differentiable with respect to τ.

Therefore,

$$z_1(\tau), z_2(\tau) \in C^1$$

at every parameter value at which the equilibrium remains interior and $g(z;\tau)$ remains strictly positive and continuously differentiable.

If the interim equilibrium is unique and remains interior throughout an open interval of values of τ, the local continuously differentiable solutions coincide and hence $z_1(\tau)$ and $z_2(\tau)$ are continuously differentiable on that entire interval. □

B.6 Positive Mass of Worker Gain when h decreases investment

Proof. A marginal decrease in h that reduces investment satisfies $d\theta^*/dh < 0$ (i.e., more agents invest when h increases), and so $\frac{d(U_1-U_0)}{dh} > 0$

Differentiating expected earnings, the moving-cutoff terms cancel because earnings are continuous across occupations. Applying the managerial envelope condition and changing variables through matching (as in the proof of Corollary 2) yields

$$\begin{aligned}\frac{d(U_1-U_0)}{dh} = &-\int_0^{z_1} \frac{dw(z)}{dh}\left[\frac{g_1(M(z))-g_0(M(z))}{g(M(z))} - \frac{g_1(z)-g_0(z)}{g(z)}\right] g(z)\,dz \\ &-\frac{1}{h}\int_{z_2}^{\bar{z}} R(m)[g_1(m)-g_0(m)]\,dm.\end{aligned}$$

The bracket is nonnegative: likelihood-ratio dominance implies that

$$\frac{g_1(z)-g_0(z)}{g(z)} = \frac{\ell(z)-1}{\Gamma(\theta^*)+[1-\Gamma(\theta^*)]\ell(z)}$$

is nondecreasing in z, and $M(z) > z$. Moreover, integration by parts gives

$$\begin{aligned}\int_{z_2}^{\bar{z}} R(m)[g_1(m)-g_0(m)]\,dm = &\, R(z_2)[G_0(z_2)-G_1(z_2)] \\ &+\int_{z_2}^{\bar{z}} R'(m)[G_0(m)-G_1(m)]\,dm > 0.\end{aligned}$$

Suppose no interior incumbent worker gains, so that $dw(z)/dh \geq 0$ for every $z \in (0, z_1)$. The first term in the premium derivative is then nonpositive and the second is strictly negative, implying

$$\frac{d(U_1-U_0)}{dh} < 0,$$

a contradiction.

Thus, $dw(z)/dh < 0$ at some interior worker knowledge level. By continuity, the inequality holds on a nondegenerate interval strictly inside $(0, z_1)$. For a sufficiently small decrease in h, these knowledge levels remain workers and receive strictly higher wages. Strict positivity of the knowledge density implies that they have positive mass. □

B.7 Positive Mass of Worker Gain when LR-better G_1 decreases investment

Proof. Consider a marginal increase in π that reduces investment, so that $d\theta^*/d\pi > 0$.

Differentiating investment indifference and applying the rent-envelope and matching equations gives

$$c'(\theta^*)\frac{d\theta^*}{d\pi} = \int_0^{\bar{z}} Y(t;\theta^*,\pi)\frac{\partial g_1(t;\pi)}{\partial \pi}\,dt + \int_0^{z_1(\theta^*,\pi)} \frac{d\,w(t;\theta^*,\pi)}{d\pi}\,g_0(t)\frac{\ell(t;\pi)-\ell(M(t;\theta^*,\pi);\pi)}{\Gamma(\theta^*)+[1-\Gamma(\theta^*)]\ell(M(t;\theta^*,\pi);\pi)}\,dt.$$

The occupational boundary terms cancel because earnings are continuous.

The left-hand side is strictly negative. The first integral is nonnegative because the improvement in G_1 raises expected earnings at an unchanged, increasing earnings schedule. Consequently, the second integral must be strictly negative.

Since $\ell(\cdot;\pi)$ is nondecreasing and $M(t;\theta^*,\pi) > t$, the fraction in the second integral is nonpositive. Thus $d\,w(t;\theta^*,\pi)/d\pi$ must be strictly positive for a positive mass of workers. By continuity, a positive mass of incumbent workers remain workers and strictly gain from a sufficiently small increase in π. □

References

Antràs, Pol, Luis Garicano, and Esteban Rossi-Hansberg (2006). "Offshoring in a Knowledge Economy." In: *The Quarterly Journal of Economics* 121.1, pp. 31–77.

Bhaskar, Venkataraman and Ed Hopkins (2016). "Marriage as a rat race: Noisy premarital investments with assortative matching." In: *Journal of Political Economy* 124.4, pp. 992–1045.

Caicedo, Santiago, Robert E Lucas Jr, and Esteban Rossi-Hansberg (2019). "Learning, Career Paths, and the Distribution of Wages." In: *American Economic Journal: Macroeconomics* 11.1, pp. 49–88.

Caliendo, Lorenzo, Giordano Mion, Luca David Opromolla, and Esteban Rossi-Hansberg (2020). "Productivity and Organization in Portuguese Firms." In: *Journal of Political Economy* 128.11, pp. 4211–4257.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg (2015). "The Anatomy of French Production Hierarchies." In: *Journal of Political Economy* 123.4, pp. 809–852.

Caliendo, Lorenzo and Esteban Rossi-Hansberg (2012). "The Impact of Trade on Organization and Productivity." In: *The Quarterly Journal of Economics* 127.3, pp. 1393–1467.

Carmona, Guilherme, Krittanai Laohakunakorn, and Hitoshi Tsujiyama (Mar. 3, 2026). "Knowledge Economies." Working paper. URL: https://klaohakunakorn.com/tu260216.pdf.

Chade, Hector and Ilse Lindenlaub (2022). "Risky Matching." In: *The Review of Economic Studies* 89.2, pp. 626–665.

Cole, Harold L, George J Mailath, and Andrew Postlewaite (2001). "Efficient non-contractible investments in large economies." In: *Journal of Economic Theory* 101.2, pp. 333–373.

Fuchs, William and Luis Garicano (2010). "Matching Problems with Expertise in Firms and Markets." In: *Journal of the European Economic Association* 8.2-3, pp. 354–364.

Fuchs, William, Luis Garicano, and Luis Rayo (2015). "Optimal Contracting and the Organization of Knowledge." In: *The Review of Economic Studies* 82.2, pp. 632–658.

Garicano, Luis (2000). "Hierarchies and the Organization of Knowledge in Production." In: *Journal of Political Economy* 108.5, pp. 874–904.

Garicano, Luis and Thomas N Hubbard (2005). "Hierarchical sorting and learning costs: Theory and evidence from the law." In: *Journal of Economic Behavior & Organization* 58.2, pp. 349–369.

Garicano, Luis and Esteban Rossi-Hansberg (2004). "Inequality and the Organization of Knowledge." In: *American Economic Review* 94.2, pp. 197–202.

Garicano, Luis and Esteban Rossi-Hansberg (2006). "Organization and Inequality in a Knowledge Economy." In: *The Quarterly Journal of Economics* 121.4, pp. 1383–1435.

— (2012). "Organizing Growth." In: *Journal of Economic Theory* 147.2, pp. 623–656.

— (2015). "Knowledge-based hierarchies: Using organizations to understand the economy." In: *Annual Review of Economics* 7.1, pp. 1–30.

Ide, Enrique and Eduard Talamas (2025). "Artificial Intelligence in the Knowledge Economy." In: *Journal of Political Economy* 133.12, pp. 3762–3800.

Kapička, Marek and Ctirad Slavík (2022). "Organization of knowledge and taxation." In: *CERGE-EI Working Paper Series* 705.

López, José Joaquín and Jesica Torres (2020). "Size-dependent policies, talent misallocation, and the return to skill." In: *Review of Economic Dynamics* 38, pp. 59–93.

Mailath, George J, Andrew Postlewaite, and Larry Samuelson (2013). "Pricing and investments in matching markets." In: *Theoretical Economics* 8.2, pp. 535–590.

Nöldeke, Georg and Larry Samuelson (2015). "Investment and competitive matching." In: *Econometrica* 83.3, pp. 835–896.

Pieri, Fabio and Massimiliano Vatiero (2025). "Firm hierarchy and the market for knowledge." In: *Journal of Economics & Management Strategy* 34.3, pp. 714–742.

Shaked, Moshe and J George Shanthikumar (2007). *Stochastic orders*. Springer.

Teschl, Gerald (2012). *Ordinary differential equations and dynamical systems*. Vol. 140. American Mathematical Soc.

United States Census Bureau (Aug. 2025). *2022 SUSB Annual Data Tables by Establishment Industry*. U.S. Department of Commerce. URL: https://www.census.gov/data/tables/2022/econ/susb/2022-susb-annual.html (visited on 10/07/2026).

Wu, Yanhui (2015). "Organizational structure and product choice in knowledge-intensive firms." In: *Management Science* 61.8, pp. 1830–1848.

Zhang, Hanzhe (2020). "Pre-matching Gambles." In: *Games and Economic Behavior* 121, pp. 76–89.