Cross-Country Differences in the Optimal Allocation of Talent and Technology*

Tommaso Porzio[†]
University of California, San Diego

February 20, 2017

Abstract

I model an economy inhabited by heterogeneous individuals that form teams and choose an appropriate production technology. The model characterizes how the technological environment shapes the equilibrium assignment of individuals into teams. I apply the theoretical insights to study cross-country differences in the allocation of talent and technology. Their low endowment of technology, coupled with the possibility of importing advanced technology from the frontier, leads poor countries to a different economic structure – one with a stronger concentration of talent and a larger cross-sectional productivity dispersion. As a result, the efficient equilibrium in poor countries resembles economic features that are often cited as evidence of misallocation. Micro data from countries of all income levels document cross-country differences in the allocation of talent that support the theoretical predictions. A quantitative version of the model suggests that a sizable fraction of the larger productivity dispersion documented in poor countries is due to differences in the efficient allocation.

*This paper is based on my PhD dissertation at Yale University. This project deeply benefited from numerous conversations with colleagues and advisors. First of all, I am extremely grateful to my advisors, Mikhail Golosov, Giuseppe Moscarini, Nancy Qian, Aleh Tsyvinski, and Chris Udry, for their invaluable guidance and support and to Benjamin Moll, Michael Peters and Larry Samuelson for helpful suggestions and stimulating discussions. I also thank for very helpful comments Joseph Altonji, Costas Arkolakis, David Atkin, Lorenzo Caliendo, Douglas Gollin, Pinelopi Goldberg, Samuel Kortum, David Lagakos, Ilse Lindenlaub, Costas Meghir, Marti Mestieri, Nicholas Ryan, Alessandra Voena, Nicolas Werquin, and Noriko Amano, Gabriele Foà, Sebastian Heise, Ana Reynoso, Gabriella Santangelo, Jeff Weaver, and Yu Jung Whang. Last, but not least, I received many insightful comments from seminar participants at BU, Bocconi, Columbia, CMU (Tepper), Georgetown, IIES, Wisconsin-Madison, McGill, Princeton, Stanford, Toronto, UCSD, Yale, and the NBER Fall Development Meeting.

[†]9500 Gilman Drive, La Jolla, CA 92093, email: tporzio@ucsd.edu.

1 Introduction

Low-income countries have low labor productivity and, thus, produce lower output for the same amount of input. At the same time, cross-sectional productivity dispersion is larger in low-income countries: while most individuals see their labor produce very little output, some others enjoy fairly high productivity.[1] Often, this excessive productivity dispersion is interpreted as evidence of market frictions that prevent reallocation of inputs to more-productive firms and sectors or adoption of the best available technologies.[2]

In this paper, I explore an alternative view. I argue that the larger productivity dispersion may be a natural side effect of poor countries having access to advanced technologies imported from the distant technology frontier. The economic mechanism I propose entails two insights.

First, low-income countries have the unique opportunity to adopt technologies, discovered elsewhere, that are much more advanced than their current level of development. As long as individuals differ in their returns from adoption – due to heterogeneous ability, for example – only some would take advantage of this opportunity. Thus, heterogeneous adoption leads naturally to the dispersion of used technology and, as a result, of productivity. Take India, for example. It has approximately the level of GDP per capita that the United States had at the beginning of the 20^{th} century. Nonetheless, we would not be surprised to spot a person on the street in Kolkata using a cell phone to shop online with Flipkart, the Indian alternative to Amazon. And yet, still in Kolkata, another person could be cruising through the city on a pulled rickshaw, which was a common mode of transportation already in the 19^{th} century. A wide range of technologies coexist in India today - arguably, many more than in the United States at the beginning of the 20^{th} century, when cell phones and the internet were still decades away from being invented.

Second, I show that the opportunity to adopt advanced technology determines the way in which individuals of heterogeneous ability form production teams – that is, the allocation of talent – in low-income countries and that, at the same time, the different allocation of talent affects the dispersion of used technology itself. In fact, high-skilled individuals have higher returns from using advanced technology, and so they cluster together in the few economic niches where modern technology is pervasive. But, in equilibrium, if all high-skilled individuals form teams with each other, the rest of the economy is left with only low-skilled individuals who have low inherent productivity and few incentives to improve their technology. The concentration of similarly talented individuals within the same teams is, thus, both a consequence and a catalyst of productivity and technology dispersion.

The interaction of these two insights shapes the economic structure of poor countries. Some economic niches attract most of the high-skilled individuals and adopt the advanced technology. The rest of the economy is left without talent and, thus, finds it convenient to rely on backward, and cheaper, technology. Summing up, the possibility of adopting technology from the frontier endogenously generates cross-sectional dispersion and dual economies in poor countries.

[1] See Caselli (2005), and Hsieh and Klenow (2009) (Figure I of TFPQ, not the more famous Figure II of TFPR).

[2] Notable exceptions are Lagakos (2016), Lagakos and Waugh (2013), Caunedo and Keller (2016), and Young (2013).

The described economic forces are not mere theoretical possibilities. In this paper, I use micro data from several countries to show that talent is more concentrated in less-developed ones, as theory predicts. Moreover, I show, through the lens of a quantitative version of the model, that these differences are sizable and that the outlined mechanism can contribute to explaining the larger productivity dispersion in poor countries.

Overview. The paper is organized in four sections.

In Section 2, I develop a theoretical framework to analyze the joint determination of the allocation of talent and technology. The economy is inhabited by a continuum of individuals of heterogeneous ability. Production requires three inputs: a manager, a worker, and a technology. The production function satisfies two assumptions. First, output is more sensitive to manager's ability than to the worker's. Second, there is complementarity between technology and the ability of both the manager and the worker. Each individual chooses his occupation: whether to be a worker or a manager. Managers choose the ability of the worker to hire and the level of technology to operate. More-advanced technologies are costlier but allow higher labor productivity. The competitive equilibrium of this setting decentralizes the Pareto efficient allocation; I characterize it. The endogenous sets of managers and workers are the main objects of interest. Given these two sets, complementarity dictates positive assortative matching – more-able managers are matched with more-able workers – and so the unique equilibrium is pinned down. As is standard, each individual chooses the occupation in which he has a comparative advantage. The main difficulty of this setting is that the comparative advantages are endogenous and depend on the technology choice of each team. If all the teams pick the same technology, the higher skill sensitivity of output dictates that the most skilled individuals would have a comparative advantage in being managers. Endogenous technology choice introduces a new trade-off: in equilibrium, each individual uses a higher technology if he decides to be a worker. Hence, to use a more advanced technology – that increases the marginal value of one's skills, – one needs to choose an occupation that is less skill-intensive. This trade-off is modulated by the technological environment. The main theoretical result is that the more important is the choice of an appropriate technology – i.e., the larger is the elasticity of the optimal technology choice to ability, relative to the production's asymmetry in skill sensitivity – the smaller is the ability gap between two individuals working together. This result shows that the allocations of talent and technology (and, thus, productivity) are tied together: when technology is dispersed, there is a lower ability gap between managers and workers.

In Section 3, I explicitly model cross-country differences in the cost of technology. When I use the term "technology," I refer to the input of the production function, which is simply a productivity term that multiplies labor input.[3] This notion of technology depends on two elements: (i) the choice of which vintage of technology to use – e.g., to rely on animal or electric power; and (ii) the amount of capital of the used vintage – e.g., how many bullocks are purchased. More modern vintages have a lower variable cost of technology, but a larger fixed one to be

[3] I use technology to distinguish it from productivity, which is output divided by the number of workers, and, thus, takes into account worker ability.

operated. Countries differ in their local technology vintage, and in each country, every production team can use the local vintage without paying any fixed cost. However, to import more modern vintages from other countries, teams must pay a fixed cost that increases in the distance between the chosen vintage and the local one. The described environment yields a country-specific cost of technology. Moreover, the elasticity of the optimal technology to the ability of team members depends on the distance of a country from the technology frontier, defined as the gap between the local vintage and the most advanced one available. This result is generated by the heterogeneity of returns to technology adoption: even in countries far from the frontier, the most able teams pay the fixed cost to import modern vintages, while the less able ones prefer to rely on local vintages.

Combining the insights from Sections 2 and 3, the theory provides sharp predictions on cross-country differences in economic structure. First, the model predicts that the efficient competitive equilibrium generates larger technology and labor productivity dispersion in poor countries. Large productivity dispersion is often associated with misallocation. In this framework, instead, it is generated through differences in endowments: poor countries are endowed with less-advanced local technology vintages and this leads, in equilibrium, to larger productivity dispersion. Second, the theory predicts cross-country differences in the way in which people form production teams. Specifically, talent should be more concentrated in countries far from the frontier: an empirical prediction, to the best of my knowledge, unique to this model.

In Section 4, I confront the model's predictions on the allocation of talent with the data. In the main empirical exercise, I use large-sample labor force surveys for 63 countries to show that in the countries farther from the technology frontier – i.e., with lower relative GDP per capita – the concentration of talent is higher. In the data, neither teams nor individual skills are directly observed. I show that under two assumptions – (i) education is correlated with ability; and (ii) individuals within the same industry use a more similar technology than those in different industries – we can construct, using the distribution of schooling within and across industries, an empirical measure of concentration of talent that aligns with the one defined in the model. I construct this measure for each country-year pair and show how it varies as a function of distance to the frontier, both in the cross-section and in the time-series. I first compare countries in the cross-section. The concentration of talent is negatively correlated with GDP per capita, and the magnitude of cross-country differences is sizable. I then compare middle-income countries today (such as Mexico and Brazil) with the United States in 1940, which was at the same level of development – but, critically, was closer to the technology frontier. Middle-income countries today have a higher concentration of talent than the U.S. did in the past. This result alleviates the concern that cross-sectional differences are merely capturing differences in levels of development. Finally, I compare the growth paths of South Korea and the United States in the past seven decades. In the United States, the concentration of talent has remained relatively unchanged, consistent with the fact that is has grown steadily as a world leader (i.e., on the frontier). In contrast, South Korea has seen a rapid decline in the concentration of talent as it has approached the technology frontier.

I then provide further supporting evidence from occupation data: managers are, on average, less skilled in poor countries, while workers are more skilled, as the theory predicts. However, the definitions of managers and workers in the data are not perfectly comparable across countries. For this reason, I also focus on two specific occupations: retail cashiers and motor vehicle drivers. Both of these occupations require simple tasks (they would correspond to workers in the model) but use advanced technologies – i.e., the cash register and the combustion engine. The model predicts that in poor countries, retail cashiers and motor vehicle drivers should be relatively more skilled. This prediction is confirmed in the data.

I end Section 4 by using use the World Bank Enterprise Surveys to indirectly validate the main model assumptions – the tasks and skill-technology complementarities – and to confirm that the allocation of talent is different in poor countries even at the firm level.

Finally, In Section 5, I write a computational version of the model and show that the cross-country differences in the concentration of talent are consistent with sizable differences in within-country productivity dispersion. Specifically, I define a notion of agriculture in the model, and I estimate the model to target cross-country differences in the concentration of talent and the agricultural productivity gap in rich countries. I then predict agricultural productivity gaps in poor countries and show that the model, once disciplined with the cross-country differences in the allocation of talent, accounts for approximately one third of the larger agricultural productivity gap in poor countries.

Related Literature. The theoretical contribution of the paper is to combine two strands of the literature and show that they interact non-trivially. I consider an environment in which individuals form production teams – following the seminal work of Lucas (1978), Kremer (1993) and, more recently, Garicano and Rossi-Hansberg (2006) – and, at the same time, choose an appropriate production technology – along the lines of the work of Atkinson and Stiglitz (1969), Basu and Weil (1998), and Acemoglu and Zilibotti (2001). I show that allowing teams to choose a production input, technology, that is complementary to the team members' ability changes the equilibrium assignment. To the best of my knowledge, this is a new insight for the literature on team formation. At the same time, taking endogenous team formation into consideration is relevant to understanding the distribution of technology within a country. Specifically, equilibrium team formation links the right and left tails of the technology distribution: the possibility of adopting advanced technology generates a right tail of high-technology teams that gather all the talent, thus crowding out high-skilled individuals from the rest of the economy; and creating a left tail of low technology and, thus, low productivity, teams. In other words, modern technology in some economic niches within a country comes at the cost of backward technology elsewhere. To the best of my knowledge, this is a new insight for the literature on appropriate technology and technology adoption. In studying the distribution of technology and productivity as the result of agents' maximization behavior, my work is similar to Perla and Tonetti (2014). In allowing countries to import more-advanced technologies from abroad, my work is similar to Buera and Oberfield (2016). Neither of these papers considers the endogenous allocation of talent.

A growing literature studies the allocation of talent in Roy models of occupational choice, in which individuals draw a vector of occupation-specific abilities that determine their comparative advantage; – one recent application is Hsieh et al. (2016). My work, instead, considers heterogeneity along one single trait, a general skill. Most of the literature with occupational choice and unidimensional skills, starting from the seminal Lucas (1978), makes assumptions on the production function such that the highest skilled have a comparative advantage in being managers. In my setting, instead, the comparative advantages and, thus, the pattern of matching, is endogeneous. Specifically, it depends on the equilibrium distribution of technology. In this respect, my work mostly resembles Kremer and Maskin (1996). While Kremer and Maskin (1996) studies how changes in skill distributions can impact the pattern of matching, I keep the skill distribution constant and show, instead, that endogenous choice of technology shapes the pattern of matching. I also provide a new solution method and develop a quantitative version of the model.

Similarly to Kremer (1993), this paper uses a frictionless model to understand cross-country differences in organization of production. However, Kremer (1993) focuses on *average* differences across countries – e.g., in poor countries, firms are, on average, smaller – while I focus on cross-country differences in the within-country *distribution* of economic activity – e.g., in poor countries, large and small firms coexist, while in rich ones, all firms are similar in size.

The idea that distance to the frontier may impact the organization of production is present in Acemoglu et al. (2006), which studies selection into entrepreneurship with credit constraints. I see my work as complementary to theirs. Roys and Seshadri (2014) also studies differences across countries in the way in which production is organized. It uses a quantitative version of Garicano and Rossi-Hansberg (2006) in which human capital is endogenously accumulated, as in Ben-Porath (1967). Cross-country differences are generated by changes in the distribution of talent, and not by changes in the pattern of matching, which, in their work, is fixed ex-ante.

The application of the theory to developing countries fits into the debate on dual economies. Through the lens of the model, the fact that individuals in poor countries have the ability to adopt advanced technology from the frontier leads to the endogenous formation of dual economies. This view is original but resembles most closely that of La Porta and Shleifer (2014), which emphasizes how duality is tightly linked to economic development.

Finally, two other papers, Acemoglu (1999) and Caselli (1999), argue that the technological environment and the allocation of workers to jobs are connected. Acemoglu (1999) focuses on the interaction between labor market frictions and the fact that firms have to commit ex-ante whether to create jobs for high- or low-skilled workers, and, it shows the conditions for a separating equilibrium to exist. Caselli (1999) shows that when new technologies are adopted, the most skilled individuals are the ones most likely to start using them, thus separating themselves from the rest of the economy. Neither of these papers considers complementarity between individuals working together. My work focuses exactly on this latter channel and characterizes how the properties of the technological environment change the overall production complementarity, thus changing the assignment of workers to jobs.

2 A Model of Technology Choice and Allocation of Talent

I develop an assignment model in which heterogeneous individuals form production teams and choose an appropriate and costly technology.

2.1 Environment

The economy is inhabited by a continuum of mass one of individuals, indexed by their ability $x \sim U[0,1]$. Individuals with higher x are more able. Each individual supplies inelastically one unit of labor and has a non-satiated and increasing utility for the unique final good produced in the economy.

Production. Production of labor input requires two individuals to form a team: a manager and a worker. A manager of ability x' paired with a worker of ability x produces $f(x', x)$ units of labor inputs, where $f(x', x)$ is strictly increasing in both arguments and twice continuously differentiable.[4] Production of the final good requires the labor input to be combined with a production technology $a \in \mathbb{R}$, that multiplies labor input. There is a continuum of available technologies in the economy, and technology a has a cost $c(a)$, in units of output. This cost is increasing, convex, and twice continuously differentiable. A production team (a, x', x) where a is the technology, x' is the ability of the manager, and x is the ability of the worker produces $g(a, x', x)$ units of the final output

$$g\left(a, x', x\right) \;\; = \;\; af\left(x', x\right) - c\left(a\right).$$

Assumptions on Production of Labor Input. I assume that the production function of labor input satisfies the following three[5] properties:

1. $f_1(x, x) > f_2(x, x)$, for all $x \in [0,1]$;
2. $f_{12}(x', x) > 0$, for all $(x', x) \in [0,1] \times [0,1]$;
3. $f_1(x, y) > f_2(z, x)$, for all $(x, y, z) \in [0,1] \times [0,1] \times [0,1]$.

(1) captures the fact that, for a given technology and partner, the individual's skills are more useful (has a greater effect on output) if he is employed in a managerial position. This assumption is common in the literature. See, for example, the seminal paper by Lucas (1978), which assumes that only the manager's ability matters for production, and Garicano and Rossi-Hansberg (2006), which builds from primitives a production function that features this property. (2) captures complementarity in production between tasks. This is a natural assumption, pervasive in the literature. (3) is a Spence-Mirrlees condition that separates types into occupations by imposing that, if all teams use the same technology, the complementarity between types is sufficiently weak relative to the skill asymmetry, and, thus, high-skilled individuals have a comparative advantage in

[4]Note the slight abuse of notation: I refer to x as the ability of an individual in general; however, when I want to distinguish explicitly between the ability of the manager and that of the worker, I refer to the former as x' and to the latter as x.

[5]Assumption (3) is stronger than (1), which is, formally, a redundant assumption. Nonetheless, it is conceptually convenient to separate the two.

the managerial occupation. As I will show, (3) is not sufficient to separate types if the technology choice is endogenous.

Complementarity Skill-Technology. Assumption (3) implies that

4. $g_{12}(a,x,y) > g_{13}(a,z,x) \geq 0$ for all $(a,x,y,z) \in \mathbb{R} \times [0,1] \times [0,1] \times [0,1]$.

Technology and skills are complementary, consistent with most of the evidence for both developed and developing countries (see Goldin and Katz (1998), Foster and Rosenzweig (1996) and Suri (2011)). Further, the manager's ability is more relevant than the worker's for generating high returns from technology, as emphasized in recent studies that highlight the role of managers in technology adoption (see Bloom and van Reenen (2007) and Gennaioli et al. (2013)).

The functional form assumption of g implicitly assumes a specific strength of the complementarity between skills and technology. This restriction is useful for tractability but is not necessary for the main results. In Appendix C, I consider a more general specification for g.[6]

Assignment of Talent to Teams. Production requires one manager and one worker. Individuals' ability x is observable, and individuals are not restricted ex-ante to belong to either group. As a result, an allocation in this setting comprehends an occupational choice function, $\omega(x) : [0,1] \to [0,1]$, which defines the probability that an individual x is a worker, and a matching function, $m(x) : [0,1] \to [0,1]$, which assigns to each individual x the type of the manager he would be matched with if he decides to be a worker.[7]

2.2 Competitive Equilibrium

The competitive equilibrium of this economy is given by five functions: optimal technology $\alpha(x',x) : [0,1] \times [0,1] \to \mathbb{R}$; profit $\pi(x) : [0,1] \to \mathbb{R}$; wage $w(x) : [0,1] \to \mathbb{R}$; occupational choice $\omega(x) : [0,1] \to [0,1]$; and matching $m(x) : [0,1] \to [0,1]$ such that

1. each team chooses the optimal technology

$$\alpha(x',x) = \arg\max_{a \in \mathbb{R}} a f(x',x) - c(a);$$

2. each manager chooses the type of worker to hire, taking into account the optimal technology that the pair would choose and taking as given the equilibrium wage schedule

$$\pi(x) = \max_{z \in [0,1]} \alpha(x,z) f(x,z) - w(z) - c(\alpha(x,z));$$

3. the matching function is consistent with manager's optimality

$$m(z^*(x)) = x$$

[6] The Appendix is available online at https://sites.google.com/a/yale.edu/tommaso-porzio/.

[7] The definition of $m(x)$ as a function rather than a correspondence might seem restrictive. However, it is not. In fact, I prove in a previous version of this paper, (Porzio (2016)), that for the currently considered case in which $x \sim U[0,1]$, $m(x)$ is indeed a function and not a correspondence. This is a standard result in the literature, and for this reason I omit the details here.

where $z^*(x)$ is the solution to the manager's problem;

4. each individual chooses the occupation that pays him the higher income or randomizes among them if $\pi(x) = w(x)$: $\omega(x)$ satisfies

$$\omega(x) \in \arg\max_{z\in[0,1]} (1-z)\,\pi(x) + zw(x)\,;$$

5. the sum of wage and profit of a team equals its produced output, for all x

$$\pi(m(x)) + w(x) = g(\alpha(m(x), x), m(x), x)\,.$$

This restriction also guarantees that the goods market clears; and

6. labor market clears for each individual type x, and all individuals are employed. That is, for each type, the mass of less-skilled workers must equal the mass of managers matched with these workers according to m, and the mass of managers and workers must sum to

$$\begin{aligned}\int_0^x \omega(z)\,dz &= \int_0^x (1-\omega(m(z)))\,dz \\ \int_0^1 \omega(z)\,dz + \int_0^1 (1-\omega(z))\,dz &= 1.\end{aligned}$$

Proposition 1: Existence and Pareto Efficiency.
A competitive equilibrium exists and is Pareto Efficient.
Proof. See appendix. □

Equilibrium uniqueness is not guaranteed in this environment. Nonetheless, this is not a concern since the properties of the equilibrium that I characterize in the next section hold for any equilibrium.

2.3 Equilibrium Characterization

I now characterize the equilibrium and show how the assignment of talent to teams and the choice of technology are related.

2.3.1 Choice of Technology

A team of a manager of ability x' and a worker of ability x picks the technology that maximizes net output; that is,

$$\alpha(x', x) = c'^{-1}(f(x', x))\,.$$

The complementarity between technology and labor input and the assumptions on the functional form of f give the following result.

Lemma 1: Appropriate Technology

The appropriate technology of a team increases in the skills of both the manager and the worker, but more so in the skills of the manager: $\alpha_1 > \alpha_2 \geq 0$.

Proof. See appendix. $\square$

The manager and the worker agree on the choice of technology since it does not affect how they share output.

2.3.2 Manager Problem

Consider a manager of ability x. He picks the optimal type of workers to maximize his profit $\pi(x)$; that is,

$$\pi(x) = \max_{z \in [0,1]} \alpha(x,z) f(x,z) - w(z) - c(\alpha(x,z)).$$

The solution of this maximization problem yields the matching function, which assigns to each worker his manager and satisfies $m(z^*(x)) = x$. The skill complementarity between managers and workers ($f_{12} > 0$) implies that the matching function is increasing.

Lemma 2: Matching Function

The matching function m *is increasing: for all* $(x', x) \in [0,1]^2$*, if* $x' > x$ *and* $\int_x^{x'} \omega(z)\, dz > 0$*, then* $m(x') > m(x)$.

Proof. See appendix. $\square$

The envelope and first-order conditions give the slopes of the profit and wage functions

$$\begin{aligned} \pi'(x) &= \alpha\left(x, m^{-1}(x)\right) f_1\left(x, m^{-1}(x)\right) \\ w'(x) &= \alpha(m(x), x) f_2(m(x), x), \end{aligned} \tag{1}$$

where $m^{-1}(x)$ is the inverse of the matching function and, thus, assigns each manager his type of worker. By definition, $m^{-1}(x) = z^*(x)$, and the function m is invertible due to the fact that it is strictly increasing over the relevant domain. The slopes of the profit and wage functions determine the marginal values of skills in each occupation, which, as I will show, drive the occupational choice. These values depend on the technology used, the occupation-specific skill sensitivity, and the production partner. In general, an individual has different partners and, thus, different appropriate technologies whether he is a worker or a manager. The higher the technology used, the more skills are valued – due to skill-technology complementarity. At the same time, the manager's tasks are different from the worker's; thus, each occupation is going to have a specific skill sensitivity that affects its overall marginal value of skills.

2.3.3 Occupational Choice

I first discuss the optimal assignment to occupations within a team. Two individuals that work together use, by assumption, identical technology. As a result, the relative marginal value of skills in either occupation depends only on the asymmetry in skill sensitivity. Due to the Spence-Mirrlees assumption, the managerial task uses skills more efficiently, and, thus, the more skilled individual of the team must be the manager.

Lemma 3: Occupational Choice within a Team

The manager of the team is more skilled than the worker of the team: $m(x) \geq x$ *for all* $x \in [0,1]$.

Proof. See below, after Lemma 4. □

The *Technology-Occupation* Tradeoff. Lemma 3 implies that an individual x is matched with a more-skilled partner and, thus, uses a more advanced technology if he decides to be a worker rather than a manager. There is a *technology-occupation* trade-off: an individual would use a higher technology, which gives his skills, *ceteris paribus*, a higher marginal value, if he selects into the less-skill-intensive occupation. As a result, an individual x sees his skills having a higher marginal value as a manager – i.e., $\pi'(x) \geq w'(x)$ – if and only if the gap in skill sensitivity across occupations is larger than gap in technology used:

$$\underbrace{\frac{f_1\left(x, m^{-1}(x)\right)}{f_2\left(m(x), x\right)}}_{\text{Skill Sensitivity Gap}} \geq \underbrace{\frac{\alpha\left(m(x), x\right)}{\alpha\left(x, m^{-1}(x)\right)}}_{\text{Technology Gap}}.$$

Individuals, as usual, select into the occupation in which they have their comparative advantage. High-skilled individuals have a comparative advantage in the occupation that has the highest marginal value of skills. Most of the literature[8] focuses on functional form assumptions such that management is more skill-intensive – that is, $\pi'(x) \geq w'(x)\ \forall x$. For example, Lucas (1978) uses a production function that gives $w'(x) = 0$: the high-skilled have a comparative advantage in being managers; thus, the shape of the occupational choice function is known ex-ante and is given by a cutoff policy that separates types into two connected sets of managers and workers. In my setting, instead, the comparative advantage of each individual x is endogenous since it depends on the optimal technology choice of each team. For example, some high-skilled individuals may find their skills more rewarded by being workers with a high technology rather than managers with a lower one. To solve this complex fixed point, I develop a method to use the necessary conditions for optimality, together with market clearing, to characterize the overall equilibrium assignment and to show how it depends on the shape of f and α.

Necessary Conditions. The occupational choice function $\omega(x)$ maximizes individual income:

$$\omega(x) \quad \in \quad \arg\max_{z\in[0,1]} (1-z)\,\pi(x) + zw(x). \tag{2}$$

$\omega(x)$ divides the type space into subsets in which individuals are managers, workers, or randomize between the two occupations. Since the type space is a compact set, (i) individuals that are at the boundaries between two subsets in which two different occupations are chosen, must be indifferent between being a manager or a worker; and (ii) individuals that randomize between the two occupations must also be indifferent. For these two groups of individuals, the maximization problem (2) provides useful necessary conditions that link the occupational choice to the marginal

[8] The one notable exception that I am aware of is Kremer and Maskin (1996).

values of skills in either occupation.

Lemma 4: Necessary Conditions of Occupational Choice Function
The occupational choice function satisfies:

1. *for all x such that* $\lim_{\epsilon\to 0}\omega(x-\epsilon)\in(0,1)$ *and* $\lim_{\epsilon\to 0}\omega(x+\epsilon)\in(0,1)$*:* $\pi'(x)=w'(x)$*;*

2. *for all x such that* $\lim_{\epsilon\to 0}\omega(x-\epsilon)=0$ *and* $\lim_{\epsilon\to 0}\omega(x+\epsilon)>0$*:* $\pi'(x)\leq w'(x)$*;*

3. *for all x such that* $\lim_{\epsilon\to 0}\omega(x-\epsilon)=1$ *and* $\lim_{\epsilon\to 0}\omega(x+\epsilon)<1$*:* $\pi'(x)\geq w'(x)$.

Proof. See appendix. □

Formally, these conditions are derived from the fact that the wage and profit functions must cross in proximity of an ability type x that is indifferent between either occupation.

I use these necessary conditions to characterize the equilibrium. The proofs of the results are left to the online appendix. Nonetheless, in order to illustrate the general method I use, I show here how the necessary conditions can be used to prove Lemma 3.

***Proof* of Lemma 3.** Let $x>x'$, $\omega(x)>0$ and suppose that $m(x)=x'$. Then, let $\hat{x}$ be the lowest type larger than x' with $\omega(\hat{x})>0$. Since $\omega(x)>0$ and – by market clearing – $\omega(x')<1$, then $\hat{x}\in[x',x]$. By the necessary conditions, $w'(\hat{x})\geq\pi'(\hat{x})$. Substituting in Equation (1), $w'(\hat{x})\geq\pi'(\hat{x})$ becomes $\frac{\alpha(m(\hat{x}),\hat{x})}{\alpha(\hat{x},m^{-1}(\hat{x}))}\geq\frac{f_1\left(\hat{x},m^{-1}(\hat{x})\right)}{f_2(m(\hat{x}),\hat{x})}$. Lemma 1 shows that m and m^{-1} are increasing. As a result, $\hat{x}\leq x$ implies that $m(\hat{x})\leq m(x)=x'\leq\hat{x}$, and, also, $\hat{x}\geq x'$ implies that $m^{-1}(\hat{x})\geq m^{-1}(x')=x\geq\hat{x}$. Therefore, $\frac{\alpha(m(\hat{x}),\hat{x})}{\alpha(\hat{x},m^{-1}(\hat{x}))}\leq 1$ since α is increasing in both of his arguments. The Spence-Mirrlees assumption (3) guarantees that $\frac{f_1\left(\hat{x},m^{-1}(\hat{x})\right)}{f_2(m(\hat{x}),\hat{x})}>1$. This leads to a contradiction and, thus, to $m(x)\geq x$. □

2.3.4 Technology Choice and Equilibrium Assignment of Talent to Teams

The main characterization result ties the properties of the optimal technology choice α and the production function f to the assignment of talent to teams.

Proposition 2: Assignment of Talent Across Teams
In a competitive equilibrium, for any worker $x\in[0,1]$ the ability gap between him and his manager, $m(x)-x$, is bounded above by $\Upsilon(x)$ and below by $\Lambda(x)$, where $\Upsilon(x)$ and $\Lambda(x)$ depend on f and α as follows:

1. *Consider two functions f and $\tilde{f}$, if for all $(x',y,z)\in[0,1]^3$ such that $x'\geq y\geq z$ $\frac{f_1(y,z)}{f_2(x',y)}\geq\frac{\tilde{f}_1(y,z)}{\tilde{f}_2(x',y)}$, then $\Upsilon(x)\geq\tilde{\Upsilon}(x)$ and $\Lambda(x)\geq\tilde{\Lambda}(x)$; and*

2. *consider two functions α and $\tilde{\alpha}$, if for all $(x',y,z)\in[0,1]^3$ such that $x'\geq y\geq z$ $\frac{\alpha(x',y)}{\alpha(y,z)}\geq\frac{\tilde{\alpha}(x',y)}{\tilde{\alpha}(y,z)}$, then $\Upsilon(x)\leq\tilde{\Upsilon}(x)$ and $\Lambda(x)\leq\tilde{\Lambda}(x)$ $\forall x\in[0,1]$.*

Proof. See appendix. □

In the appendix, I include the implicit functions that define each bound. The proposition shows that $\Lambda(x)\leq m(x)-x\leq\Upsilon(x)$ and that these upper and lower bounds depend on the shape of f and α. f and α modulate the technology-occupation trade-off and, thus, change the equilibrium assignment of talent to teams. When there is stronger asymmetry in skill sensitivity, so that high-skilled individuals have a stronger comparative advantage in being managers, then the gap between workers and managers widens, since more of the high-skilled find it optimal to be managers. When, instead, the elasticity of optimal technology choice to the team members' ability increases – i.e., when teams formed by individuals of different ability choose very different optimal technologies – the choice of occupation becomes a less important, and the technology used becomes a more important, determinant of the marginal value of an individual's skills. As a result, some high-skilled individuals find it optimal to be workers and, in equilibrium, the skill gaps between workers and their managers decrease.

It is simple to see that when $m(x) - x$ is smaller, talent is more concentrated, as some teams bring together the high-skilled, and, thus, in equilibrium, other teams are left with low-skilled managers and workers. Formally, I define the concentration of talent as follows.

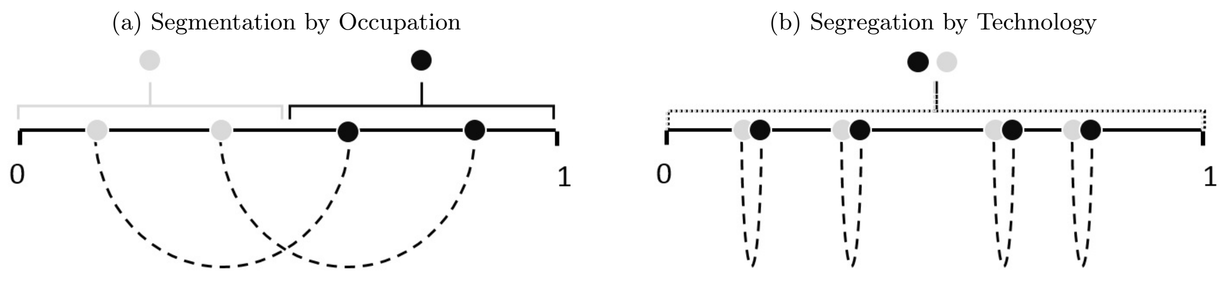
Definition 1: Concentration of Talent
Consider two matching functions m and $\tilde{m}$. Talent is more concentrated according to m if $\int [m(x) - x]\,\omega(x)\,dx < \int [\tilde{m}(x) - x]\,\omega(x)\,dx$. I define $1 - \int [m(x) - x]\,\omega(x)\,dx$ the concentration of talent.

Proposition 2 shows that the concentration of talent is tightly linked to the strength of either side of the technology-occupation trade-off. I next consider two polar cases with respect to this trade-off and show the conditions under which they emerge.

Definition 2: Segmentation by Occupation
Talent is segmented by occupation if $x' > x$ and $\omega(x) < 1$ imply that $\omega(x') = 0$.

Definition 3: Segregation by Technology
Talent is segregated by technology if $x' > x$ implies that $\alpha\left(x', m^{-1}(x')\right) \geq \alpha(m(x), x)$ with probability one.

Figure 1: Two Polar Cases of Allocation of Talent



Notes: A grey dot represents a worker, while a black dot represents a manager. Dashed lines connect workers to their managers.

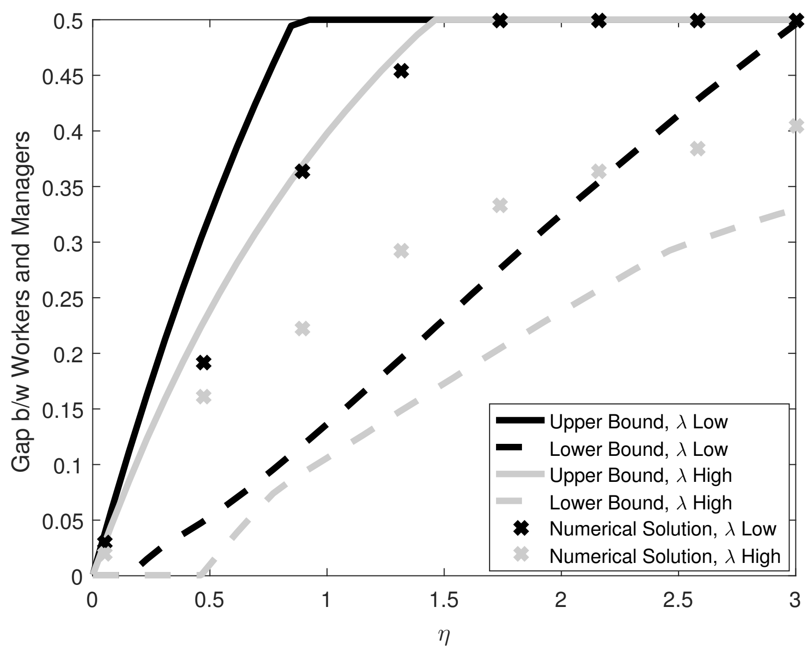
When talent is segmented by occupation, there is a cutoff type such that all individuals with ability above this cutoff are managers, and all of those with ability below the cutoff are workers.

This case is illustrated in Figure 1a below. When talent is segregated by technology, more-skilled individuals use a higher technology than lower-skilled individuals use. When the function $\alpha(x', x)$ strictly increases in both its arguments, this requires that $m(x) \to x$. This second case is illustrated in Figure 1b.

Corollary 1: Conditions for Segregation and Segmentation
If α and f satisfy $\frac{\alpha(x',y)}{\alpha(y,z)} < \frac{f_1(y,z)}{f_2(x',y)}$ for all $(x', y, z) \in [0,1]^3$, with *$x' > y > z$, then talent is segmented by occupation and $m(x) = \frac{1}{2} + x$. If α satisfies $\frac{\alpha(x',y)}{\alpha(y,z)} \to \infty$ for all $(x', y, z) \in [0,1]^3$,* with *$x' > y > z$, then talent is segregated by technology and $m(x) \to x$.*

Proof. See appendix. □

Figure 2: Lower and Upper Bounds of Ability Gap between Managers and Workers



Notes: The figure shows the bounds according to Proposition 2 compared with actual values from the numerical solution.

Proposition 2 bounds the concentration of talent, while Corollary 1 shows two limit cases of it. It is natural to wonder how wide the bounds provided are. The answer depends on the specific choice of the functional form. To explore whether the bounds can be informative, I provide an example with simple functional forms. Let $f(x', x) = x'(1 + \lambda x)$, with $\lambda \leq 1$, and $c(a) = \frac{a^{1+\eta}}{1+\eta}$, with $\eta \geq 0$, which implies that $\alpha(x', x) = (x'(1 + \lambda x))^{\frac{1}{\eta}}$. This system of functional forms has two parameters: lower λ generates stronger skill asymmetry, and lower η generates larger technology gaps. In Figure 2, I plot the upper and lower bounds for the median individual $x = \frac{1}{2}$ as a function of η and for two different values of λ. These bounds are informative. In the same figure, I also plot the actual gaps $m\left(\frac{1}{2}\right) - \frac{1}{2}$, computed numerically using the Hungarian algorithm.

This section has characterized how the optimal technology choice and the allocation of talent are linked. The results are characterized in terms of one intuitive property of α, the technology gap, rather than in terms of the primitive $c(a)$. In the next section, I model $c(a)$.

3 Cross-Country Differences in the Allocation of Talent

I describe a technological environment that yields a country-specific functional form for the cost of technology. I characterize the properties of the cost function and obtain sharp predictions for the relationship between the distance of a country from the technology frontier and its allocation of talent.

3.1 Distance to the Technology Frontier and Cost of Technology

Recall that the gross output of a team (x', x) that uses technology a is $af(x', x)$. Technology a is a multiplicative productivity term.[9] More specifically, I think of a as a reduced-form term that captures two separate elements that determine the team's productivity: (i) the technology vintage that a team uses; and (ii) the capital intensity at which the specific technology vintage is operated. Consider an example. In order to grind one hundred pounds of grain into flour in an hour, a team could use either many horse-powered mills or fewer but faster electric mills. The team can achieve the same level of "technology" a – one hundred pounds of grain per hour – with either many units of an old vintage of technology (horse-powered mills) or fewer units of a newer vintage of technology (electric mills). Thus, the cost of achieving technology a needs to take into consideration (i) the optimal choice of technology vintage and its fixed cost; and (ii) the amount of capital of that technology vintage that is necessary to achieve a and its variable cost.

I describe, consistent with this interpretation of technology a, the world technological environment.

The Technological Environment. The most advanced vintage of available technology is $\bar{t}$, which represents the technology frontier at a given point in time. Each country is indexed by its level of development $t \leq \bar{t}$. All individuals in country t know how to use technology vintage t; hence, they do not need to pay any fixed cost to use it. I will refer to t as the local technology vintage. The distance of a country from the technology frontier is given by $\bar{t} - t$ – that is, the distance at a given point in time between that country's level of development and world knowledge. Individuals in each country can decide to import and learn how to use better vintages of technology, up to the frontier one $\bar{t}$, paying a fixed cost. The lower the country's level of development, the higher the fixed cost. A more modern technology vintage allows a team to achieve the same level of productivity a at a lower variable cost, but it has a higher associated fixed cost. Thus, there is a trade-off that leads to the optimal choice of technology vintage.

I summarize this technological environment using a system of equations, from which I then derive a country-specific cost of technology. The choice of functional forms is guided by parsimony and tractability, rather than by generality.

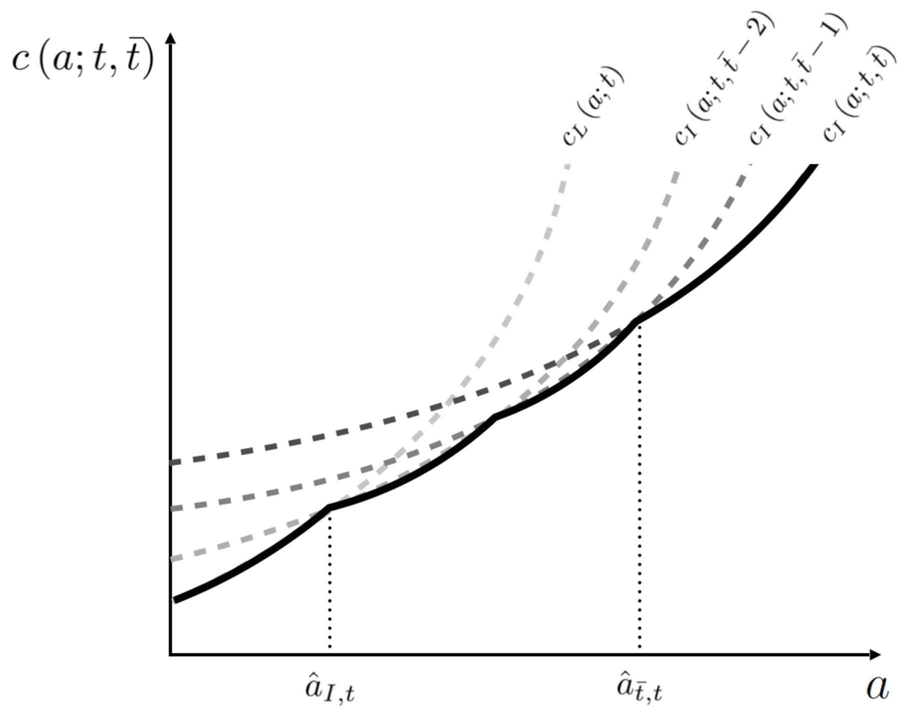
Cost of Technology. The cost of achieving productivity level a in country t using the local technology vintage is

$$c_L(a; t) = \gamma^{-\eta t} \kappa_1^{-\eta} \frac{(a - a_t)^{1+\eta}}{1+\eta},$$

[9] I call it technology to distinguish it from labor productivity, that is simply defined as output divided by labor.

where γ captures the productivity improvements across technology vintages; η captures the within-vintage cost elasticity of technology; and $a_t = \nu\gamma^t$ with $\nu \in [0,1]$ is a non-homotheticity term that becomes useful in the limit case described below, but in general can be assumed to be 0. Since individuals already know how to use the local vintage, there is no associated fixed cost. Hence, we should think of $c_L(a;t)$ as the variable cost component linked to the chosen capital intensity. κ_1 is a constant term that scales the cost of technology. I choose $\kappa_1 \equiv \frac{\eta\varepsilon-1}{\eta\varepsilon}$ to guarantee that the marginal type that decides not to use the local technology does not depend on either t or $\bar{t}$. While restrictive, this provides tractability.

Figure 3: Cost of Technology



Notes: dashed lines represent the cost of technology for one specific technology vintage. The lightest gray line is the cost for the local vintage, while the darkest gray line is the cost for the frontier vintage. The black solid line is the cost of technology that takes into account the choice of technology vintage. $\hat{a}_{I,t}$ is the level of technology below which the local vintage minimizes the cost. $\hat{a}_{\bar{t},t}$ is the level of technology above which the frontier vintage minimizes the cost.

The cost of achieving productivity level a in country t using an imported technology vintage $\tilde{t}$ is given by

$$c_I\left(a,\tilde{t};t\right) = \gamma^{-\eta\tilde{t}}\frac{\left(a-a_t\right)^{1+\eta}}{1+\eta} + \kappa_2^{-(\eta\varepsilon-1)}\frac{\gamma^t\gamma^{\varepsilon\eta\left(\tilde{t}-t\right)}}{\varepsilon\left(1+\eta\right)},$$

where $\kappa_2^{-(\eta\varepsilon-1)}\frac{\gamma^t\gamma^{\varepsilon\eta(\tilde{t}-t)}}{\varepsilon(1+\eta)}$ is the fixed cost associated with learning technology vintage $\tilde{t}$; ε captures the elasticity of the cost of technology across vintages; $\kappa_2 \equiv \chi_0^{-\frac{\eta+1}{\eta}\frac{1}{\eta\varepsilon-1}}$ is defined so that the marginal team that decides to import technology rather than using the local vintage is given by

χ_0. I assume that $\chi_0 < f\left(\frac{1}{2}, 0\right)$, so that not all teams decide to import, even when talent is segmented. The fixed cost is increasing in the gap between the current level of development and the vintage of technology $\tilde{t}$ to capture that individuals in less-developed countries may have a harder time learning and using more-advanced technologies, and extracting proper returns from it; coherently with the view that frontier technologies are targeted to the level of development of rich countries, as in Acemoglu and Zilibotti (2001).

Finally, the cost of achieving productivity level a in country t, taking into consideration the optimal vintage choice, is given by

$$c(a;t,\bar{t}) \quad = \quad \min\left\{c_L(a;t), \min_{\tilde{t}\leq\bar{t}} c_I\left(a,\tilde{t};t\right)\right\}.$$

I assume that $\eta\varepsilon < 1$ in order for the technology vintage choice to have, possibly, an interior solution – i.e., some teams may choose nor the local, nor the frontier vintage.

Optimal Technology. This system of functional forms provides simple solutions for the optimal technology α.

The cost function $c(a;t,\bar{t})$ is the lower envelope of the costs of technology for each vintage. More-modern vintages have a lower variable cost of productivity, but a higher fixed cost. As a result, as Figure (3) shows, teams that choose a higher technology minimize costs by choosing a more advanced vintage. The next three Lemmas summarize the cost of technology, the resulting optimal technology choice, and its properties.

Lemma 5: Cost of Technology

The cost of a technology $c(a;t,\bar{t})$ is given by

$$c(a;t,\bar{t}) = \begin{cases} \gamma^{-\eta t}\kappa_1^{-\eta}\frac{(a-a_t)^{1+\eta}}{1+\eta} & \text{if } a \leq \hat{a}_{I,t} \\ \gamma^{-\eta_\varepsilon t}\kappa_2^{-\eta_\varepsilon}\frac{(a-a_t)^{1+\eta_\varepsilon}}{1+\eta_\varepsilon} & \text{if } a \in \left(\hat{a}_{I,t}, \hat{a}_{\bar{t},t}\right) \\ \gamma^{-\eta\bar{t}}\frac{(a-a_t)^{1+\eta}}{1+\eta} + \kappa_2^{-(\eta\varepsilon-1)}\frac{\gamma^t\gamma^{\varepsilon\eta(\bar{t}-t)}}{\varepsilon(1+\eta)} & \text{if } a \geq \hat{a}_{\bar{t},t}, \end{cases}$$

where $\eta_\varepsilon \equiv \frac{\eta\varepsilon-1}{\varepsilon+1} < \eta$, $\hat{a}_{I,t} = a_t + \gamma^t\kappa_1\chi_0^{\frac{1}{\eta}}$ and $\hat{a}_{\bar{t},t} = a_t + \gamma^t\gamma^{\left(\frac{1+\varepsilon}{1+\eta}\right)\eta(\bar{t}-t)}\chi_0^{\frac{1}{\eta}}$.

Proof. See appendix. □

Lemma 6: Optimal Technology Choice

The optimal technology choice of a team (x', x) is given by

$$\alpha\left(x', x; t, \bar{t}\right) = \begin{cases} a_t + \gamma^t\kappa_1 f(x', x)^{\frac{1}{\eta}} & \text{if } f(x', x) \leq \chi_0 \\ a_t + \gamma^t\kappa_2 f(x', x)^{\frac{1}{\eta_\varepsilon}} & \text{if } f(x', x) \in (\chi_0, \chi_1(t,\bar{t})) \\ a_t + \gamma^{\bar{t}} f(x', x)^{\frac{1}{\eta}} & \text{if } f(x', x) \geq \chi_1(t,\bar{t}), \end{cases}$$

where $\chi_1(t,\bar{t}) = \chi_0\gamma^{\frac{\eta(\eta\varepsilon-1)}{\eta+1}(\bar{t}-t)}$. Also, teams with $f(x', x) \leq \chi_0$ use the local vintage, and teams with $f(x', x) \geq \chi_1(t,\bar{t})$ use the frontier vintage.

Proof. See appendix. □

Lemma 7: Properties of Optimal Technology
The optimal technology choice of a team (x', x) satisfies:

1. *each team would use a higher technology in a country closer to the frontier: for all $(x', x,)$ $\alpha(x', x; t', \bar{t}) \geq \alpha(x', x; t, \bar{t})$ if and only if $t' \geq t$;*

2. *for a given level of development, the more advanced is the frontier, the higher the technology used: for all (x', x) $\alpha(x', x; t, \bar{t}') \geq \alpha(x', x; t, \bar{t})$ if and only if $\bar{t}' \geq \bar{t}$;*

3. *the technology gap between teams in a country depends only on the distance from the frontier: $\frac{\alpha(x,y;t,\bar{t})}{\alpha(y,z;t,\bar{t})} = \frac{\tilde{\alpha}(x,y;\bar{t}-t)}{\tilde{\alpha}(y,z;\bar{t}-t)}$, where $\tilde{\alpha}(x', x; \bar{t}-t) = \alpha(x', x; t, \bar{t})\,\gamma^{-t}$; and*

4. *the cutoff $\chi_1(t, \bar{t})$ increases in the distance from the frontier.*

Proof. See appendix. □

A few comments on Lemma 7 are in order. Properties 1 and 2 show that in more-developed countries, individuals use more-advanced technologies, but at the same time, the presence of modern technology vintages at the frontier also benefits less-developed countries. The third property shows that it is not the absolute level of development that matters for the technology gap, but, rather, the distance from the technology frontier. The fourth property says that, everything else equal, the number of individuals using a modern vintage is increasing in the level of development of a country. Nonetheless, some teams in less-developed countries also use the most advanced vintage available. This property of the cost function is consistent with the fact that most modern technologies are used even in the poorest regions of the world, with the main difference across countries being that, in rich ones, many more individuals use them; as documented in Comin and Mestieri (2016). Cell phones and computers are two salient examples.

Last, and critical, Proposition 3 below shows that in countries far from the frontier, the technology gaps across teams are higher. The possibility of countries far from the frontier to choose whether to import more-advanced technology vintages from abroad leads naturally to larger gaps in the optimal technology across teams. This result emerges because not everyone finds it optimal to choose the same technology vintage. At the frontier, however, all individuals use the same vintage, the frontier one, since there is nothing better available. Putting this result together with the characterization of Section 2, and especially Proposition 2, implies that the concentration of talent is stronger in countries far from the frontier.

Proposition 3: Technology Gap and Distance to the Frontier
The technology gap increases in the distance from the technology frontier: for all $(x', y, z) \in [0,1]^3$ such that $x' \geq y \geq z$, $\frac{\alpha(x',y;t,\bar{t})}{\alpha(y,z;t,\bar{t})} \geq \frac{\alpha'(x',y;t',\bar{t}')}{\alpha(y,z;t',\bar{t}')}$ if and only if $\bar{t} - t \geq \bar{t}' - t'$.
Proof. See appendix. □

Corollary 2: Assignment of Talent and Distance to the Frontier
Both the upper and lower bounds of the ability gap between a worker and his manager are decreasing in the distance to the technology frontier.
Proof. See appendix. □

The distance of a country from the frontier is not directly observable in the data. Lemma 8 shows that we can use the GDP per capita of a country to proxy for the distance to the frontier.

Lemma 8: GDP per capita and Distance to the Frontier
The GDP per capita of a country is given by

$$Y(t,\bar{t}) = \int_0^1 g\left(\alpha\left(m(x), x; t, \bar{t}\right), m(x), x\right)\omega(x)\, dz$$

and satisfies

$$Y(t,\bar{t}) = \gamma^{\bar{t}}\tilde{Y}(\bar{t}-t),$$

where $\frac{\partial \tilde{Y}(\bar{t}-t)}{\partial(\bar{t}-t)} < 0$.
Proof. See appendix. □

3.2 A Tractable Case

Using the general functional forms in Section 2, I characterized the equilibrium in terms of bounds to the matching function. Next, I consider a functional form for net output $g(a, x', x)$ that, combined with a limit case for $c(a; t, \bar{t})$, allows me to solve for the equilibrium analytically.

I use

$$g\left(a, x', x\right) = ax'(1+\lambda x) - c(a; t, \bar{t})(1+\lambda x),$$

where $\lambda \leq 1$, and $c(a; t, \bar{t})$ is the cost function as derived in Lemma 5, with $a_t = \gamma^t$ and in the limit case when $\eta \to \infty$ and $\varepsilon\eta \to 1$.[10] Since the within-vintage cost elasticity of technology (η) goes to infinite, conditional on a technology vintage, all teams would pick identical technology a. Since the vintage cost elasticity (ε) goes to 0, – as implied by $\eta \to \infty$ and $\varepsilon\eta \to 1$ – conditional on deciding to use a foreign vintage, all teams would choose the frontier vintage. As a result, in this limit case, only two technologies are chosen. Finally, I assume that $\chi_0 \geq 0.75$, implying that more than half of the individuals use the local vintage. The functional form of g gives two other convenient properties: (i) the optimal technology choice depends only on the ability of the manager; (ii) the marginal value of both the managers and workers depends only on their partners' ability and not on their own – that is, for all (x', x'') $f_1(x', x) = f_1(x'', x)$ and $f_2(x', x) = f_2(x'', x)$. Notice that, *in equilibrium*, one's own type x does affect the marginal value of one's skills, but it does so through the matching function that assigns more-skilled managers to more-skilled workers due to

[10] Two features of g are worth notice. First, $f(x', x) = x'(1+\lambda x)$ satisfied the Spence-Mirrlees assumption described in Subsection (2.1) as long as $\lambda \leq 1$. Second, I am departing slightly from the production function defined in (2.1) to the extent that I am now allowing the cost of technology – given by $c(a; t, \bar{t})(1+\lambda x)$ – to depend on the type of workers. As I discuss below, this change is convenient for tractability. The results of Section (2) hold for this functional form.

complementarity (Lemma 2).

Characterization. Lemmas 1-4 and Proposition 1 hold for this functional form. The main difference of this tractable case is the discrete nature of the optimal choice of technology.

Lemma 9: Optimal Technology for the Tractable Case
The optimal technology for a team (x', x) is given by

$$\alpha\left(x', x\right) = \begin{cases} \gamma^{t} & x' < \chi_0 \\ \gamma^{t} + \gamma^{\bar{t}} & x' \geq \chi_0 \end{cases}.$$

Proof. See appendix. □

Therefore, the technology gap takes only two values: $\frac{\alpha(m(x),x)}{\alpha(x,m^{-1}(x))} \in \left\{1; 1+\gamma^{\bar{t}-t}\right\}$. It is equal to $1+\gamma^{\bar{t}-t}$ if $x < \chi_0$ and $m\left(x\right) \geq \chi_0$ – that is, if the ability x of an individual is such that he would use the local technology if he chooses to become a manager, but if he chooses to become a worker, he would be matched with a manager sufficiently skilled to use the frontier technology.

I now define a notion of equilibrium shape, which captures the structure of equilibrium with respect to the occupational choice sets of managers and workers.

Definition 4: Equilibrium Shape
Let $\mathbb{X}_\omega$ be the set of values of x at which the individual occupational choice changes; that is, $\mathbb{X}_\omega = \{x \in (0,1) : \lim_{\varepsilon\to 0} \omega\left(x-\varepsilon\right) \neq \lim_{\varepsilon\to 0} \omega\left(x+\varepsilon\right)\}$. Let $\mathbb{I}_\omega$ be the cardinality of such a set, $\mathbb{I}_\omega = |\mathbb{X}_\omega|$. An equilibrium shape is a sequence $\{i_0, i_1, .., i_{\mathbb{I}_\omega}\}$ where $i_0 = \omega\left(0\right)$ and for each component $\hat{x}_j \in \mathbb{X}_\omega$, ordered such that $\hat{x}_1 < \hat{x}_2 < .. < \hat{x}_{\mathbb{I}_\omega}$, $i_j = 1$ if $\lim_{\varepsilon\to 0} \omega\left(x_j+\varepsilon\right) = 1$, $i_j = 0$ if $\lim_{\varepsilon\to 0} \omega\left(x_j+\varepsilon\right) = 0$, and $i_j = \frac{1}{2}$ if $\lim_{\varepsilon\to 0} \omega\left(x_j+\varepsilon\right) \in (0,1)$.

With general functional forms, the equilibrium can take any one of infinite many shapes. The functional forms assumed in this section – by restricting the choice of technology to be discrete and the marginal value of skills to not depend on the individual's ability – limit considerably the number of attainable shapes.

Lemma 10: Possible Equilibrium Shapes
The equilibrium has, depending on parameter values, one of the following five shapes: $S_1 = (1,0)$, $S_2 = \left(1, \frac{1}{2}, 0\right)$, $S_3 = \left(1, \frac{1}{2}, 1, 0\right)$, $S_4 = \left(1, 0, \frac{1}{2}, 1, 0\right)$, $S_5 = (1,0,1,0)$.
Proof. See appendix. □

The five equilibrium shapes are illustrated in Figure 4. Keeping all other parameters constant, the shape of the equilibrium allocation and the resulting concentration of talent are pinned down by the distance of a country from the frontier. In particular, talent is segmented by occupation in countries very close to the frontier, while is segregated by technology in countries very far from it.[11]

[11]Notice, in fact, that S_1 has talent segmentation, while S_5 has talent segregation. There is one main difference with respect to Corollary 1 in Section 2: here, talent segregation is obtained without the need for the technology

Proposition 4: Equilibrium Shapes and Distance to the Frontier

Let $S(d)$ be the equilibrium shape for a country at distance $d = \bar{t} - t$ from the frontier. There exist four constants $d_1 < d_2 < d_3 < d_4$ such that if $\bar{t} - t \leq d_1$, $S(d) = S_1$; if $\bar{t} - t \in \left(d_1, d_2\right]$, $S(d) = S_2$; if $\bar{t} - t \in \left(d_2, d_3\right]$, $S(d) = S_3$; if $\bar{t} - t \in \left(d_3, d_4\right]$, $S(d) = S_4$; and if $\bar{t} - t > d_4$, $S(d) = S_5$.

Proof. See appendix. □

Corollary 3: Concentration of Talent and Distance to the Frontier

The concentration of talent increases in the distance from the technology frontier.

Proof. See appendix. □

Corollary 4: Conditions for Segregation and Segmentation in the Tractable Case

(i) If $\bar{t} - t \leq d_1$, talent is segmented by occupation; (ii) if $\bar{t} - t \geq d_4$, talent is segregated by technology.

Proof. See appendix. □

A few features of Figures 4 and 5 are worth discussing.

First, the equilibrium may dictate that some types x are indifferent between being managers or workers – that is, $w(x) = \pi(x)$ and $w'(x) = \pi'(x)$ (applying Lemma 4). How can $\pi'(x) = w'(x)$ be satisfied? The slope of the matching function $m(x)$ depends on the fraction of workers at x, $\omega(x)$, and fraction of managers at $m(x)$, $1 - \omega(m(x))$. The restriction $w'(x) = \pi'(x)$, therefore, pins down the unique value of $\omega(x)$ in the indifference region. *Second*, just as in the general case, the skill-asymmetry parameter λ plays an important role.[12] The higher is λ, the stronger the task-complementarity is and, therefore, the lower the skill asymmetry. As a result, concentration of talent increases in λ. To visualize the role of λ and d together, in Figure 5, I plot the contour plots of parameter values for which each equilibrium shape attains, as well as the corresponding values of the concentration of talent. *Third*, as the distance from the frontier (or, similarly, task complementarity) increases, the economic structure changes smoothly, in contrast to most frictionless matching models that feature a discrete jump between two polar cases.[13] How does the equilibrium evolve as $\bar{t} - t$ increases? Consider the case with segmentation of talent in Figure 4a. When a country is sufficiently close to the technology frontier, the gap between the frontier and the local technology is small; thus, every team uses a similar technology and high-skilled individuals are assigned to be managers. As $\bar{t} - t$ increases, it becomes more rewarding for an individual to become a worker and get access to the frontier technology, rather than being a manager and using the local vintage. The reward from being a manager rather than a worker depends also on the production partner and, thus, on the matching function m. At first, only the lowest-skilled among the managers finds it optimal to become workers. As $\bar{t} - t$ increases further, more and more individuals who, if they were managers, would use local technology become workers in order

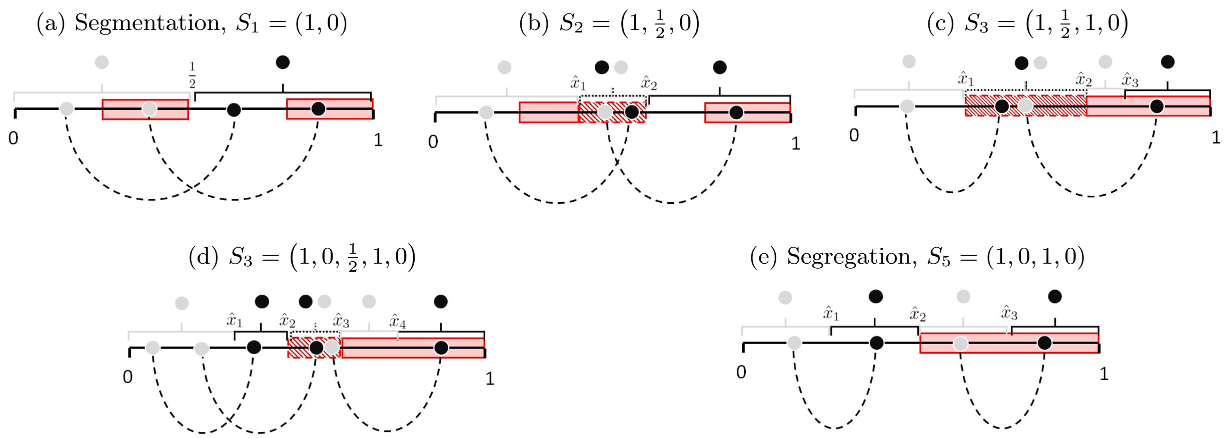
gap to go to infinite. This result is obtained because, in this limit case, the optimal technology function, $\alpha(x', x)$ is not differentiable. Also notice that this result does not contradict Corollary 1 since Corollary 1 provides sufficient but not necessary conditions for segregation and segmentation.

[12] With this functional forms, $\frac{f_1\left(x, m^{-1}(x)\right)}{f_2(m(x), x)} = \frac{1+\lambda m^{-1}(x)}{\lambda m(x)}$ that is decreasing in λ.

[13] A CES production function provides a simple example of discrete jump in matching pattern: depending on the value of the elasticity of substitution it leads to either perfectly positive or perfectly negative assortative matching.

to get access to the frontier technology. When $\bar{t} - t$ is sufficiently large, access to the frontier technology drives the assignment. Therefore, the optimal allocation resembles a dual economy: within each technology, there is talent segmentation, but skills are segregated by technology.

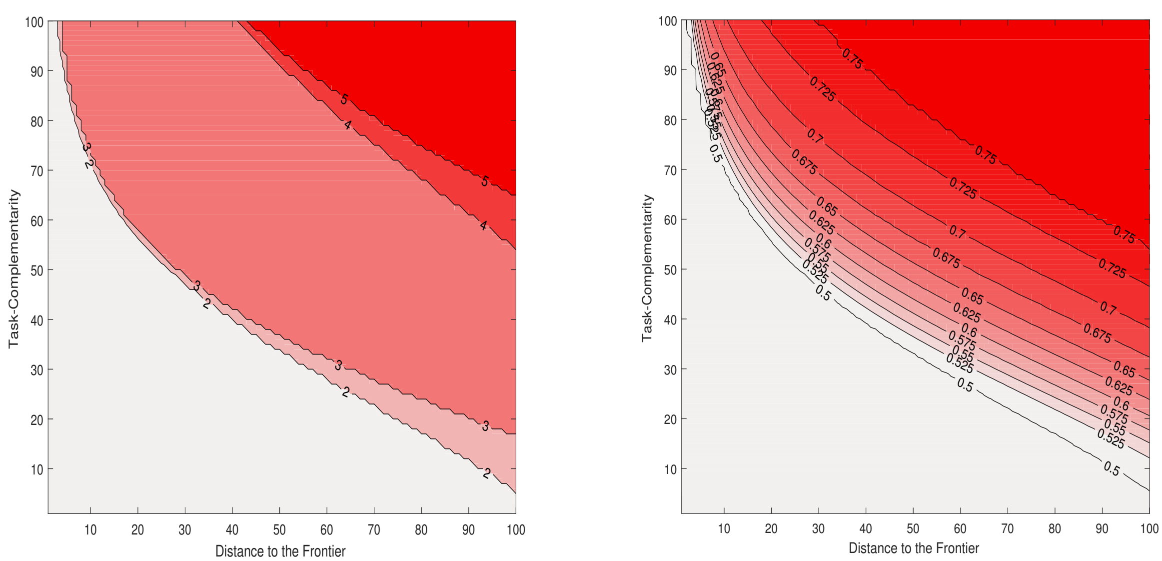
Figure 4: Five Equilibrium Shapes



Notes: the squared brackets put together individuals with the same occupation. Workers are highlighted with light grey square brackets, and managers with black ones. Dotted brackets signal mixing: some are workers and some managers. The red regions covers the set of individuals using the frontier technology vintage. The red striped regions represent mixing areas, in which the workers use the frontier technology, while the managers use the local technology. Dotted lines connect examples of workers and managers that are together in a team.

Figure 5: Roles of Task Complementarity and Distance to the Frontier

(a) Equilibrium Shape (b) Concentration of Talent



Notes: the left figure shows the contour plots of the subsets of parameter values in the space of λ (Task Complementarity) and d (Distance to the Frontier) for which each equilibrium shape attains. Specifically, the gray area highlights the set of pairs (λ, d) for which talent is segmented (S_1). The line "2" then separates the region of S_2 from the one of S_1, and so on until the last shape, S_5. The right figure shows the contour plots, in the same space, for the concentration of talent. When talent is segmented, the concentration of talent is equal to 0.5, its minimum value. When, instead, either d or λ is sufficiently high for talent not to be segmented, the concentration of talent increases smoothly in both parameters. Last, notice that both axes for λ and d are normalized on a scale from 1 to 100, even though λ only takes values in $[0, 1]$, and I use values of d between 0 and 50.

3.3 Taking Stock: Some Insights for Economic Development

In this subsection, I summarize the theoretical insights that are of interest for economic development. I also discuss how the empirical predictions of the model fit within the literature and the role that cross-country differences in the distribution of skills may play.

Role of Team Production and Distance to the Frontier. Section 2 showed that the equilibrium allocation of talent and the cross-sectional distribution of technology are inherently intertwined. Section 3 further showed that the possibility to use the frontier technology vintages in less-developed countries leads them to larger technology dispersion and, consequently, to a different allocation of talent. Specifically, in countries close to the technological frontier, most teams use a similar technology, and the allocation resembles the familiar structure from occupational choice problems: the low-skilled are workers and the high-skilled are managers. The main purpose of team production is to put together differently-skilled individuals to allow the most able ones to specialize in the most skill-sensitive task. As a result, in countries close to the technological frontier, all teams are fairly similar, and there is low productivity dispersion across them, whereas in countries far from the technological frontier, the allocation is asymmetric. Some teams attract skilled individuals (both managers and workers) and use frontier technologies, while others are left with low-skilled individuals and use traditional technologies. Teams now concentrate similarly-skilled individuals to reap the benefits of the complementarity between skills and technology. As a result, there is larger dispersion of talent, technology, and productivity in the economy. In the limit case depicted in Figure 4e, the possibility of adopting frontier technology leads to an endogenous formation of a dual economy in poor countries. Teams that adopt advanced technology attract the most-skilled individuals, leaving the rest of the economy with low talent, and, thus lower productivity.[14] Theoretically, it is even possible that some individuals would use a higher technology in autarchy than when countries gain access to the frontier, due to the polarization of talent generated by the possibility of technology adoption. In fact, when talent concentrates, low-skilled workers are matched with lower-ability managers, and thus – ceteris paribus – would use a lower technology. For the same reason, technological leapfrogging is also possible in this framework.

Empirical Predictions and Existing Literature. The model predictions on the economic structure in developing countries are qualitatively consistent with a large body of empirical evidence. The larger productivity dispersion in poor countries has been noted by, among others, Caselli (2005), Hsieh and Klenow (2009), and Adamopoulos and Restuccia (2014). The model also predicts that in developing countries, some very low-skilled individuals are employed in managerial positions. Bloom and Van Reenen (2010) shows the existence of a thick left tail of poorly managed firms and that firms with more-educated manager have better management practices. More broadly, the asymmetric equilibrium resembles a dual economy, and duality is a feature often associated with developing countries (see, for example, La Porta and Shleifer (2014)). Many existing theories that provide an explanation for these empirical facts attribute them to larger

[14] This feature resembles a mechanism outlined by Acemoglu (2015) (page 454) for the case of physical capital.

market frictions in developing countries. Instead, in the context of this paper, they emerge as a result of differences in endowment that lead to differences in optimal allocations. This paper also departs from the prior literature in linking these previously documented cross-country differences to the different assignment of individuals to teams. Differences in the allocation of talent are a new feature of economic development that has been previously overlooked. In Section 4, I show that it has empirical content.

Cross-Country Differences in Ability Distribution. The assumption that all countries have identical ability distribution – $x \sim U[0,1]$ – seems in contrast with the abundant empirical evidence that average years of schooling are lower in poor countries. I intend x to capture the relative ability rank within a country, thus comparable only within countries and not across them. The reason for this choice is that there is an intrinsic isomorphism between the level of ability x and the cost of technology. A higher ability is isomorphic to a lower cost of technology. Consider an example. Let h^t be a human capital term that captures the average ability of a country with a level of development t. Also, consider, for simplicity, the choice of technology for the tractable case of Section 3.2. Keeping the cost of technology constant and letting ability change, is identical to keeping ability fixed and letting the cost of technology change by a properly scaled factor:

$$\max_a axh^t - \frac{a^{1+\eta}}{1+\eta} \quad = \quad \max_a ax - \frac{1}{h^{(1+\eta)t}} \frac{a^{1+\eta}}{1+\eta}.$$

For this reason, the cross-country differences in the local technology vintage can be interpreted as cross-country differences in the level of human capital. I charge all cross-country differences to the cost of technology for the sake of clarity.[15]

4 Empirical Evidence on the Allocation of Talent

I document cross-country differences in the allocation of talent.

The main empirical challenge is to construct, for each country, a scalar statistic that summarizes the information in the data on the allocation of talent. To directly compute the measure of concentration of talent defined in the model, we would need to observe the ability of all individuals in the economy and their production partners. Additionally, we would need such data to be comparable for several countries around the world. These ideal data do not exist.[16]

Therefore, in the main empirical exercise, I take an indirect approach that exploits one of model's assumptions: the complementarity between skills and technology. This assumption im-

[15] Allowing the distribution of ability to vary across countries would be more problematic. The reason is that the distribution of ability may impact the matching patterns (See Kremer and Maskin (1996)). In a previous version of this paper (Porzio (2016)), I show that if individuals are allowed to invest in their ability – through schooling, for example – then the distribution of schooling will depend on the matching pattern and the distribution of technology. The stronger concentration of talent and dispersion of technology in poor countries imply a larger cross-sectional dispersion of education, as observed in the data.

[16] Matched employer-employee datasets are available for a few countries around the world. However, for less-developed countries, they are not representative of the whole economy.

plies that more-able *teams* use a more advanced technology. Observing the average ability of individuals that use each technology then becomes sufficient to make an inference on the matching function and, thus, the concentration of talent. In order to implement this strategy, we need an empirical measure of individual ability and the technology used. I use micro data from censuses and labor force surveys. In these data, I observe the education years and the working industry of each individual. I can then compute an empirical measure of concentration of talent under two assumptions: (i) education is increasing in individual ability; and (ii) the industry in which an individual works is a proxy for the technology that he uses. I will discuss potential concerns, but I first provide details on the data and on the exact construction of my empirical measure.

I also explore two alternative empirical strategies. In Section 4.5, using occupation data, I compare the average ability of managers and workers across countries. In Section 4.6, using the available firm-level data, I compare the distributions of workers to firms across countries.

4.1 Data

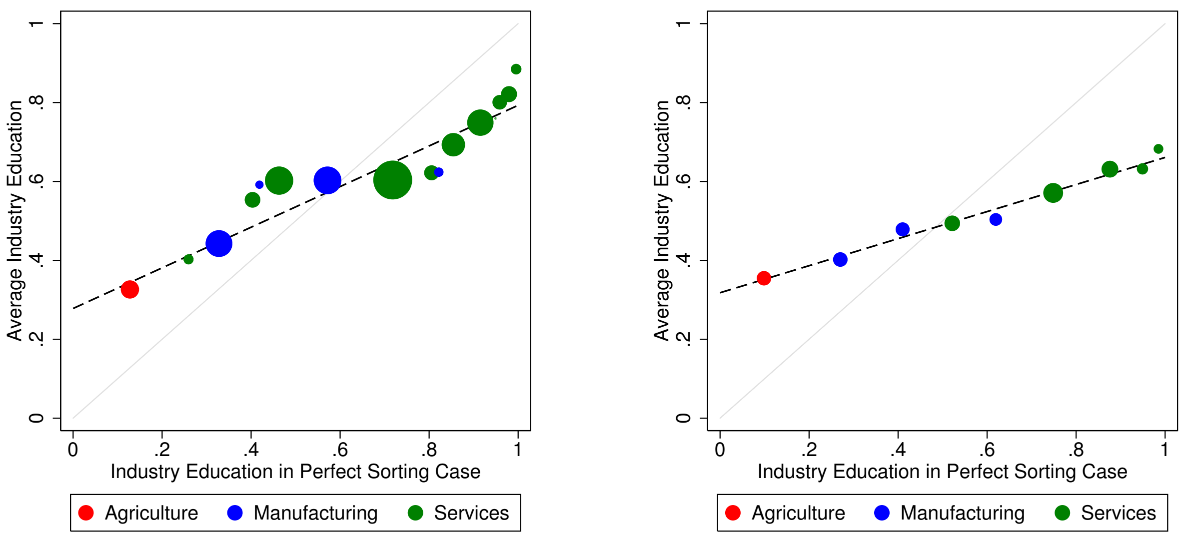
I use labor force surveys and censuses available from the Integrated Public Use Micro-data Series, International (IPUMS). The data cover 63 countries of different income levels, from Rwanda and Tanzania to Switzerland and the United States. For most countries, the datasets have a very large sample. Merging all countries and years together, I have more than 600 million individuals. Some countries have larger datasets, but – in order to avoid comparability concerns – I extract for each country-year pair a random sample of 500,000 individuals. For 13 countries, I have fewer than 500,000 individuals, but for all countries I have at least 10,000 individuals that satisfy the sample selection criteria described. Part of my analysis focuses on the South Korean growth experience, for which I use data from the Korean Longitudinal Study of Aging (KLoSA) and, to perform robustness checks, the Korean Labor and Income Panel Study (KLIPS). All GDP per capita data are taken from the Penn World Table version 8.0.

In the IPUMS data, there are three variables of interest: education, industry, and employment status. Completed years of education is coded from the educational attainment variable, and industries are standardized by IPUMS to be comparable across countries. Their industry definition spans 12 industries. Employment status indicates whether an individual is a wage-worker, own-account self-employed or an employer. Following most of the literature, I restrict the sample to include males, head of households and those between 18 and 60 years old. For the baseline results, I al exclude own-account self-employed since they do not work in teams. All data are representative of the entire population from which they are drawn. Robustness checks and alternative sample selections are in Section 4.4. Data details are in Appendix G.

KLoSA is a survey gathered with the purpose of understanding population aging in Korea. It has a sample size of approximately 10,000 individuals and represents individuals older than 45. The survey, which started in 2006, is conducted biannually. Hence, it does not allow me to directly trace the growth miracle in South Korea. However, a 2006 job supplement contains the complete employment history of each individual. In particular, this contains information on the industry in which the respondent works, his employment status, and his education. Using these data, I retroactively construct cross-sections for each year from 1953 to 2006. There is one obvious concern with this procedure: the average age of the individuals in my dataset changes over time by construction, and, thus, I may confound life-cycle and time-series trends. Robustness checks in Appendix H address this concern.

Figure 6: Construction of the Measure of Concentration of Talent

(a) Brazil, 2010 (b) United States, 1940



Notes: in each figure, I plot the average normalized measure of skills in an industry as a function of the average normalized measure of skills in a counterfactual scenario in which there is perfect sorting of individuals across industries. Each dot correspond to an industry and the size of the dot is increasing in the number of individuals employed in that industry. The dotted lines are the prediction from a linear regression weighted by the number of individuals in each industry. The slopes of the regression lines are the measures of the concentration of talent.

4.2 Measure of Concentration of Talent

I build, *separately* for each country-year dataset, the empirical measure of concentration of talent in five steps. First, I compute a normalized measure of skill: $\hat{x}_i = F(s_i)$, where s_i denotes the schooling years of individual i, and F is the country-year specific cumulative density function; to ease notation, I omit the country-year subscript, both here and in each of the following steps.[17] Second, I compute the average skill in each industry j: $\bar{\hat{x}}_j = E[\hat{x}_i | I_{ij} = 1]$, where I_{ij} is an indicator function equal to 1 if individual i works in industry j. Following the model assumption of skill-technology complementarity, I interpret industries with higher *average* education as having higher technology.[18] Third, I build a perfect sorting counterfactual in which talent is segregated by technology. Specifically, keeping industry size constant, I assign all of the most-skilled individuals to the industry with the highest average education, all the highest skilled ones among the remaining workforce to the second-highest and so on.[19] Fourth, I compute the average skill

[17]The variable s takes a only finite number of values. Therefore, I renormalize $\hat{x}$ in such a way that the lowest-skilled individuals have ability $\hat{x} = 0$ and the highest-skilled ones ability $\hat{x} = 1$. Results are robust to alternatives.

[18]The ranking of industries is, by definition, country-year specific. However, most countries display a similar ranking. For example, agriculture is consistently the less-skilled industry, while government is very high-skilled.

[19]Formally, let $\hat{\mathbb{X}}_j$ be the observed set of individuals in industry j, of mass $v\left(\hat{\mathbb{X}}_j\right)$; then, $\bar{\hat{x}}_j = E\left[\hat{x} | \hat{x} \in \hat{\mathbb{X}}_j\right]$. The counterfactual sets are given by $\hat{\mathbb{P}}_j \equiv \left\{\hat{x} : \ \hat{x} \in \left[\hat{P}_1(j), \hat{P}_2(j)\right]\right\}$ where $\hat{P}_1(j) \equiv \sum_{k:\, \bar{\hat{x}}_k < \bar{\hat{x}}_j} v\left(\hat{\mathbb{X}}_k\right)$ and $\hat{P}_2(j) \equiv$

in each industry under the perfect sorting counterfactual: $\bar{\hat{p}}_j = E\left[\hat{x}_i | I_{ij}^C = 1\right]$, where I_{ij}^C is the constructed indicator function. Fifth, and last, I regress $\bar{\hat{x}}_j = B_0 + B_1\bar{\hat{p}}_j + \varepsilon$ – weighting by the number of individuals in each industry. The measure of concentration of talent is $\hat{\pi} = \hat{B}_1$. By the definition of the least squares estimator, $\hat{\pi} = \frac{E\left[\bar{\hat{x}}_j - \bar{\hat{x}}_{j'}\right]}{\bar{\hat{p}}_j - \bar{\hat{p}}_{j'}}$. $\hat{\pi}$ captures the expected ability gap across industries relative to the benchmark case in which workers sort perfectly across industries.

In Figure 6, I show two examples: Brazil in 2010 and United States in 1940. I plot the average skill in an industry, $\bar{\hat{x}}_j$, as a function of the skill in the perfect sorting counterfactual, $\bar{\hat{p}}_j$. Each dot in the figure corresponds to an industry, and its size increases in the number of individuals employed in that industry. A linear regression $\bar{\hat{x}}_j = \alpha + \pi\bar{\hat{p}}_j + \varepsilon$ fits the data well.[20] Brazil in 2010 had GDP per capita similar to that of the U.S. in 1940; however, it had a higher concentration of talent (the regression line in the figure is steeper). This is consistent with the fact that Brazil in 2010 was farther from the technology frontier than the United States was in 1940. In Section 4.3, I build this measure of concentration of talent for each country-year pair in my sample and systematically document differences across countries depending on their distance from the frontier, but first, I further discuss its interpretation.

Link to the Model Definition of Concentration of Talent. The measure $\hat{\pi}$ captures how close the observed allocation of talent is to the case in which individuals are perfectly sorted across industries. Under the working assumption that industry can be used to proxy for technology used, we can interpret $\hat{\pi}$ as capturing how close the allocation of talent is to the case of segregation by technology. To fix ideas, consider that only two technologies are used: the local and the frontier. Assume, also, that we observe only two industries. As discussed, we measure ability with years of education. Following the model, the industry with higher average education has a larger fraction of teams using the frontier technology. For simplicity, assume that all teams in this industry use the frontier technology and that all teams in the other use the local technology.[21] The argument generalizes. Due to skill-technology complementarity, the most-skilled *teams* use the frontier technology and, hence, are in the same industry. But how are teams formed? Consider, first, the case in which talent is segmented by occupation – that is, when talent concentration is low. All managers are high-skilled, but some of them use the local technology and, thus, are in the low-technology industry. As a result, the ability gap across industries is small. Consider, next, the case in which talent is instead segregated by technology – that is, when talent concentration is high. All of the most-skilled individuals now use the frontier technology, some of them being managers and other being workers. Only the lowest-skilled individuals are left in the low-technology industry. As a result, in this case, the ability gap across industries is, instead, large. Summing up, this argument shows that the ability gap across industries captures how close the data are to either segregation by technology or segmentation by occupation.

$\hat{P}_1(j) + v\left(\hat{\mathbb{X}}_j\right)$. Then, $\bar{\hat{p}}_j = E\left[\hat{x} | \hat{x} \in \hat{\mathbb{P}}_j\right]$.

[20] The average R^2 across countries from this regression is ~ 0.9 similarly in rich and poor countries.

[21] This extreme assumption is not necessary, but we need that the set of *teams* in each industry to be connected; that is, we cannot have both the very good and very bad teams select into the same industry. This is a still a strong assumption, but in 4.4 I discuss how the failure of it would likely bias my results downward.

Formally, I can show the equivalence between the empirical and theoretical measures of concentration of talent under additional assumptions. Assume that there is a continuum of industries. Let $\tilde{a}$ be the technology rank of an industry, by construction, $\tilde{a} \sim U[0,1]$. Let $\bar{x}(\tilde{a})$ be the average ability of individuals in $\tilde{a}$. Due to the definition of $\tilde{a}$, the counterfactual average ability is equal to $\bar{p}(\tilde{a}) = \tilde{a}$. The empirical measure of concentration of talent is given by the slope in the regression of $\bar{x}(\tilde{a})$ on $\tilde{a}$. Next, assume that the average ability of workers in $\tilde{a}$ – call it $x_w(\tilde{a})$ – can be well approximated by a linear function with unknown slope: $x_w(\tilde{a}) = \beta\tilde{a}$. Last, assume that the ability gap between managers and workers is not correlated with technology. Under this set of assumptions, the empirical measure of concentration of talent is equal to β. And we can see – using market clearing – that β is also identical to the model measure of concentration of talent defined in 2.3:

$$\begin{aligned}
\frac{1}{2}\int_0^1 x_w(\tilde{a})\, d\tilde{a} + \frac{1}{2}\int_0^1 x_m(\tilde{a})\, d\tilde{a} &= \int_0^1 x dx \\
\frac{1}{4}\beta + \frac{1}{4}\beta + \frac{1}{2}\int_0^1 \left[x_m(\tilde{a}) - x_w(\tilde{a})\right] d\tilde{a} &= \frac{1}{2} \\
\beta &= 1 - \int_0^1 (m(x) - x)\,\omega(x)\, dx,
\end{aligned}$$

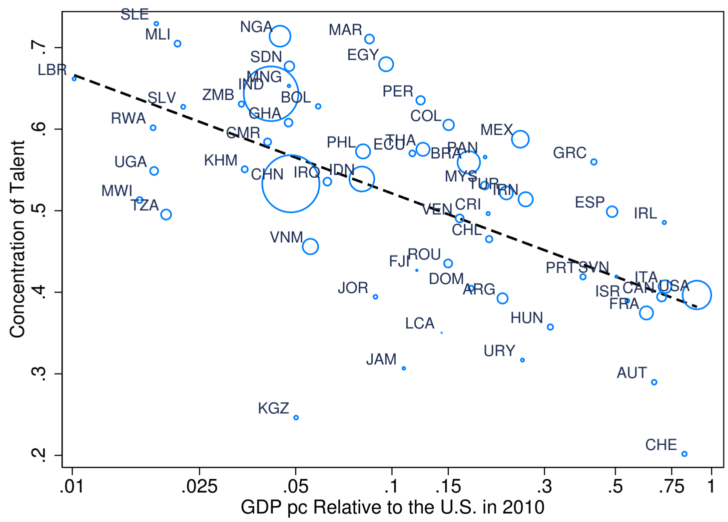
where I use the fact that $\int_0^1 x_w(\tilde{a})\, d\tilde{a} = \frac{\beta}{2}$, and that – by market clearing and by the definition of $m(x)$ – $\int_0^1 x_w(\tilde{a})\, d\tilde{a} = \int_0^1 x\omega(x)\, dx$ and $\int_0^1 x_m(\tilde{a})\, d\tilde{a} = \int_0^1 m(x)\,\omega(x)\, dx$.

4.3 Results

I now show that the concentration of talent varies systematically across countries as a function of the distance to the technology frontier. Lemma 8 in Section 3, – which characterizes the relationship between GDP per capita, the technology frontier and the country's level of development – shows that there are three cross-country comparisons to identify differences in distance to the frontier. *First*, compare two countries in the same year, but with different GDP per capita. Both countries face the same frontier; hence, the richer one is closer to the frontier. *Second*, compare two countries with the same GDP per capita, but in different years. Since the frontier grows over time, the country observed in the past was closer to the frontier than the one observed today with the same level of GDP per capita. *Third*, follow two countries over time. The one that grows faster is approaching the frontier.

1^{st} **Comparison: Poor and Rich Country Today.** In Figure 7, I plot, for countries observed in the period 2000 to 2010, my measure of concentration of talent as a function of GDP per capita relative to that of the United States in 2010. Poor countries – i.e., those farther from the technology frontier – have larger concentration of talent. In order to interpret the magnitude of cross-country differences, consider the following thought experiment. Let a country have two industries and two types of workers, high- and low-skilled. Both industries are of equal size, and half of the population is high-skilled and half is low-skilled. If $\hat{\pi}$ in this economy is equal to 0, it means that half of the high-skilled individuals are in each industry. If $\hat{\pi} = 1$, it means that all the high-skilled individuals are in one industry, which I call the modern industry. If, instead, $\hat{\pi} = \frac{1}{2}$, 75% of the high-skilled are in the modern industry, and, hence, a high-skilled individual is three times more likely to work in the modern industry. In this hypothetical world, the estimates imply that high-skilled individuals in poor countries would be approximately five times as likely to work in the modern industry as in the traditional one, while in rich countries, they would be only twice as likely. In Section 5, I build a quantitative version of the model that allows me to further interpret the magnitude of these cross-country differences.

Figure 7: 1^{st} Comparison: Cross-county Differences in Concentration of Talent



Notes: light blue circles show how populated each country is. The dotted line is the fit from a regression of concentration of talent on log of GDP per capita. The regression is not weighted by population since I treat each country as one observation. The regression line has a positive slope that is significant at the 1% level.

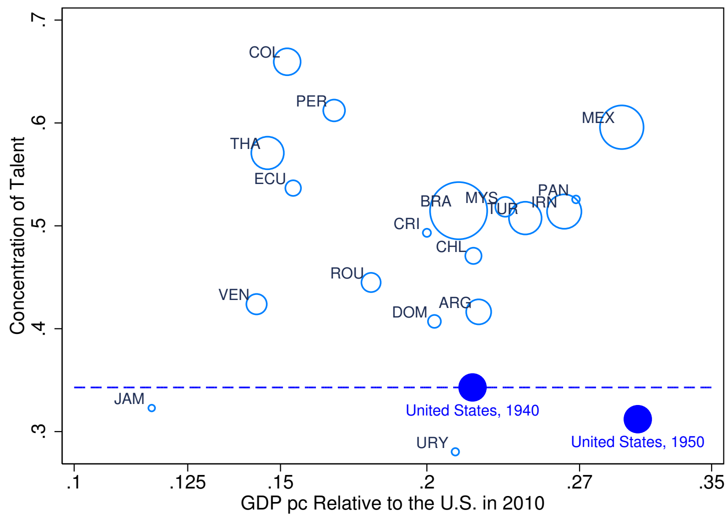
2^{nd} **Comparison: Poor Countries Today and U.S. in the Past.** We have usable census data for the United States every ten years from 1940 to 2010, since before the 1940 census, data did not report education years. GDP per capita in the United States in 1940 is comparable to that of many middle-income countries – such as Brazil, Mexico, Turkey, and Argentina – that I observe in my sample between 2000 and 2010. In fact, I observe 18 such countries and, among them, 16 have a higher concentration of talent than the U.S. used to have, as shown in Figure 8.[22] This result alleviates the concern that the observed differences might be driven by differences in the level of development rather than by the distance to the technology frontier.

3^{rd} **Comparison: South Korea Convergence to the Frontier.** Finally, I study the growth path of South Korea, a particularly interesting country since it has converged to the frontier over the past 50 years. South Korea's GDP per capita relative to that of the United States increased

[22] I have computed a similar comparison for France, for which I have data to calculate the measure of concentration of talent back to 1962. The results are very similar.

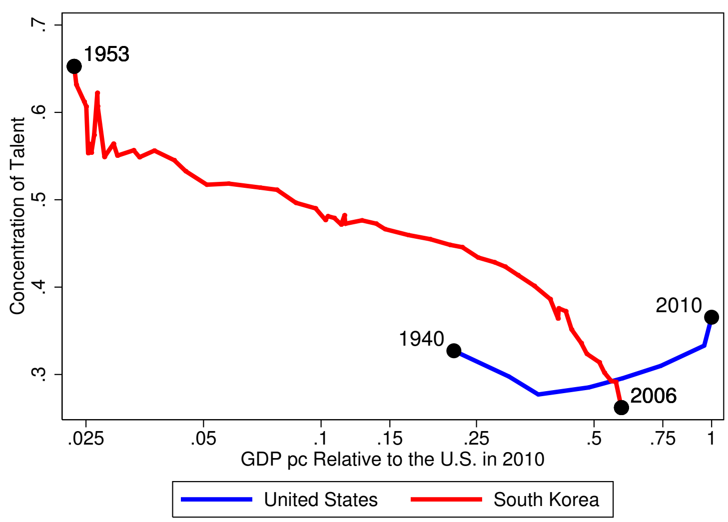
from only 7% to almost 60%. In Figure 9, I plot the growth path for concentration of talent across sectors for both countries.[23] U.S. concentration of talent remained fairly constant along the growth path, consistent with fact that the U.S. has been growing steadily over this period as a world leader – i.e., on the technology frontier. South Korea's concentration of talent, instead, decreased steeply. This comparison alleviates the concern that cross-country differences might be driven by time-invariant country characteristics that are correlated with GDP per capita.

Figure 8: 2^{nd} Comparison: Developing Countries today and U.S. in the past



Notes: light blue circles show how populated each country is. The blue dotted line is at the level of concentration of talent of United States in 1940. A permutation test of the null hypothesis that the U.S. is not different from the other countries rejects the null hypothesis more than 99% of the time.

Figure 9: 3^{rd} Comparison: South Korea as it Approaches the Frontier



Notes: the figure shows the growth paths – over time – of the concentration of talent as a function of GDP.

[23] I use concentration of talent across sectors (hence aggregating industries to agriculture, manufacturing, or services) because industry measure is not comparable for the United States and South Korea. Results with the concentration of talent across industries are comparable.

4.4 Robustness

Next, I explore the robustness of the empirical results.

The empirical strategy is sound if two assumptions hold: (i) individual ability and education years are positively correlated; and (ii) individuals that work in higher-skilled industries use, on average, a more advanced technology. One concern is that the underlying patterns of matching are identical across countries, and the observed differences are driven by mismeasurement resulting from the failure of one of the two working assumptions. I argue, however, that failures of the assumptions would most likely attenuate my results.

Assumption (i) allows me to measure skill using an individual's years of education.[24] There are few concerns. First, education might measure skill with white noise.[25] This is not a problem, however. Since I compare average skill by sector, measurement error does not bias the results: because there are a large number of individuals within each sector, this type of measurement error washes out. Second, the documented cross-country patterns can be observed if individuals are similarly matched on ability in all countries; and in poor countries, more-able individuals have more education, while in rich countries, the relationship between education and skills is non-monotonic. This hypothesis seems at odds with the often made claim that in developing countries, schooling choices are more constrained. (See, for example, Mestieri (2016)).

Assumption (ii) allows me to measure technology using the industry in which an individual works. Specifically, according to my interpretation, if an industry has a more educated workforce, it should have a more advanced technology. This is consistent with the model's prediction that more-skilled teams sort into higher production technologies. It is also consistent with the previous literature that argues that some sectors use technologies with a higher degree of skill complementarity (see, for example, Buera et al. (2015)) and that documents large productivity gaps across sectors in developing countries (see, for example, Caselli (2005), Gollin et al. (2014), and Acemoglu and Zilibotti (2001)). Nonetheless, one concern remains: my results will be biased if industry is a worse proxy for technology in countries *closer* to the frontier. My empirical results can be obtained in an hypothetical scenario in which individuals are identically matched in all countries, but in *rich* countries, dispersion of technology within industries is *larger* than in poor ones. This seems at odds with empirical evidence that documents, even in narrowly defined industries, larger dispersion of productivity in poor countries (see, for example, Hsieh and Klenow (2009) and Asker et al. (2014)).

Next, I explore robustness to alternative sample selections or measures of concentration of talent. All the results are reported in Table 1, which I refer to below, focusing on the cross-sectional comparison.

[24] A prominent alternative measure of skills is individual income. I choose to use education years for three reasons: (i) in most countries, only wage income is available, and only a fraction of the developing country workforce receives a formal wage; (ii) even when non-wage income data are available, they are hard to compare with wage income since they might capture non-individual returns (e.g., family labor); (iii) income measures are available only for a subset of countries.

[25] I use education to make an ordinal comparison across individuals within a country. As a result, any cross-country comparability concern is alleviated. For example, the concern of Hanushek and Woessmann (2008) – i.e., that cross-country education differences understate those in cognitive ability – does not apply.

Table 1: Robustness Table

		Point Estimate	R^2
(1)	**Baseline**	-0.0708	36.3%
	Level of Industry Aggregation		
(2)	Sectors (Agr, Mfg, Ser)	-0.0861	44.0%
(3)	Unharmonized Industries	-0.0472	17.3%
	Sample Selection		
(4)	Include Non Households Heads	-0.0681	35.5%
(5)	Include Women	-0.0719	37.9%
(6)	Only Women	-0.0610	15.0%
(7)	Include Self-Employed (Own-Accounts)	-0.0715	36.4%
	Role of Agriculture		
(8)	Drop Agriculture	-0.1710	55.3%
(9)	Only Individuals non in Agriculture	-0.0266	8.7%
	Measure of Concentration of Talent		
(10)	Correlation	-0.0407	24.9%
(11)	Correlation Using Normalized Skills	-0.0246	10.3%

Notes: The point estimates are from the cross-country regressions of the concentration of talent on log of GDP per capita. All Coefficients are significant at less than 1%, with the exception of rows (9) and (11) that are significant at less than 5%.

First, I explore alternative definitions of industries. I aggregate industries at the sector level (agriculture, manufacturing, services), or I use, when available, a finer definition of industries. This second alternative comes at the cost of a lack of cross-country comparability since the level of aggregation varies. The results are robust (rows 2 and 3).[26]

Second, I explore alternative sample selections. Results are robust to the inclusion of male non-household heads (row 4), or females (row 5). I then restrict the sample to *only* females (row 6). The fit is weaker, but the point estimates are significant and similar. Last, I include own-account self-employed: the result does not change (row 7).

Third, I investigate whether the results are driven mainly by the prevalence of agriculture in developing countries.[27] I address this point through two exercises. I start by recomputing the measure of concentration of talent, dropping agricultural industries in the previously described cross-industry regression used to compute $\hat{\pi}$ – i.e., $\bar{\bar{x}}_j = B_0 + B_1\bar{\bar{p}}_j + \varepsilon$. Row 8 shows that the larger measure in poor countries does not come purely from the gap between agriculture and non-agriculture, but, rather, holds also within other sectors. I then consider only individuals who are not in agriculture and recompute both the normalized measure of skill and the concentration of talent. This exercise calculates cross-country differences in concentration of talent if, suddenly, all individuals in agriculture were to drop out of the labor force. The results (row 9) are weaker and smaller in magnitude but still show more concentration in poor countries.

Fourth, and last, I compute an alternative measure of concentration of talent, along the lines

[26]Examples of Brazil 2010 and the United States 1940 are in Figures A.7 and A.8 in the Appendix G. We can appreciate that the nice fit of the measure of concentration of talent is present at any level of aggregation.

[27]The comparison between currently poor countries and the U.S. in the past already hinted towards the fact that the result cannot be driven solely by differences in agricultural share, since – Herrendorf et al. (2014) show – most countries follow a structural transformation pattern similar to that of the U.S. in the past.

of the one used in Kremer and Maskin (1996) – namely, the correlation between individuals' education and the average education in an industry:

$$\hat{\pi}_2 = \mathrm{Corr}\left(s_{ij}, \mathrm{E}\left(s_i | I_{ij} = 1\right)\right).$$

Also under this alternative measure, which is equivalent to a variance decomposition exercise, poor countries have a stronger concentration of talent. The results hold both if computed with raw education (row 10) or with normalized skills $\hat{x}$ (row 11).

4.5 Evidence from Occupation Data

IPUMS data report information on individual occupations, coded following the International Standard Classification of Occupations (ISCO). ISCO occupations do *not* correspond to the notion of managers and workers in my model. ISCO's definition of occupations seems to depend on the technology used: a manager that uses a backward technology would likely not be coded as a manager but, rather, as someone in an "elementary occupation." In fact, the occupation of manager, according to ISCO, is an occupation with skill level 4. The ISCO report, available at ilo.org, states: "Occupations at this skill level (4) generally require extended levels of literacy and numeracy, sometimes at very high level . . . and typically involve the performance of tasks that require complex problem-solving, decision-making and creativity based on an extensive body of theoretical and factual knowledge . . ." A manager of a low-technology firm in a developing country – who should be coded as a manager according to my model – does not fit well into this definition. As a result, I cannot directly use the ISCO definition of a managers to provide support for the model. Instead, I follow two alternative approaches.

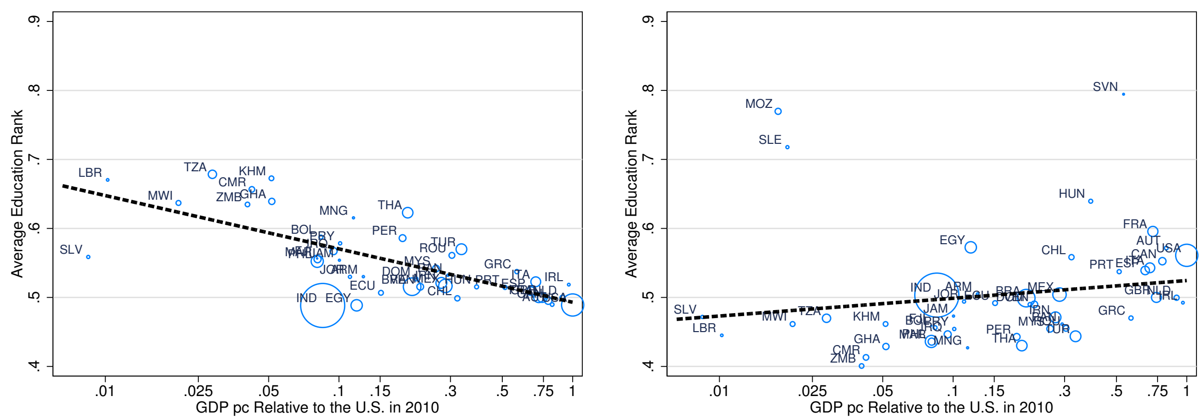
First, from the data, I build two variables that most closely resemble the model's definition of managers and workers. I code as managers individuals who are either managers according to the ISCO definition or who report being self-employed. Managers of high-technology firms are likely coded as ISCO-managers, while those at low-technology firms – which are often small, privately owned, and informal – are most likely captured in the data as self-employed. I code as workers all other employed individuals: they are wage workers and not managers according to ISCO. I then calculate, in each country, the average rank in the education distribution (my empirical measure of ability x) for managers and workers. Recall that the measure of concentration of talent in a country is given by 1 minus the average ability distance between a worker and his manager, $1 - \int \left(m\left(x\right) - x\right)\omega\left(x\right)dx$, where $\omega\left(x\right)$ is the probability that individual x is a worker. Using market clearing, I can write

$$\underbrace{1 - \int \left(m\left(x\right) - x\right)\omega\left(x\right)dx}_{\text{Concentration of Talent}} = \underbrace{1 - \int x\left(1 - \omega\left(x\right)\right)dx}_{\text{Average Manager Ability}} + \underbrace{\int x\omega\left(x\right)dx}_{\text{Average Worker Ability}}.$$

The model predicts that in poor countries, the concentration of talent is higher. Therefore, we should observe that the average ability of workers is higher in poor countries and that the average ability of managers is lower. This prediction is supported in the data, as shown in Figure 10.[28]

Figure 10: Broad Occupation Categories

(a) Workers (b) Managers



Notes: light blue circles show how populated each country is. Regression lines are not weighted by population size.

This first approach has one main limitation: the definitions of workers and managers may be not comparable across countries. Therefore, I take a second route and focus on narrowly defined occupations. In the data, an occupation is often a bundle task-technology. The model's occupation-technology[29] trade-off suggests that occupations that entail simple tasks (and, thus, would correspond to workers in the model) but use an advanced technology should be performed by relatively more skilled individuals in poor countries. Cashiers in retail stores seems to be an appropriate example. Within my model, retail cashiers are "workers" who use an advanced technology: they perform a relatively simple task within their team, working below a store or chain manager, but use an advanced technology, the cash register. Retail cashiers are comparable across countries and easy to identify in the data.[30] Moreover, they usually do not own the technology they use, which alleviates the concern that the results might be driven by binding credit constraints in poor countries. I compute average years of schooling and normalized skills of retail cashiers in each country. Figures 11a and 11b show that – consistent with the model – retail cashiers are relatively more skilled in poor countries.[31] Of course, there is a degree of

[28]The average ability of managers is only slightly lower in poor countries. This is driven by two counteracting forces. The ISCO-managers are more skilled in poor countries, while the self-employed are less skilled. On net, managers are slightly less skilled in poor countries. This additional piece of evidence shows that there is a larger dispersion of skills among managers in poor countries, consistent with the model's predictions.

[29]The model's definition of "occupation" corresponds more closely with a task, rather than with an occupation, in the data.

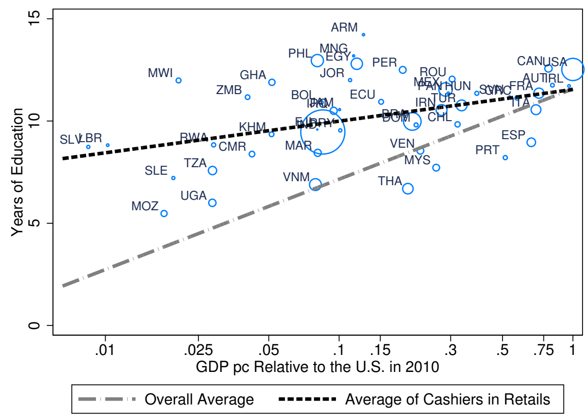
[30]In almost all countries, the codebooks for the unharmonized occupation variable encompass the category cashier. The industry variable then allows me to identify retail cashiers. Specifically, I call an individual a retail cashier if he has the variable INDGEN = 60, and he is coded as cashier in the unharmonized occupation variable (OCC in IPUMS. Cashiers are identifiable from the codebooks available for each country). I managed to code retail cashiers for 54 countries. Beware that in some countries, there are few retail cashiers. In fact, the median country has 30 observations, and the minimum number of observations is five.

[31]I can also confirm that the results are not driven by selection into retail by looking at the education of retail managers, as previously defined. Figures A.11a and A.11b show that retail managers are only mildly more skilled in poor countries.

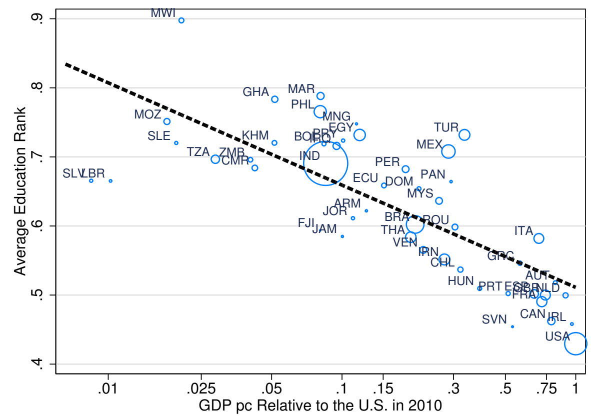
arbitrariness in the choice of occupation to focus on. Another possible alternative could be to use drivers of motor vehicles. Driving is a simple task, and a motor vehicle is an advanced technology. Figures 11c and 11d show that – again, as predicted – drivers are relatively more skilled in poor countries.

Figure 11: Narrow Occupation Categories: Retail Cashiers and Motor Vehicle Drivers

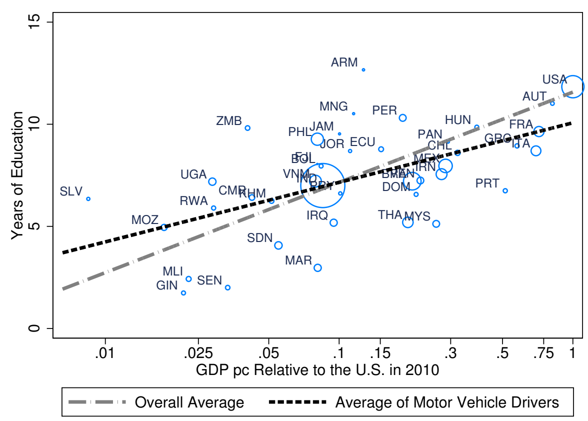
(a) Average Education of Retail Cashiers



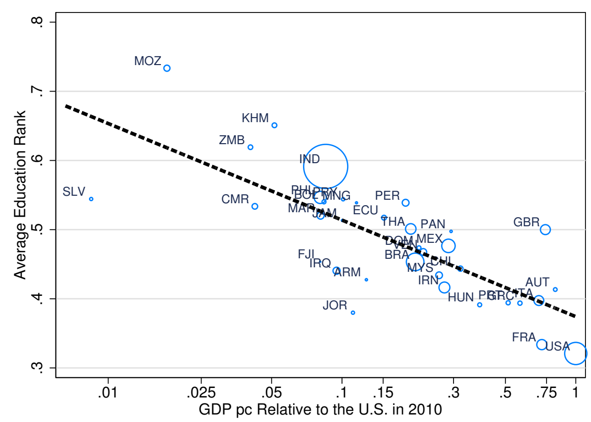
(b) Education Rank of Retail Cashiers



(c) Average Education of Motor Vehicle Drivers



(d) Education Rank of Motor Vehicle Drivers



Notes: light blue circles show how populated each country is. Regression lines are not weighted by population size.

4.6 Evidence from Firm Data

The World Bank Enterprise Surveys provide firm-level information from several countries around the world. Specifically, they report the average education of workers and managers in the firm, and measures of technology and labor productivity. Unfortunately, these data suffer from major cross-country comparability concerns; as noted by Asker et al. (2014), for example. Most importantly, they cover only registered firms, which constitute a much larger fraction of the labor share in rich countries. Despite these concerns, these data can be used to provide further validation for the model. I do so in detail in Appendix E. I interpret a team in the model as a firm in the data. The three main findings are as follows. First, I validate the main prediction of the assumptions of skill-technology and manager-worker complementarity. I show that, similarly in rich and poor countries, more-educated workers are employed in firms with higher technology (measured by computer usage), higher labor productivity, and better-educated managers. Second, I show that, consistent with Proposition 3, dispersion of used technology across firms is larger in less-developed countries. Third, I show that talent, measured as years of education, is more concentrated within firms in less-developed countries, as suggested by Corollary 2.

5 A Quantitative Exploration

As discussed in Section 3.3, the equilibrium in countries far from the technology frontier, despite being efficient, predicts several empirical features that are usually attributed to the presence of larger frictions in developing countries. In view of this result, cross-country exercises that compare micro-level allocations should be cautious in assigning all the observed differences to larger frictions in developing countries.

In this section, I take a first step in understanding whether the mechanism that this paper proposes could be quantitatively relevant in explaining cross-country differences. I ask whether the model can explain a sizable fraction of the larger agricultural productivity gaps documented in poor countries (see Caselli (2005)). Specifically, I write down a version of the model that is amenable to a quantitative exploration of the data. Next, I estimate it using data on the allocation of talent in rich and poor countries and the agricultural productivity gap in rich countries. I then show that, once fitted into the model, the observed cross-country differences in allocation of talent reconcile approximately one third of the larger agricultural productivity gaps in poor countries.

5.1 Quantitative Model

I add one feature to the model in Section 2. Each individual is characterized not only by his ability x, but also by a vector of technology-specific shocks, which make him more willing to use some technologies rather than others. The introduction of this additional source of noise allows me to develop an efficient numerical algorithm and to make the model amenable for estimation. Apart from this feature, the quantitative model mirrors the one presented in Section 2.

Each individual x chooses the technology that gives him the highest value. He solves

$$\max_{a \in \mathbb{A}} \upsilon_a y\left(a, x\right),$$

where $\mathbb{A} = \{a_1, .., a_N\}$ is a discrete approximation of the set of available technologies; υ_a is distributed according to a Frechet with dispersion parameter θ (dispersion of shocks increases as θ decreases)

$$\upsilon_a \sim e^{-\upsilon_a^{-\theta}};$$

and $y\left(a, x\right)$ is the income of an individual of ability x that uses technology a and picks his occupation optimally

$$y(a,x) = \max_{z\in[0,1]} z\pi(a,x) + (1-z)\,w(a,x).$$

Managers choose the optimal worker among the set Ω_a that gathers the individuals that are using technology a:

$$\pi(a,x) = \max_{z\in\Omega_a} g(a,x,z) - w(a,z). \tag{3}$$

The Spence-Mirrlees assumption (Assumption 3 in Section 2) guarantees that, as long as individuals use the same technology, the most-skilled ones are managers. Market clearing dictates that the mass of manager and workers for each technology should be identical. These two conditions allow me to solve – given $\phi(a,x)$, which is the joint distribution of individuals over technology a and ability x – for the cutoff types $\hat{x}(a)$ that characterize the occupational choice in each technology:

$$\int_0^{\hat{x}(a)} \phi(a,x)\,dx = \int_{\hat{x}(a)}^1 \phi(a,x)\,dx.$$

The income of an individual x that uses technology a is then given by

$$y(a,x) = \begin{cases} w(a,x) & \text{if } x \leq \hat{x}(a) \\ \pi(a,x) & \text{if } x > \hat{x}(a). \end{cases}$$

The matching function within each technology, $m(a,x)$, is derived using the fact that, due to the complementarity of f, there is positive assortative matching between managers and workers. $m(a,x)$ solves for all (x,a)

$$\int_{\hat{x}(a)}^{m(a,x)} \phi(a,z)\,dz = \int_0^x \phi(a,z)\,dz.$$

Wages and profits are calculated using the first-order and envelope conditions of the manager's problem (3), together with market clearing $\pi(a,m(x)) + w(a,x) = af(m(x),x) - c(a)$. That is,

$$\begin{aligned} w(a,x) &= \kappa + \int_0^x w_2(a,z)\,dz \\ \pi(a,x) &= w(a,\hat{x}(a)) + \int_{\hat{x}(a)}^x \pi_2(a,z)\,dz, \end{aligned}$$

where κ satisfies

$$\int_0^{\hat{x}(a)} w(a,x)\,\phi(a,x)\,dx + \int_{\hat{x}(a)}^{1} \pi(a,x)\,\phi(a,x)\,dx \quad = \quad \int_0^{\hat{x}(a)} g(a,m(a,x),x)\,\phi(a,x)\,dx.$$

The properties of the Frechet distribution (see, for example, Hsieh et al. (2016)) provide a simple equation that characterizes, together with market clearing, the distribution of individuals over technologies, $\phi(a,x)$. Specifically, the probability that an individual x selects into technology a is

$$\frac{\phi(a,x)}{\sum_{a\in\mathbb{A}}\phi(a,z)} \quad = \quad \frac{\tilde{y}(a,x)^{\theta}}{\sum_{a\in\mathbb{A}}\tilde{y}(a,x)^{\theta}}, \tag{4}$$

where $\tilde{y}(a,x) = \max\{0, y(a,x)\}$. Last, the GDP per capita is

$$Y \quad = \quad \sum_{a\in\mathbb{A}} \int_0^{\hat{x}(a)} g(a,m(a,x),x)\,\phi(a,x)\,dx. \tag{5}$$

This model can be computed for any functional form for net output g – hence, for f and c – that satisfies the assumptions in Section 2. The computing algorithm iterates on the distribution $\phi(a,x)$ until convergence. Details are in Appendix F. The numerical solution is fast, thus allowing me to estimate the model through a simulated method of moments, as I discuss next.

5.2 Model Estimation

I next describe how I estimate the model. Six steps are involved. First, I pick parsimonious functional forms that give five parameters to be structurally estimated. Second, I describe the targeted moments in the data. Third, I show how to construct equivalent moments in the model. Fourth, I provide some details on the estimation method. Fifth, I show the fit of the model and the external validation exercises. Sixth, I comment on the identification of the parameters.

Functional Forms and Parameters. I need to pick a functional form for net output g – hence, for f and c. I choose parsimonious functional forms to limit the number of parameters to be estimated. I use $f(x',x) = x'(1+\lambda x)$, where $\lambda \in [0,1]$ modulates the strength of the gap in skill sensitivity between managers and workers: $\frac{\partial}{\partial\lambda}\frac{f_1(x,y)}{f_2(y,z)} < 0$, for all (x,y,z). $\lambda \leq 1$ guarantees that the Spence-Mirrlees condition is satisfied. For the cost of technology, I use $c(a;t,\bar{t})$, as defined in Lemma 5 in Section 3. $c(a;t,\bar{t})$ is modulated by three main parameters: (i) η, the cost elasticity within technology vintages; (ii) ε, the cost elasticity across technology vintages; and (iii) χ_0, the level of the cost of importing foreign technology vintages. Additionally, I fix the non-homotheticity term ν equal to 1, just as in the tractable case of Section 3.2, and I fix $\gamma = 1.02$, which corresponds to 2% yearly growth at the frontier. The value of γ does not affect the results since it is not separately identified from $\bar{t}$ and t. Intuitively, two years away from the frontier at 2% growth per year is the same as one year away at 4.04% growth per year. As I showed in Section 3, only the distance to the frontier, $\bar{t} - t$, matters for the technology gaps and the allocation of talent: thus, I can fix $\bar{t}$ without loss of generality. I choose $\bar{t}$ such that a country on the frontier – hence, with $t = \bar{t}$ – has the level of GDP per capita of the United States.

Summing up, the production function g gives four parameters to be estimated, three for c and one for f. Additionally, we have the new parameter θ, which modulates the dispersion of the technology shocks. Therefore, I estimate five parameters within the model. Given the vector of these parameters, I can compute the equilibrium for countries at different distances to the frontier – i.e., for different values of t – and study how the allocation of talent and agricultural productivity gaps vary as a function of GDP per capita.

Data Moments. I target three sets of moments: (i) the empirical relationship between GDP per capita and concentration of talent; (ii) the agricultural productivity gap in countries close to the frontier; (iii) the percentage of agricultural workers by skill level in countries close to the frontier. For (i), I use the empirical results of Section 4, and, specifically, I target the values of concentration of talent as a function of log GDP predicted by the regression line in Figure 7. For (ii), I use the average agricultural productivity gap – i.e., the ratio between labor productivity in non-agriculture and agriculture – for the first quartile of GDP per capita, as reported in Gollin et al. (2014). For (iii), I compute the fraction of workers in agriculture, separately for different skill brackets, using the same IPUMS data as described in Section 4. For each country, I compute the normalized measure of skills x, just as described in 4, and calculate for each skill decile the fraction of people that work in agriculture. I then divide the countries in my dataset into four groups to match the average GDP of the four quartiles of Gollin et al. (2014) – the first one has GDP per capita that is 78% of the United States'; the second 21%; the third 8.7%; and the fourth 2.5% – and take the average within each group. The result of this procedure is shown in Figure 12b: rich countries have a smaller fraction of workers in agriculture, and in all countries, low-skilled people are more likely to work in agriculture. These result are consistent with evidence in the previous literature; see, for example, Caselli (2005) and Caselli and Coleman (2001).

Model Moments. First, notice that the equilibrium objects described in Section 5.1 are functions of the vector of parameters, which I define $\Theta \equiv \{\eta, \varepsilon, \chi_0, \lambda, \theta\}$, and the level of development t. I now make this dependence explicit in my notation. For example, I use $Y(t, \Theta)$ for GDP per capita, defined in equation (5), rather than simply Y.

I treat the model-generated data following the same procedure that I used for the micro data, and according to the same assumptions. I next describe in detail how I construct the moments.

For a given vector of parameters Θ, I first find $\hat{t}$ such that $\frac{Y(\hat{t},\Theta)}{Y(\bar{t},\Theta)} = 1\%$, where $Y(\bar{t}, \Theta)$ is the GDP per capita of the United States. I then compute the model for 100 countries with equally spaced values of t in $[\hat{t}, \bar{t}]$. The result is a model-generated dataset of 100 countries with GDP per capita between 1% and 100% of that of the frontier – i.e., the United States. For each country $t \in \{\hat{t}, ..., \bar{t}\}$, I compute the relative GDP per capita: $\frac{Y(t,\Theta)}{Y(\bar{t},\Theta)}$. I then divide the set of 100 countries into four groups such that their average relative GDP per capita matches the four quartiles from Gollin et al. (2014). That is, I create $\mathbb{T}_1 = \{\hat{t}, ..., t_{k_1}\}$, $\mathbb{T}_2 = \{t_{k_1+1}, ..., t_{k_2}\}$,

$\mathbb{T}_3 = \{t_{k_2+1}, ..., t_{k_3}\}$, $\mathbb{T}_4 = \{t_{k_3+1}, ..., \bar{t}\}$ such that $\frac{1}{|\mathbb{T}_1|}\sum_{j=1}^{k_1} \frac{Y(t_j,\Theta)}{Y(\bar{t},\Theta)} = 2.5\%$, $\frac{1}{|\mathbb{T}_2|}\sum_{j=k_1+1}^{k_2} \frac{Y(t_j,\Theta)}{Y(\bar{t},\Theta)} = 8.7\%$, $\frac{1}{|\mathbb{T}_3|}\sum_{j=k_2+1}^{k_3} \frac{Y(t_j,\Theta)}{Y(\bar{t},\Theta)} = 21\%$, $\frac{1}{|\mathbb{T}_4|}\sum_{j=k_3+1}^{100} \frac{Y(t_j,\Theta)}{Y(\bar{t},\Theta)} = 78\%$. These quartiles are useful later.

To calculate the concentration of talent, I use the procedure described in Section 4, but with model-generated data. I keep the assumption that an industry is a technology, and I calculate the average ability of individuals that use each technology a

$$\bar{x}(a;t,\Theta) = \frac{\int_0^1 x\phi(a,x;t,\Theta)\,dx}{\int_0^1 \phi(a,x;t,\Theta)\,dx}.$$

I then calculate the counterfactual average ability under perfect sorting

$$\bar{p}(a;t,\Theta) = \sum_{\tilde{a}=a_1}^{a_{-1}} \left[\int_0^1 \phi(\tilde{a},x;t,\Theta)\,dx\right] + \frac{\int_0^1 \phi(a,x;t,\Theta)\,dx}{2},$$

where a_{-1} is the technology just worse than a. The measure of concentration of talent, which I call $\pi^m(t,\Theta)$, is given by the coefficient $\hat{\beta}_1$ from the regression $\bar{x}(a;t,\Theta) = \beta_0 + \beta_1\bar{p}(a;t,\Theta) + \varepsilon$.

I next need to define a notion of agriculture in the model. In all countries, agriculture is the industry with the lowest average education. Therefore, following the assumptions of Section 4, I construct agriculture in the model as the partition of teams that use the least-advanced technologies. I choose such a mass of teams to match the agricultural employment in the data for the countries in the same quartile of GDP. Agriculture is, thus, the partition of teams that produce with technologies in the set $\mathbb{A}^{agr}(t,\Theta) = \{a_1, .., a_n\}$, where a_n is such that[32]

$$\sum_{a\in\mathbb{A}^{agr}(t,\Theta)} \int_0^1 \phi(a,x;t,\Theta)\,dx \simeq L^{agr}(t),$$

and if country t is in quartile x – i.e., $t \in \mathbb{T}_x$ – then $L^{agr}(t) = \hat{L}_x^{agr}$, where $\hat{L}_x^{agr}$ is the average share of employment in agriculture measured in the IPUMS data for countries in the x^{th} quartile. Output in agriculture and non-agriculture is, thus, given by

$$Y^{agr}(t,\Theta) = \sum_{a\in\mathbb{A}^{agr}(t,\Theta)} \int_0^{\hat{x}(a;t,\Theta)} g(a, m(a,x;t,\Theta), x; t,\Theta)\,\phi(a,x;t,\Theta)\,dx \tag{6}$$

$$Y^{nagr}(t,\Theta) = \sum_{a\in\mathbb{A}^{nagr}(t,\Theta)} \int_0^{\hat{x}(a;t,\Theta)} g(a, m(a,x;t,\Theta), x; t,\Theta)\,\phi(a,x;t,\Theta)\,dx, \tag{7}$$

where $\mathbb{A}^{nagr}(t,\Theta) = \mathbb{A}\backslash\mathbb{A}^{agr}(t,\Theta) = \{a_{n+1}, ..., a_N\}$. The agricultural productivity gap, $Z(t,\Theta)$, is simply, as in the data, the ratio between the labor productivity in non-agriculture and in

[32]Since $\mathbb{A}$ is a discrete set, in general, there is no a_n such that this equation holds with exact equality. For sufficiently large $|\mathbb{A}|$, the approximation error is small. To avoid any concern, I find a_n and then randomly assign to agriculture a fraction of teams of the closest technology in such a way that the model exactly matches the fraction of agricultural workers.

agriculture; that is,

$$Z(t,\Theta) \;=\; \frac{\frac{Y^{nagr}(t,\Theta)}{L^{nagr}(t)}}{\frac{Y^{agr}(t,\Theta)}{L^{agr}(t)}}.$$

Finally, the fraction of people employed in agriculture by skill level is

$$\rho(x;t,\Theta) = \frac{\sum_{a\in\mathbb{A}^{agr}(t,\Theta)} \phi(a,x;t,\Theta)\,dx}{\sum_{a\in\mathbb{A}(t,\Theta)} \phi(a,x;t,\Theta)\,dx}.$$

Estimation Method. For a given vector of parameter Θ, I compute the following measure of the distance between the model and the data

$$\mathbb{L}(\Theta) = \frac{1}{100}\sum_{j=1}^{100}\left(\frac{\pi^m(t_j,\Theta)-\pi^d(t_j)}{\pi^d(t_j)}\right)^2 + \frac{1}{10}\sum_{s=1}^{10}\left(\frac{\rho^{m,1}(x_s;\Theta)-\rho^{d,1}(x_s)}{\rho^{d,1}(x_s)}\right)^2 + \left(\frac{Z^{m,1}(\Theta)-Z^{d,1}}{Z^{d,1}}\right)^2$$

where a few objects have not been defined yet: $\pi^d(t_j)$ is the targeted empirical value of the concentration of talent for a country t_j; $\rho^{m,1}(x_j;\Theta)$ is the average fraction of individuals with a skill level in decile s that are employed in agriculture for the model-generated countries in the first quartile – i.e., $\rho^{m,1}(x_s;\Theta) = \frac{1}{|\mathbb{T}_1|}\sum_{t\in\mathbb{T}_1}\rho(x_s;t,\Theta)$; $\rho^{d,1}(x_s)$ is the same object in the data; $Z^{m,1}(\Theta)$ is the average agricultural productivity gap for countries in $\mathbb{T}_1$ – i.e., $Z^{m,1}(\Theta) = \frac{1}{|\mathbb{T}_1|}\sum_{t\in\mathbb{T}_1} Z^m(t,\Theta)$; and $Z^{d,1}$ is the corresponding empirical value from Gollin et al. (2014). The estimated vector of parameters is $\Theta^* = \arg\min \mathbb{L}(\Theta)$. To solve this minimization problem, I use the simulated method of moments. Specifically, I simulate, using the Metropolis–Hastings algorithm, a Markov chain that converges to the vector of parameters that minimizes the distance between the model and the data.[33]

Model Fit and Validation. The model fits the data fairly well. Figure 12a shows the cross-country relationships between concentration of talent and GDP per capita in the model and in the data. The dotted line represents the targeted values, π^d, while the red crosses are the model-generated data, $\pi^m(\cdot,\Theta^*)$. Figure 12b shows agricultural employment by skill, separately for each GDP quartile. The targeted data, $\rho^{d,1}$, are the points of the darkest line. The model data, $\rho^{m,1}(\cdot,\Theta^*)$, are the darkest crosses. The model matches the average agricultural employment by construction, but the estimation targets the heterogeneity by skills. The model's and the data's agricultural productivity gaps, $Z^{m,1}(\Theta^*)$ and $Z^{d,1}$, respectively, are shown in rows (4) and (7) of the first column of Table 2.

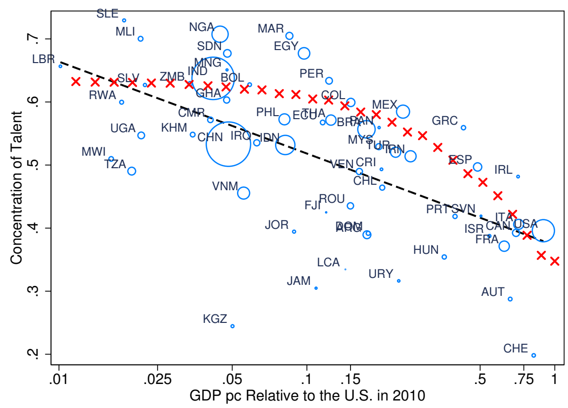
I also provide two validation exercises. First, notice that in Figure 12b, the relationships between skills and agricultural employment for GDP quartiles 2, 3, and 4 are not targeted. Nonetheless, the model matches these empirical relationships. Second, and more remarkably, the model predicts well the share of individuals in each country that use the frontier technology vintage. I construct, using the CHAT dataset (Comin and Hobijn (2009)), the number of internet users per

[33] For details on the simulated method of moments, see McFadden (1989). The simulation method that I use was developed first by Chernozhukov and Hong (2003) and was recently used in Lise et al. (2016).

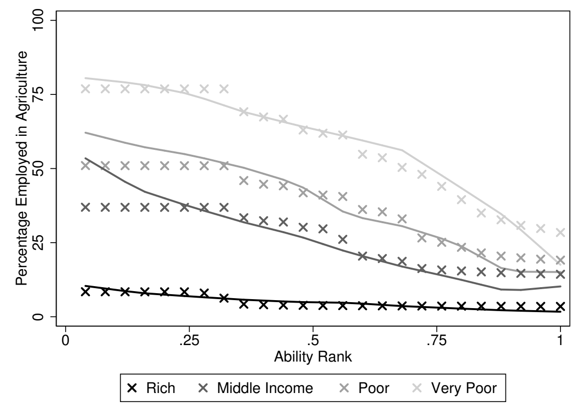
capita for several countries in 2005, when the world wide web was arguably a frontier technology. The relative fraction of internet users across the world should correspond to the model's relative fractions of individuals that use the frontier vintage.[34] I plot the data along with the model's predictions in Figure 12c.

Figure 12: Model Fit and Validation

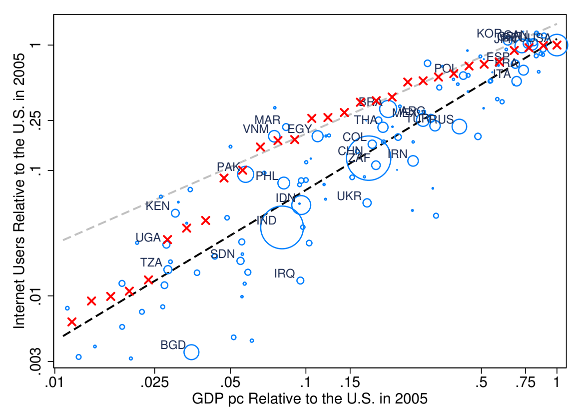
(a) Concentration of Talent



(b) Ability Distribution in Agriculture



(c) Penetration Rate of Internet Users



Notes: the top left panel plots the empirical measure of concentration of talent, calculated as described in Section 4, and the same statistic computed in the model (red crosses) as a function of relative GDP per capita. The top right panel shows the average fractions of people employed in agriculture as a function of their skills, separately for the four quartiles of GDP per capita. The solid lines are the data. The crosses are the values computed in the model. The bottom panel plots the fraction – in 2005 – of internet users relative to the U.S., calculated from the CHAT dataset together with the relative fraction of individuals – computed in the model – that use the frontier technology vintage (red crosses) as a function of relative GDP per capita. In this same figure, the black dotted line is the fit line (of relative internet users on relative GDP per capita) for 2005, while the gray dotted line is the fit line of the same regression but with data from 2010.

[34]Internet users captures only the extensive margin and not the intensive margin of technology usage. I've computed similar results for other technologies in the CHAT dataset, such as number of computers, and number of cell phones. The results look similar but are harder to interpret because they confound the number of users with the amount of technology per person.

Model and data align fairly well. The model, however, slightly overestimates the fraction of internet users relative to the data. In the same figure, I also plot the empirical relationship for 2010. The model underestimates the relative fraction of internet users when compared to 2010. This is consistent with the world wide web becoming less of a frontier technology over the period 2005 to 2010.

Comments on Identification. The five parameters are jointly determined by the targeted moments. Nonetheless, I discuss intuitively the main links between moments and parameters. λ decreases the skill asymmetry, hence increases the concentration of talent. η decreases the technology gap for teams that use the same technology vintage, hence decreases the concentration of talent. Additionally, η affects the agricultural productivity gap since a higher η implies a lower agricultural gap for a given team composition. In countries close to the frontier, η and λ are the main determinants of the concentration of talent since most teams use the same vintage. Therefore, λ and η together are relevant mainly in matching the concentration of talent and agricultural productivity gap close to the frontier. ε affects how much the concentration of talent increases as a function of the distance from the frontier. In fact, the higher is ε, the costlier it is for countries with a low local technology vintage to upgrade to the next available vintage. As a result, the higher is ε, the more similar are the technologies used and the lower is the concentration of talent. χ_0 affects the cost of importing technology vintages and, thus, the mass of people in each country that decide to not use the local technology. The higher is χ_0, the higher the decrease in GDP as we decrease the level of development t. Therefore, ε and χ_0 together match mainly the relationship between the change in concentration of talent and the change in GDP per capita. Last, the dispersion of technology shocks θ mainly pins down how many high-skilled individuals work in agriculture relative to non-agriculture. Consider that high-skilled individuals have, ceteris paribus, a comparative advantage in advanced technology. Therefore, if we observe many of them in agriculture, it must be that the dispersion of the technology shock is large, thus pushing many high-skilled into agriculture.

5.3 Results on Agriculture Productivity Gaps

I next use the estimated model to study the extent to which the larger agricultural productivity gaps in poor countries are driven by the interaction of the endogenous allocation of talent and the possibility of importing advanced technology from the frontier. I compute the model-generated agricultural productivity gap for each quartile of the income distribution – that is, $Z^{m,x}(\Theta^*) = \frac{1}{|\mathbb{T}_x|}\sum_{t\in\mathbb{T}_x} Z^m(t,\Theta^*)$ for $x \in \{1,2,3,4\}$ – and compare them with the empirical ones reported in Gollin et al. (2014). Results are in rows (4) and (7) of Table 2. The model, once disciplined by cross-country differences in the concentration of talent, reconciles a sizable amount, approximately one third, of the higher agricultural productivity gap in developing countries.[35] Therefore, a naive cross-country comparison that takes as null the hypothesis that the agricultural productivity gap should be identical, would overstate the extent to which market frictions are larger in developing countries. At the same time, even after accounting for endogenous technology choice and team

[35]The value $\frac{1}{3}$ is calculated as follows: the total increase in the data from the first to the fourth quartile is $3.6 = 5.6 - 2$; the same value in the model is, instead, $1.3 = 3 - 1.7$; and $\frac{1.3}{3.6} \simeq \frac{1}{3}$.

formation, more than half of the cross-country differences remain unexplained, thus still leaving a prominent role to larger market failures in developing countries or to any other competitive explanations that are not present in this model.

Role of Endogenous Allocation of Talent. The role of endogenous allocation of talent is twofold. First, it permits me to discipline the model: the predicted larger agricultural productivity gaps in poor countries are modulated by their larger empirically measured concentration of talent. In fact, if I estimate the model targeting identical concentration of talent in all countries, then the predicted agricultural productivity gaps are similar across countries – the remaining differences are driven by the country-specific mass of people employed in agriculture – and actually slightly smaller in countries far from the frontier. Second, the larger concentration of talent in countries far from the frontier amplifies their larger productivity dispersion. I explore the quantitative relevance of this amplification mechanism by computing counterfactual agricultural productivity gaps in a scenario in which all countries have the output function g of the average country in the first quartile, but retain their country-specific concentration of talent. That is, I replace $g\left(a, m\left(a, x; t, \Theta^{*}\right), x; t, \Theta^{*}\right)$ with $g\left(a, m\left(a, x; t, \Theta^{*}\right), x; \bar{t}^{1}, \Theta^{*}\right)$, where $\bar{t}^{1} = \frac{1}{|\mathbb{T}_1|} \sum_{t \in \mathbb{T}_1} t$, when I calculate agricultural and non-agricultural outputs using equations (6) and (7), and then I use these value to calculate the agricultural productivity gaps. The results are shown in row (8) of Table 2. The allocation of talent alone explains almost 40% of the difference in agricultural productivity between the first and the fourth quartiles.[36]

Table 2: Allocation of Talent and Agriculture to Non-Agriculture Productivity Gaps

			Quartiles of World Income Distribution			
			Q1	*Q2*	*Q3*	*Q4*
	(1)	GDP pc Relative to U.S.	78%	21%	8.7%	2.5%
Data	(2)	Concentration of Talent	38%	47%	53%	62%
	(3)	Internet Users Relative to U.S.	79%	28%	9%	2%
	(4)	Agricultural Productivity Gap	2	3.2	3.4	5.6
Model	(5)	Concentration of Talent	40%	57%	61%	63%
	(6)	Internet Users Relative to U.S.	92%	41%	16%	2%
	(7)	Agricultural Productivity Gap	1.7	2	2.4	3
	(8)	Productivity Gap with Fixed g	1.7	1.7	1.8	2.2

[36]Notice that while the concentration of talent increases between quartiles 1 and 2, the agricultural productivity gap does not increase significantly. Why does this happen? Recall that the agricultural productivity gap depends, in this counterfactual, only on the skills of people in agriculture and in non-agriculture. In quartiles 1 and 2, a small fraction of the overall population works in agriculture. As a result, although the concentration of talent increases due to changes within non-agriculture, these changes are not sufficiently large to sizably affect the agriculture to non-agriculture gap.

Robustness. In order to explore whether the results are robust to perturbations around the value of Θ^*, I compute the agricultural productivity gaps averaged across other "good fit" parameter sets from the chain simulated for the estimation. Specifically, letting $\mathbb{H} = \left\{\Theta : \frac{\mathbb{L}(\Theta)-\mathbb{L}(\Theta^*)}{\mathbb{L}(\Theta^*)} < 1\%\right\}$, I compute $\bar{Z}^{m,x} = \frac{1}{|\mathbb{H}|}\sum_{\Theta\in\mathbb{H}}\left[\frac{1}{|\mathbb{T}_x|}\sum_{t\in\mathbb{T}_x} Z^m(t,\Theta)\right]$, for $x \in \{1,2,3,4\}$. This exercise gives: $\bar{Z}^{m,1} = 1.8$, $\bar{Z}^{m,2} = 2.1$, $\bar{Z}^{m,3} = 2.7$, $\bar{Z}^{m,4} = 3.6$. The results are, therefore, similar and, if anything, stronger than those for the point estimate Θ^*.

6 Extensions of the Analytical Model

The main theoretical insight of this paper is that the allocations of talent and technology are intertwined due to a technology-occupation trade-off. This insight – once extended to richer frameworks – may have a variety of additional predictions. In this section, I informally discussed two directions that I explored: introducing self-employment, and dynamics.

We might also be concerned about the robustness of the occupation-technology trade-off to a richer framework in which managers are allowed to choose firm size. I here show that – under stronger assumptions on the skill-technology complementarity – the trade-off is robust.

Self-Employment. I extend the tractable framework of Section 3.2 allowing individuals to choose whether to produce alone or in teams. Details and empirical results are in Appendix B. I model self-employment along the lines of Garicano and Rossi-Hansberg (2004): it is an occupation more skill-intensive than a worker, but less than a manager. As a result, when everyone uses a similar technology, as in countries close to the technology frontier, the self-employed are just as skilled as the average individual. However, selection into self-employment is different in countries far from the technology frontier. I show that the farther a country is from the technology frontier, the more negatively selected the self-employed are. Intuitively, in countries far from the frontier, low-skilled individuals become self-employed because the types that would be their managers in a country close to the frontier choose to be workers themselves to get access to advanced technology. In this sense, the model provides an equilibrium explanation for the prevalence of subsistence self-employment in poor countries, a phenomenon widely documented in the literature (e.g., Schoar (2010), Ardagna and Lusardi (2010) and Banerjee and Duflo (2012)). I also use the same dataset described in Section 4 to document that, consistent with model's predictions, own-account self-employed workers are relatively less skilled in countries far from the frontier.

Growth, Time of Take-off, and Dual Economy Traps. The benchmark model is static. Individuals live one period, and everyone faces identical costs of technology. I next discuss an alternative setting, in which individuals still live one period, but the technology they use survives them and can be used by individuals in next period. In this setting, we need to define an allocation mechanism to assign existing technologies to teams. One possibility is to use a competitive method and let teams bid for existing technologies. This mechanism leads to positive assortative matching – due to skill-technology complementarity – between the technologies and the skills of team's managers. Each team can also decide whether to further improve their assigned technology.

Countries differ with respect to the exogenous time in which they take off and start importing technologies from the frontier. Before that, everyone in the country uses the local technology vintage, which does not improve over time. The frontier technology grows exogenously. Consider, first, a country that takes off early. The frontier technology vintage is still similar to the local vintage. As a result – just as in the static model – there is little technology dispersion, and talent is segmented by occupation. In the periods after the take-off, the better-skilled teams match with the best existing technology, but technology dispersion is contained, and, thus, talent remains segmented by occupation, and the whole economy grows at similar rate. Consider, instead, a country that takes off late. The frontier technology vintage had evolved, while the local one did not. As a result, the country is far from the frontier, and this leads – just as in the static model – to talent segregation by technology: high-skilled individuals pair together and adopt the most advanced technologies. Technology dispersion builds up in the economy, thus favoring further talent segregation. Segregation can be beneficial in the short run, concentrating talent and pushing some parts of the economy to quickly converge to the frontier. At the same time, though, it may lead to dual economy traps since other parts of the economy are crowded out of talent and, thus, have weak incentives to adopt the advanced technology. In this case, reallocating high-skilled workers towards more depressed parts of the economy – towards rural areas, for example – may be beneficial to long run growth.

Firm Size. In the benchmark model, a manager has only one way to leverage his talent: to hire a more-skilled worker. Managers are, in fact, not allowed to hire more than one worker. In Appendix D, I show that this restriction can be relaxed while still preserving the core theoretical insights. I study the marginal values of skills for managers and workers when managers can choose the firm size. Endogenous firm size does play a relevant role, but as long as we assume a stronger degree of complementarity between skills and technology – specifically, log-super-modularity rather than super-modularity, along the lines of results in Grossman et al. (2017) – the technology-occupation tradeoff is still present and so is the role of the cost of technology in shaping the assignment of talent.

7 Conclusion

This paper develops a theoretical framework for studying the link between the allocation of talent into production teams and the cross-sectional distribution of technology. The theory highlights that the main purpose of team production and, thus, the equilibrium assignment of talent depends on the distance of a country from the technology frontier. In countries close to the frontier, all teams use a similar technology, and the main purpose of team production is to gather together high- and low-skilled individuals to allow for task specialization. In contrast, in countries far from the frontier, the main purpose of team production is to match individuals with their appropriate production technology. As a result, high-skilled individuals gather together in teams that adopt the most advanced technology, while the low-skilled are left to match with one another and use backward technology. The theoretical predictions are supported by cross-country micro data on the allocation of talent. Moreover, a quantitative analysis demonstrates that the observed cross-country differences in the allocation of talent can account for a sizable fraction of larger productivity dispersion in poor countries. Thus, the model suggests that we should be cautious about assigning all the observed cross-country differences to larger frictions in developing countries, as is often done in the extensive literature that studies the (mis)allocation of resources in developing countries. Through the lens of the theory here proposed, those exercises are misspecified because they do not take into account the endogenous team formation and technology choice.

More broadly, this paper brings attention to a new feature of economic development that has been previously overlooked: individuals in poor countries form production teams differently than those in rich countries. This paper focuses on the implications of this phenomenon for cross-sectional productivity dispersion. However, the cross-country differences in "who matches with whom" may matter for several other core questions of economic development. For example, the different patterns of matching may influence the formation of social networks, or the degree through which information and knowledge spread out within an economy, thus ultimately affecting the growth rate of a country.[37]

[37] Specifically, a recent literature (e.g., Lucas and Moll (2014), Perla and Tonetti (2014), and Buera and Oberfield (2016)) develops theoretical models in which growth is generated through the diffusion of ideas. If embedded within these models, the different patterns of assignment in developing countries would affect the overall speed of learning and possibly generate a trade-off between the allocation of talent that maximizes output and the one that maximizes technology diffusion. Evidence of slow accumulation of human capital over the life-cycle, which might be driven by slow learning due to lack of interaction of individuals of different skills, is documented in Lagakos et al. (2017).

References

Acemoglu, Daron, "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence," *American Economic Review*, 1999, pp. 1259–1278.

_ , "Localised and Biased Technologies: Atkinson and Stiglitz's New View, Induced Innovations, and Directed Technological Change," *The Economic Journal*, 2015, *125* (583), 443–463.

_ **and Fabrizio Zilibotti**, "Productivity Differences," *Quarterly Journal of Economics*, May 2001, *116* (2), 563–606.

_ **, Philippe Aghion, and Fabrizio Zilibotti**, "Distance to Frontier, Selection, and Economic Growth," *Journal of the European Economic association*, 2006, *4* (1), 37–74.

Adamopoulos, Tasso and Diego Restuccia, "The Size Distribution of Farms and International Productivity Differences," *The American Economic Review*, 2014, *104* (6), 1667–1697.

Ardagna, Silvia and Annamaria Lusardi, "Explaining International Differences in Entrepreneurship: The Role of Individual Characteristics and Regulatory Constraints," in "International Differences in Entrepreneurship," University of Chicago Press, 2010, pp. 17–62.

Asker, John, Allan Collard-Wexler, and Jan De Loecker, "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 2014, *122* (5), 1013–1063.

Atkinson, Anthony B and Joseph E Stiglitz, "A New View of Technological Change," *The Economic Journal*, 1969, pp. 573–578.

Banerjee, Abhijit and Esther Duflo, *Poor economics: A Radical Rethinking of the Way to Fight Global Poverty*, PublicAffairs, 2012.

Basu, Susanto and David N Weil, "Appropriate Technology and Growth," *Quarterly Journal of Economics*, 1998, pp. 1025–1054.

Ben-Porath, Yoram, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 1967, *75*, 352.

Bloom, Nicholas and John Van Reenen, "Why Do Management Practices Differ Across Firms and Countries?," *The Journal of Economic Perspectives*, 2010, pp. 203–224.

Bloom, Nick and John van Reenen, "Measuring and Explaining Management Practices Across Countries and Firms," *Quarterly Journal of Economics forthcoming*, 2007.

Buera, Francisco J and Ezra Oberfield, "The Global Diffusion of Ideas," Technical Report, National Bureau of Economic Research 2016.

_ **, Joseph P Kaboski, and Richard Rogerson**, "Skill Biased Structural Change," Technical Report, National Bureau of Economic Research 2015.

Caselli, Francesco, "Technological Revolutions," *American economic review*, 1999, pp. 78–102.

_ , "Accounting for Cross-Country Income Differences," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 1, Elsevier, 2005, chapter 9, pp. 679–741.

_ **and Wilbur John Coleman**, "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, 2001, *109* (3), 584–616.

Caunedo, Julieta and Elisa Keller, "Capital Obsolescence and Agricultural Productivity," *Cornell University Working Paper*, 2016.

Chernozhukov, Victor and Han Hong, "An MCMC Approach to Classical Estimation," *Journal of Econometrics*, 2003, *115* (2), 293–346.

Chiappori, Pierre-André, Alfred Galichon, and Bernard Salanié, "The Roommate Problem is More Stable than you Think," 2014.

_ , Robert J McCann, and Lars P Nesheim, "Hedonic Price Equilibria, Stable Matching, and Optimal Transport: Equivalence, Topology, and Uniqueness," *Economic Theory*, 2010, *42* (2), 317–354.

Comin, Diego A and Bart Hobijn, "The CHAT Dataset," Technical Report, National Bureau of Economic Research 2009.

_ and Martí Mestieri, "If Technology Has Arrived Everywhere, why Has Income Diverged?," Technical Report, Darmouth College Working Paper 2016.

Eeckhout, Jan and Philipp Kircher, "Assortative Matching with Large Firms: Span of Control over More versus Better Workers," *Universitat Pompeu Fabra Working Paper*, 2012.

Feenstra, Robert C, Robert Inklaar, and Marcel P Timmer, "Penn World Table Version 8.0," *The Next Generation of the Penn World Table, available for download at www. ggdc. net/pwt*, 2013.

Foster, Andrew D and Mark R Rosenzweig, "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution," *The American economic review*, 1996, pp. 931–953.

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

_ and _ , "Organization and Inequality in a Knowledge Economy," *The Quarterly Journal of Economics*, 2006, *121* (4), 1383–1435.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "Human Capital and Regional Development*," *The Quarterly journal of economics*, 2013, *128* (1), 105–164.

Goldin, Claudia and Lawrence F Katz, "The Origins of Technology-Skill Complementarity*," *The Quarterly journal of economics*, 1998, *113* (3), 693–732.

Gollin, Douglas, David Lagakos, and Michael E. Waugh, "The Agricultural Productivity Gap," *Quarterly Journal of Economics*, 2014, *129* (2), 939–993.

Grossman, Gene M, Elhanan Helpman, and Philipp Kircher, "Matching and Sorting in a Global Economy," *Journal of Political Economy*, 2017, *125* (1).

Hanushek, Eric A and Ludger Woessmann, "The Role of Cognitive Skills in Economic Development," *Journal of economic literature*, 2008, pp. 607–668.

Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi, "Growth and Structural Transformation," *Handbook of Economic Growth*, 2014, *2*, 855–941.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

_ **, Erik Hurst, Charles I. Jones, and Peter J. Klenow**, "The Allocation of Talent and U.S. Economic Growth," Stanford Working Paper 2016.

Kremer, Michael, "The O-ring Theory of Economic Development," *The Quarterly Journal of Economics*, 1993, pp. 551–575.

_ **and Eric Maskin**, "Wage Inequality and Segregation by Skill," Technical Report, National Bureau of Economic Research 1996.

Lagakos, David, "Explaining Cross-Country Productivity Differences in Retail Trade," *Journal of Political Economy*, 2016, *124* (2), 579–620.

_ **and Michael E Waugh**, "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, 2013, *103* (2), 948–80.

_ **, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman**, "Life-Cycle Wage Growth Across Countries," *Journal of Political Economy*, 2017, (Forthcoming).

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin, "Matching, Sorting and Wages," *Review of Economic Dynamics*, 2016, *19*, 63–87.

Lucas, Robert E. and Benjamin Moll, "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 2014, *122* (1), 1 – 51.

Lucas, Robert E Jr., "On the Size Distribution of Business Firms," *Bell Journal*, 1978, *9*, 508–523.

McFadden, Daniel, "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica*, 1989, pp. 995–1026.

Mestieri, Martí, "Wealth Distribution and Human Capital: How Do Borrowing Constraints Shape Educational Systems," Technical Report, Northwestern Working Paper 2016.

Minnesota Population Center, *Integrated Public Use Microdata Series, International: Version 6.1 [Machine-readable database]*, University of Minnesota, 2011.

Perla, Jesse and Christopher Tonetti, "Equilibrium Imitation and Growth," *Journal of Political Economy*, 2014, *122* (1), 52 – 76.

Porta, Rafael La and Andrei Shleifer, "Informality and Development," *The Journal of Economic Perspectives*, 2014, pp. 109–126.

Porzio, Tommaso, "Distance to the Technology Frontier and the Allocation of Talent," *Yale University Working Paper*, 2016.

Roys, Nicolas and Ananth Seshadri, "Economic Development and the Organization of Production," Technical Report, University of Wisconsin Madison Working Paper 2014.

Schoar, Antoinette, "The Divide between Subsistence and Transformational Entrepreneurship," in "Innovation Policy and the Economy, Volume 10," University of Chicago Press, 2010, pp. 57–81.

Suri, Tavneet, "Selection and Comparative Advantage in Technology Adoption," *Econometrica*, 2011, pp. 159–209.

Young, Alwyn, "Inequality, the Urban-Rural Gap and Migration*," *The Quarterly Journal of Economics*, 2013, p. qjt025.