THE ALLOCATION OF TALENT: FINANCE VERSUS ENTREPRENEURSHIP

Kirill Shakhnov

European University Institute
Department of Economics

The Allocation of Talent: Finance versus Entrepreneurship

Kirill Shakhnov

The Allocation of Talent: Finance versus Entrepreneurship

Kirill Shakhnov [*]

First version: January 2014
This version: November 2014

Abstract

The rapid growth of the US financial sector has driven policy debate on whether it is socially desirable. I propose a heterogeneous agent model with asymmetric information and matching frictions that produces a tradeoff between finance and entrepreneurship. By becoming bankers, talented individuals efficiently match investors with entrepreneurs, but do not internalize the negative effect on the pool of talented entrepreneurs. Thus, the financial sector is inefficiently large in equilibrium, and this inefficiency increases with wealth inequality. The model explains the simultaneous growth of wealth inequality and finance in the US, and why more unequal countries have larger financial sectors.

Keywords: talent, financial sector, matching, productivity.

JEL Classification: E44, G14, L26, O15.

Introduction

"We are throwing more and more of our resources, including the cream of our youth, into financial activities remote from the production of goods and services, into activities that generate high private rewards disproportionate to their social productivity."

— James Tobin (1984)

The growth of the financial sector is well known and well documented. Figure 1 shows that the share of finance in GDP as well as employment has increased substantially since the Second World War. The figure shows that finance accounts for a higher share of GDP than of employment before the Second World War and after the 1980s (Philippon and Reshef, 2012). More importantly, while the share of finance in employment has stabilized after the 1980s, the share of finance in GDP has continued to rise.
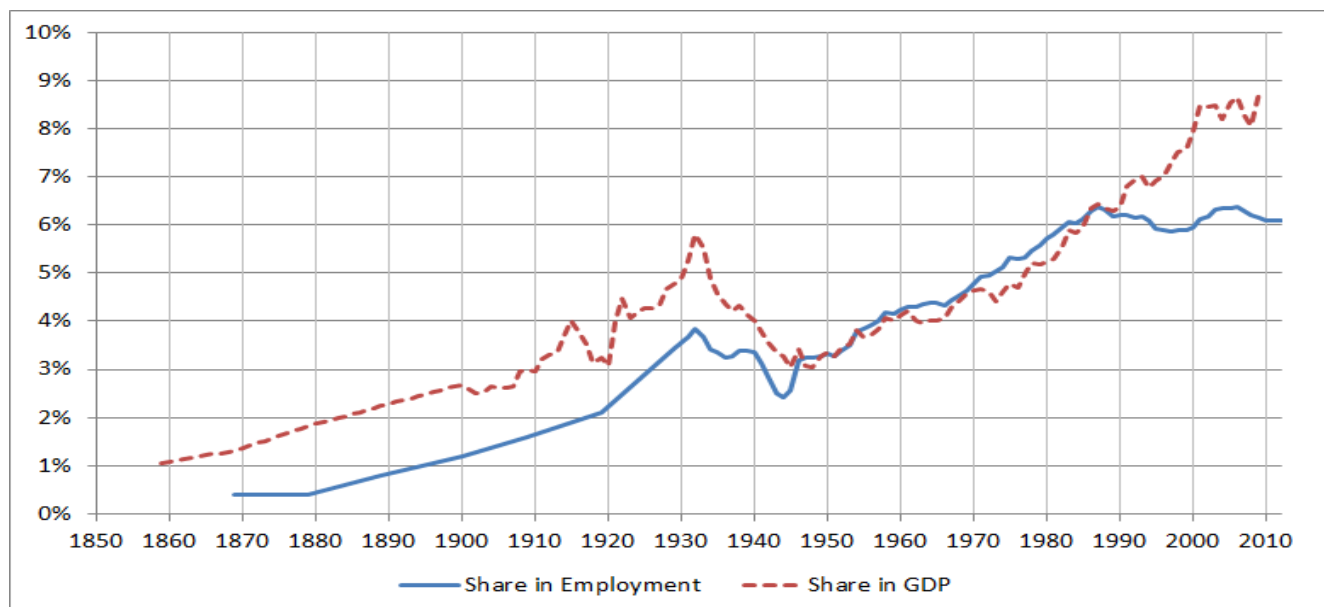

Figure 1: The growth of the financial sector in the US

The substantial expansion of the financial sector has driven a debate on whether this expansion is socially desirable. On the one hand, the former chairman of the Federal Reserve, Alan Greenspan (2002) stated: "[M]any forms and layers of financial intermediation will be required if we are to capture the full benefit of our advances in technology and trade." This idea is related to a vast literature arguing that financial development causes economic growth, because by relaxing financial constraints the financial sector corrects capital misallocation and consequently mitigates productivity losses from financial frictions. (See Schumpeter (1934) for an early contribution and also Merton (1995). Brunnermeier et al. (2012) review the macroeconomic implications of financial frictions, while Levine (2005) provides a survey of an even larger empirical literature.)

On the other hand, critics of the financial sector suggest that it might have negative implications for the allocation of talent. Another former chairman of the Federal Reserve, Paul Volcker (2010)

2

clearly stated the issue: "[I]f the financial sector in the United States is so important that it generates 40% of all the profits in the country...What about the effect of incentives on all our best young talent, particularly of a numerical kind, in the United States?" Furthermore, this concern has been vividly expressed on both sides of the Atlantic, in particular by Lord Turner, the former chairman of the UK's Financial Services Authority, who stated in 2009 that the financial sector had increased "beyond a socially reasonable size." Barack Obama (2012) and James Tobin (1984) tend to agree. This concern has been supported by empirical findings. For example, Berkes et al. (2012) suggest that finance starts having a negative effect on output growth when credit to the private sector reaches 100% of GDP. Other authors, such as Goldsmith (1995) and Lucas (1988), claim that the role of finance has been overstated, and argue that it responds passively to economic growth.

In order to evaluate these claims in a structured way, I build a model in which financial intermediation potentially enhances welfare but draws some talented individuals away from production. The model includes three key elements: (a) heterogeneous agents who differ in terms of capital and talent; (b) an occupational choice between being a banker or an entrepreneur; (c) financial frictions. Heterogeneity and an occupational choice provide a framework to study the allocation of capital (wealth) and talent. Talent is important for both industry and the financial sector: more talent in industry means more output is produced per unit of capital, while more talent in finance means capital is potentially allocated more efficiently. Financial frictions in the form of private information cause the misallocation of capital, because investors cannot distinguish between talented and ordinary entrepreneurs. Since talented bankers can make this distinction, the financial sector can potentially correct this misallocation.

The model generates four important insights about the financial sector. First, it implies that the optimal (constrained efficient) size of the financial sector is larger for countries or periods with higher wealth and talent inequality, because in these cases the potential productivity losses from capital misallocation are particularly severe. The planner faces a tradeoff between the misallocation of capital and the misallocation of talent. Second, the decentralized equilibrium exhibits a misallocation of talent: the financial sector absorbs talent beyond the socially desirable level, because it provides talented agents with an opportunity to extract an excessive informational rent due to the presence of externalities. When agents make their occupational choice between finance and entrepreneurship, they do not internalize the negative externality that they impose on investors: the more bankers there are, the fewer talented entrepreneurs and good investment opportunities there are. Third, even though the equilibrium is generically *inefficient*, efficiency can be restored by taxing the financial sector. Fourth, the model provides a novel explanation for the growth of finance by linking it to an increase in *wealth inequality*. In the dynamic framework, this effect is self-reinforcing: small initial differences in wealth among investors cause substantial income inequality among entrepreneurs, which is translated into greater wealth inequality next period. Wealthy investors are willing to pay a higher premium for financial services that increase the return on their savings, and so the greater is the dispersion of wealth, the higher is the price of financial services. This higher price induces a larger fraction of talented agents to pursue careers in finance. Hence, the growth of finance and the increase in wealth inequality go hand in hand.

Some papers provide indirect empirical evidence on the misallocation of talent. Data from college graduates in the US suggests that the financial sector has become one of the most popular destinations for graduates of elite universities, regardless of their major. For example, Shu (2012), studying the career choices of MIT graduates, concludes that careers in finance attract students with high levels of raw academic talent. She concludes that the overall allocation of talent is inefficient. (See also Goldin and Katz (2008) for Harvard graduates, and Wadhwa et al. (2006) for Engineering Management graduates at Duke University.) In addition, Kneer (2012) finds that US banking deregulation reduces labor productivity disproportionately in industries that are relatively skill-intensive. Finally, MGI (2011) estimates that the United States may face a shortfall of almost two million technical and analytic workers over the next ten years.

This paper is related to a vast literature on misallocation, particularly to papers attributing the misallocation of capital to financial frictions (Buera and Shin, 2013; Midrigan and Xu, 2014). Whereas most papers focus on the impact of frictions on output and the allocation of capital, and abstract away from its impact on the labor market and the allocation of human capital (Jovanovic (2014) is one of the exceptions), this paper argues that financial development has an important impact on the allocation of both capital *and* talent, which cannot be neglected. The issue of allocative efficiency has also been studied theoretically in relation to venture capital. For example, Jovanovic and Szentes (2013) show that the competitive equilibrium is always socially optimal, while in search-matching models such as Michelacci and Suarez (2004), the Hosios condition must hold for the equilibrium to be efficient.

Apart from the current paper, three recent papers have analyzed whether the expansion of the financial sector is efficient. The financial sector is inefficient in all three papers, but the source of the inefficiency is different. Murphy et al. (1991) argue that the flow of talented individuals into law and finance might not be entirely desirable, because even though private returns in these occupations are high, social returns might be higher in other occupations. However, they provide no reason for the disparities between social and private returns. The study of Philippon (2010) is the first that acknowledges the meaningful role of the financial sector, a monitoring device that helps to overcome the opportunistic behavior of entrepreneurs. The allocation is not optimal in his model, because the projects developed by entrepreneurs have higher social benefits than private ones; therefore, they need to be subsidized with respect to workers and bankers. Bolton et al. (2011) focus on financial innovations, in the sense that the financial sector creates a new over-the-counter (OTC) market. Informed dealers in the OTC market extract excessive rents, and consequently the financial sector attracts too many individuals. However, none of these papers seek to explain the growth of the financial sector; none of them consider the financial sector as financial intermediaries connecting investors and entrepreneurs; neither Murphy et al. (1991) nor Philippon (2010) allow for excessive informational rent extraction; and finally, neither Philippon (2010) nor Bolton et al. (2011) have a role for talent in either finance or industry.

Many studies analyze the causes of the expansion of the financial sector. Several explanations have been suggested: the fluctuation of trust in financial intermediaries (Gennaioli et al., 2013); the increasing efficiency of the production sector (Bauer and Mora, 2014); structural change in

finance (Cooley et al., 2013); and asset bubbles (Cahuc and Challe, 2012). None of them connect the expansion of the financial sector and the increase in wealth inequality. The only paper that partially attributes the growth of finance to capital accumulation is Gennaioli et al. (2013). I focus not on aggregate capital accumulation, but rather on increasing wealth inequality. I show that the growth of wealth inequality alone is enough to fully explain the growth of finance. This is in line with Piketty and Zucman (2014)'s argument that the primary reason for increased inequality is the fact that financial services associated with asset management generate superior returns and disproportionately affect the wealthy. According to Greenwood and Scharfstein (2013), much of the growth of the financial sector comes from asset management, which is mostly a service for wealthy individuals.

The calibrated model qualitatively replicates well other features of the US data: the increase in wealth inequality, the productivity slowdown, and the growth of the financial sector as a share of both employment and GDP. The model predicts that the financial sector would continue to grow as a share of GDP, but not of employment. It also provides an additional explanation for the US productivity slowdown. Furthermore, cross-country regressions show that, in line with the predictions of the model, inequalities of wealth and talent are positively associated with the size of the financial sector.

The paper is structured as follows. Section 1 describes the static version of the model and policy results. Section 2 provides the dynamic version of the model and quantitative analysis. Section 3 performs a cross-country analysis to confirm the findings. The last section discusses the paper, concludes, and motivates further research.

1   Static model

There are two opposing views on finance. On the one hand, a large literature on finance and development establishes a positive link between finance and aggregate output. From the theory side, the standard way to think about the issue is that, due to financial frictions, there is misallocation of capital and consequently output losses, which can be severe. The financial sector plays an important role in overcoming or at least mitigating the effect of these frictions. Based on this view, the main policy prescription is to promote the development of the financial sector. On the other hand, the Great Recession has cast doubt on the efficiency of the rapid growth of finance, suggesting that possible rent seeking behavior might be involved. The model presented below features financial frictions that generate capital misallocation. The financial sector can correct this misallocation at the cost of talent misallocation.

I adopt the "classical" view of financial intermediaries as institutions that connect surplus agents (investors) and deficit agents (entrepreneurs). Financial intermediaries are efficient at obtaining information, but they require talent to acquire this information. A talented banker can screen entrepreneurs to discover the best investment opportunities, and sells this information to an investor. The financial sector in the model is clearly a productive sector, because it mitigates informational frictions.

1.1 Environment

The economy consists of two types of agents: investors and entrepreneurs. To produce output, two inputs are required: capital and an idea. Investors have wealth but no investment opportunities of their own, while entrepreneurs have ideas but need external funding.

Agents are heterogeneously endowed with talent and wealth. (Since capital is the only asset in the economy, the terms "wealth" and "capital" are used interchangeably.) Investors can be capital-abundant or capital-scarce, while entrepreneurs can be talented or ordinary. Entrepreneurs can choose whether to remain entrepreneurs or to become bankers instead. In industry, talent translates into capital productivity. The more talented is the entrepreneur, the more output is produced from a unit of capital. In finance, talent affects bankers' ability to distinguish between talented and ordinary entrepreneurs or to sort them, as we shall see below.

I consider a two-sided one-to-one matching market: to produce, one entrepreneur needs to be matched with one investor. The economy is subject to financial frictions: two-sided private information, meaning that the types of entrepreneurs (investors) are not publicly observable. When investors are looking for investment opportunities, they do not know whether an entrepreneur that they meet is talented or ordinary. The same holds for entrepreneurs: entrepreneurs do not know whether an investor they are dealing with is capital-abundant or -scarce. Even though the latter assumption seems questionable at first, in the venture capital industry it is common for entrepreneurs to be imperfectly informed about the total wealth of investors.[1] Two-sided private information guarantees that the outcome is random matching. The literature on assortative matching states that as long as the private information is one-sided, there is a separating equilibrium that supports the same positive assignment as in the full-information equilibrium assignment. In the economy with private information, but without matching, the aggregate outcome is exactly as in random matching, because investors optimally allocate equal shares to every entrepreneur. Matching simply ensures that all funds are not allocated to one entrepreneur. Alternatively, we can simply assume that *without financial intermediation*, the investment technology in the economy is random matching.

All agents are assumed to be risk-neutral and discount the future at a zero rate, so all agents maximize their incomes.

---

[1]In the model, the wealth of investors is invested fully; immediately afterwards, a one-time investment output is produced. In reality, it is more complicated. Even after engaging with a venture capitalist, the entrepreneur faces a substantial degree of uncertainty about the total amount of investment, because of staging. Staging is one of the central incentive mechanisms used in the venture capital industry (Sahlman, 1990). As shown by Bienz and Hirsch (2011), staging is frequently implemented through multiple negotiated financing rounds. Furthermore, the venture capital literature often assumes that *neither* the inputs of the investor nor those of the entrepreneur are contractible. The standard feasible contract in the venture capital literature specifies only a sharing rule and an initial investment, but not the total investment, which, like entrepreneurial inputs, is assumed to be noncontractible.

1.2 Simple model without finance

This subsection presents a simple static general equilibrium model with unobserved heterogeneity. The model without finance and full information is a variant of the standard static model of two-sided matching in which a Becker–Brock type of assignment problem arises (Becker, 1973). I add to this framework two features: two-sided private information and intermediation. Two-sided private information ensures that the assignment should be random—*without intermediation* (the financial sector), there is no mechanism to enforce positive assignment (assortative matching). The full dynamic model presented in the next subsection will incorporate this same static model into a dynamic framework.

In this section, for the sake of simplicity, I consider a *very particular distribution of wealth and talent:* there is a unit mass of agents with talent and no capital, who can be talented z^H or ordinary z^L; there is a unit mass of agents with capital and no talent, who can be capital-abundant k^H or capital-scarce k^L. The share of capital-abundant investors (talented entrepreneurs) is denoted as β^i (β^e). Hence, the mass of agents with capital is equal to the mass of agents with talent. Agents with capital and no talent are potential investors, while agents with no capital and talent can be either entrepreneurs or bankers. Every investor can be matched with at most one entrepreneur. Hence, I consider the simplest case of matching, which is *one-to-one matching.* Furthermore, I assume that all short-sided agents are matched with certainty.[2] The outcome of the match is given by a *strictly supermodular function* $F(z, k)$ depending on both capital and talent. The strict supermodularity in the discrete case is given by:

$$F(z^H, k^H) + F(z^L, k^L) > F(z^H, k^L) + F(z^L, k^H) \tag{1}$$

Condition (1) suggests that positive assortative matching maximizes the sum of match outputs when the entrepreneur's type and the investor's type are complements in the match output function.

Figure 2 shows the outcome of matches in this economy. Since investors and entrepreneurs can be of only two types, we have four possible outcomes (sky blue, yellow, pink and orange). I introduce an additional notation $F_{IJ} = F(z^I, k^J)$, where $I, J = \{H, L\}$, I stands for the entrepreneur's type and J stands for the investor's type. For example, F_{HL} is the outcome of a match between a high-type entrepreneur and a low-type investor; the yellow area is the combination of two colors: green z^H and brown k^L.

1.3 First best vs. random matching

In this section, I define the *first best* as an optimal allocation under the constraint of the matching technology. Since the financial sector mitigates information frictions but does not directly contribute to production, the first best in this economy is the allocation in which nobody

---

[2]One-to-one matching can be viewed as a technological constraint. Many-to-one matching, different specifications of the matching function and a continuum of types over talent and wealth are discussed in section 4 below.
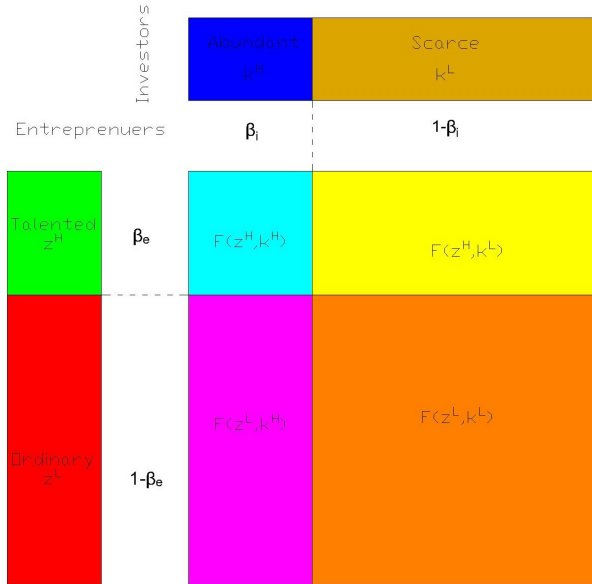
Figure 2: Model without finance

is a banker and all talented agents are matched with investors. Under the supermodularity assumption on the production function (outcome of the match) and observability of types, the most efficient outcome in this economy is positive assortative matching—when high-type entrepreneurs are matched with high-type investors, and low-type entrepreneurs are matched with low-type investors (see the Becker–Brock efficient matching theorem). However, assortative matching cannot be achieved due to two-sided private information about types, so I consider the assortative matching outcome as the first-best allocation. The only possible outcome in the economy with private information and without a financial sector is random matching.

The simple example below shows the disparity between the first best and random matching: the loss of aggregate output due to the misallocation of capital caused by private information in this economy can be severe. I consider the case in which the production function is simply the product of two inputs $F = zk$. I assume that the value of the high type is one with probability one-quarter, while the value of the low type is zero with the complementary probability for both the distribution of talent and the distribution of wealth. Hence, only if two high types are matched is any output (one unit) produced. It happens with probability 1/16 in the case of random matching and with probability 1/4 in the case of assortative matching (the first best). Table 1 summarizes the information described above. As we can see, output is four times lower in the case of random matching compared to the first best due to capital misallocation. This brings us to the first main question of whether the financial sector can mitigate this capital misallocation.

1.4 The role of finance in the model

The financial sector clearly provides many useful functions to the economy, as discussed in section 4. This paper focuses on two services: intermediation and sorting between investors and entrepreneurs. It is important to remember that the most desirable outcome is assortative match-

8

Table 1: The simple example

	value	probability
z^H	1	1/4
z^L	0	3/4
k^H	1	1/4
k^L	0	3/4
Random matching	1/16	
Assortative matching	1/4	

ing. All investment goes to industry. Bankers are good at sorting, but they do not directly produce any output. The quality of sorting depends on talent. Both finance and industry require talent. While talent in industry increases the firm's productivity, talent in finance gives an *advantage in obtaining information* and therefore increases the quality of sorting. By the latter, I mean that the financial sector brings the allocation as close as possible to the allocation under assortative matching. However, the allocation under assortative matching cannot be achieved. As a reminder, the first-best allocation is the allocation under assortative matching; the allocation without financial intermediation is the allocation under random matching. I call the allocation with financial intermediation the allocation under *intermediated matching*. It is important to distinguish between the constrained efficient allocation under intermediated matching, discussed in the next subsection, and the decentralized allocation under intermediated matching, discussed later.

For most of this paper, if not specified otherwise, I consider an extreme case in which only the high-type z^H banker can match a talented entrepreneur and a capital-abundant investor for sure, while the low-type z^L banker can only match randomly. This assumption has two possible interpretations. Under the first interpretation, the quality of sorting depends on the talent of the agent who does the sorting. A banker with ability z can distinguish between ideas with productivity z and $z' < z$. Hence, the planner would only consider allocating talented z^H agents to finance.

Under the second interpretation, there is a cost of screening $\psi(z)$ per project discovered, which depends on talent. If this cost is high enough for the low type while low enough for the high type, $\psi(z^L) \gg \psi(z^H)$, then the planner might find it optimal to allocate to intermediation some of the talented agents, who can provide efficient matches at a small cost, while she would not allocate any of the ordinary agents to intermediation because of their higher matching costs. In other words, the financial sector provides a useful service (sorting) because it has an information advantage, but requires talent to realize this advantage. This accords with Philippon and Reshef's (2012) empirical observation that working in a world of innovative finance requires talent.[3]

I introduce an additional technical assumption: limited capacity. Each banker can only provide

---

[3]In other words, talented bankers provide an investment opportunity with a superior return because of their informational advantage. We can also think of agents as having different search costs in the case of search frictions.

transaction support for one deal at a time. This assumption is to ensure that one banker cannot undo all private information frictions. This assumption is discussed in detail in section 4. To sum
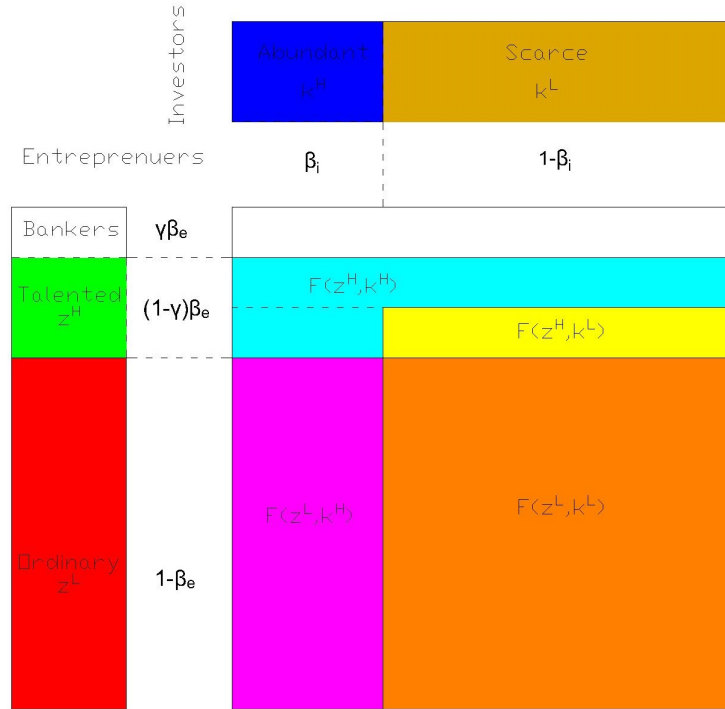
Investors

| Abundant k^H | Scarce k^L |

Entreprenuers β_i $1-\beta_i$

Bankers $\gamma\beta_e$

Talented z^H $(1-\gamma)\beta_e$ $F(z^H,k^H)$ $F(z^H,k^L)$

Ordinary z^L $1-\beta_e$ $F(z^L,k^H)$ $F(z^L,k^L)$

Figure 3: Model with finance

up, the two assumptions imply that if the share γ of talented agents β^e is allocated to the financial sector, they can match at most $\gamma\beta^e$ talented entrepreneurs. To be precise, $\min[\gamma, 1-\gamma]\beta^e$ talented entrepreneurs are matched by bankers with capital-abundant investors and $\max\{1-2\gamma, 0\}\beta^e$ are left for random matching. Figure 3 summarizes the situation stated above and describes the outcome of matches in the case of $\gamma \leq 1/2$. It is very similar to Figure 2, but with the addition of the financial sector. Out of talented agents β^e, the share γ is allocated to the financial sector, while the share $1-\gamma$, together with all ordinary agents $1-\beta^e$, is allocated to industry. We observe losses (the white area) because some investors remain unmatched, and gains (the sky blue area) because the number of efficient matches increases.

1.5 Constrained efficiency

In this subsection, I introduce the notion of constrained efficiency. A social planner faces the same private information constraints as individuals do. To overcome these constraints, the planner can choose to allocate a fraction of agents to the financial sector. Since only talented agents z^H can distinguish between good and bad projects, they are the only agents that need to be considered as possible bankers. For simplicity, I assume that the number of investors is always greater than the number of entrepreneurs.[4] Hence, some investors are matched with nobody. By allocating

---

[4]I prove that this is necessary for the existence of a decentralized equilibrium. See the proof of Proposition 2 in appendix A.2

the fraction γ of talented agents to finance, the planner gains the value of intermediated matches between talented entrepreneurs and capital-abundant investors F_{HH} and incurs two costs: the direct cost is due to the fact that $\gamma\beta^e$ investors become unmatched; the indirect one is that the probability of being randomly matched with talented entrepreneurs drops substantially. Because of risk neutrality, the constrained efficient allocation is one that maximizes aggregate output. The precise expression for aggregate output is given by

$$
\mathbf{Y} = \max_\gamma \left\{ \frac{(\beta^i - \min \gamma, 1 - \gamma\beta^e)}{(1 - \min \gamma, 1 - \gamma\beta^e)} \left[\max\{1 - 2\gamma, 0\}\beta^e F_{HH} + (1 - \beta^e)F_{LH} \right] + \right.
$$
$$
\left. \frac{(1 - \beta^i)}{(1 - \min \gamma, 1 - \gamma\beta^e)} \left[\max\{1 - 2\gamma, 0\}\beta^e F_{HL} + (1 - \beta^e)F_{LL} \right] + \min \gamma, 1 - \gamma\beta^e F_{HH} \right\}. \tag{2}
$$

As soon as γ exceeds $1/2$, all talented entrepreneurs are matched with capital-abundant investors. There is no gain to allocating an additional talented agent to the financial sector. Therefore, the constrained efficient allocation γ^* cannot exceed $1/2$; otherwise we would observe pure losses in the quantity of talented entrepreneurs without any additional gains from matching, which cannot be efficient.

Proposition 1 describes the solution of problem (2):

Proposition 1. *The constrained efficient allocation γ^* is always the corner solution of problem (2), i.e. γ^* can be either 0 or $1/2$.*

Proof. See Appendix A.1.

I calculate $\Delta\mathbf{Y}$, the difference between the values of the planner's objective (2) with $\gamma = 1/2$ and $\gamma = 0$. This difference is given by

$$
\Delta\mathbf{Y} = (1/2 - \beta^i)\beta^e(F_{HH} - F_{HL}) - \frac{\beta^e}{2}F_{HL} - \frac{(1 - \beta^i)(1 - \beta^e)\beta^e}{2 - \beta^e}(F_{LH} - F_{LL}). \tag{3}
$$

After analyzing expression (3) above, we can conclude that if $\beta^i \geq 1/2$, $\gamma^* = 0$ is the only possible solution of the planner's problem. For $\gamma^* = 1/2$ to be the solution, two conditions must be satisfied: $\beta^i < 1/2$, and F_{HH} needs to be high enough. In other words, it is efficient to have a financial sector if two requirements are met: the probability of a random match between a talented entrepreneur and a capital-abundant investor is relatively low, but the value of this match is relatively high. I provide two potential interpretations of this result. On the one hand, one might think that the level of development affects the optimal size of the financial sector. In a developing country with weak institutions, it is difficult for an investor to meet the "right" entrepreneur. Hence, it is essential for such countries to develop their financial sectors to mitigate the effect of underdeveloped institutions. The conclusion might be that the more developed a country is, the less likely it is to benefit from the financial sector. This conclusion seems at best to be counter-factual. However, Mayer-Haug et al. (2013) observe that entrepreneurial talent is more relevant in developing economies. Furthermore, empirical evidence suggests that the misallocation of capital is a particularly acute problem in developing countries. On the other hand, having a financial sector is efficient for countries with higher degrees of wealth or talent inequality. The more unequal a country is, the higher are the benefits from the presence of the financial sector.

I provide empirical support for the latter interpretation in section 3. (See also Restuccia and Rogerson (2008) for the argument that resource misallocation shows up as low TFP, and Hsieh and Klenow (2009) for empirical evidence on misallocation in China and India.)

If we go back to the simple example in Table 1 and calculate aggregate output in the constrained efficient case, we obtain $1/2\beta^e F_{HH} = 1/8$, which is twice as large as in the case of random matching (the economy without finance), but still two times lower than in the first best. In the case of the simple example, we can say that the financial sector undoes half of the financial friction.

1.6 Decentralized equilibrium

In this subsection, I study the decentralized equilibrium (DE) and compare it to the constrained efficient allocation to answer the question of whether the financial sector attracts the right amount of talent. The main difference between the DE and the constrained efficient allocation is the fact that the occupational choice of agents depends on the *private* returns in the two sectors, as opposed to *social* returns in the planner's case. The planner chooses how many talented agents to allocate to finance and at the same time how much consumption they should get. Given the information structure, it is a complicated task for the market to solve, because the number of talented agents in finance affects the way the surplus is shared between three parties: an investor, an entrepreneur and a banker. On the one hand, surplus is created by agents in industry (entrepreneurs). On the other hand, private information frictions create an information rent that can be captured by agents in the financial sector (bankers). In addition, due to matching it is important to understand how the outcome of the project is split between the investor and the entrepreneur. The most natural way to do this is *Nash bargaining*, where the bargaining power of the entrepreneur $\delta \in [0, 1]$ is exogenously given, and the bargaining power of the investor is the complement $1 - \delta$.

The rest remains as outlined in subsection 1.1. The banker with talent z^H can distinguish between entrepreneurs with productivity z^H and $z^L < z^H$. She can sell this information to an investor for price p_i and an entrepreneur for price p_e. Each talented banker can discover at most one talented entrepreneur z^H and consequently makes at most one match between a capital-abundant investor and a talented entrepreneur. If an investor (entrepreneur) pays p_i (p_e), she knows that she will be matched with a high-type counterpart with certainty; otherwise, she can always choose to be randomly matched for free. I assume that if there are more bankers than talented entrepreneurs, $\gamma > 1/2$, some of the bankers discover nothing and therefore receive zero income. In this case, bankers bear all the risk and need to be compensated for this. Equilibrium prices are set competitively.

Returning to the Nash bargaining problem, to solve the problem, I need to define the bargaining power, the surplus of the match and the outside options of the two counterparties. The outside option to intermediated matching is random matching. Hence, the problem must be solved *backwards*. First, I provide the solution for random matching with a given size of the financial sector γ. Then, I use the solution for random matching as the outside options for the intermediated matching problem.

To solve a Nash bargaining problem following Nash (1950, 1953), I need to define the set of feasible utility payoffs from an agreement U and the utility payoffs to the players from a disagreement D. Since preferences are linear, the sets U and D are given by

$$U = \left\{ (x^e, x^i) | x^e + x^i = F(z, k), x^j \geq 0 \right\}, \tag{4}$$

$$D = \left\{ (d^e, d^i) | \right\}, \tag{5}$$

$$x^e = \arg\max \left[(x - d^e)^\delta (F(z, k) - x - d^i)^{1-\delta} \right]. \tag{6}$$

The solutions are:

$$x^e = \delta \left(F(z, k) - d^i \right) + (1 - \delta) d^e, \tag{7}$$

$$x^i = (1 - \delta) \left(F(z, k) - d^e \right) + \delta d^i. \tag{8}$$

As every banker can discover at most one good project, the total number of discovered good projects that are different from each other is $\min \gamma, 1 - \gamma$. It is worth mentioning that, contrary to the planner's solution to problem (2), $\gamma^* \leq 1/2$, the market outcome can be any number in the interval $[0, 1]$.

Assume that investors have no access to a storage technology, while entrepreneurs have no opportunity for outside borrowing. Thus, the outside options for a random match—the set D in (5)—are $(0, 0)$. The solution of the Nash bargaining problem gives the value of random matching for a capital-abundant investor. Note that not all investors are matched. The value of random matching is equal to the probability of matching with somebody Pr^m multiplied by the sum of products of the probability of being matched with a talented (ordinary) entrepreneur Pr^H (Pr^L) and the value of the match for a capital-abundant investor $(1 - \delta) F(z^I, k^H)$. It turns out that:

$$Pr^m = \frac{1 - \gamma\beta^e - \min \gamma, 1 - \gamma\beta^e}{1 - \min \gamma, 1 - \gamma\beta^e},$$

$$Pr^H = \frac{(1 - \gamma - \min \gamma, 1 - \gamma)\beta^e}{1 - \gamma\beta^e - \min \gamma, 1 - \gamma\beta^e},$$

$$Pr^L = \frac{1 - \beta^e}{1 - \gamma\beta^e - \min \gamma, 1 - \gamma\beta^e}.$$

Hence the outside option for intermediated matching is

$$d^i = \frac{1 - \delta}{1 - \min \gamma, 1 - \gamma\beta^e} \left[(1 - \gamma - \min \gamma, 1 - \gamma)\beta^e F_{HH} + (1 - \beta^e) F_{LH} \right] + \frac{\gamma\beta^e}{1 - \gamma\beta^e} 0. \tag{9}$$

Equation (9) defines the value of random matching for a capital-abundant investor, which is the outside option of a capital-abundant investor when negotiating a deal with a talented entrepreneur after intermediated matching. It is important to note that an increase in the size of the financial sector γ worsens the outside option of the capital-abundant investor, because it affects the relative proportions of agents. I return to this point later on.

Similar to (9), the value of random matching for a talented entrepreneur, which is the outside option for bargaining in the case of intermediated matching, is

$$d^e = \frac{\delta}{1 - \min \gamma, 1 - \gamma\beta^e} \left[(\beta^i - \min \gamma, 1 - \gamma\beta^e) F_{HH} + (1 - \beta^i) F_{HL} \right]. \tag{10}$$

Applying once again the solution of Nash bargaining (7) to the intermediated matching case, I obtain the restriction on the prices that can be extracted from investors (11) and entrepreneurs (12):

$$p_i \leq (1 - \delta)(F_{HH} - d^i - d^e), \tag{11}$$

$$p_e \leq \delta(F_{HH} - d^i - d^e). \tag{12}$$

Conditions (11) and (12) are the participation constraints of a capital-abundant investor and a talented entrepreneur. They state that both an investor and an entrepreneur being matched by a banker cannot be worse off in comparison to the random matching scenario. However, these inequalities are not necessarily binding. It depends on which agents are on the *short side* of the market. In addition, the prices obviously should be non-negative.

To complete the description of equilibria, I need an additional condition (13). For the solution to be interior, $\gamma \in (0, 1)$, the talented agent ($z^H > 0$) should be indifferent between being an entrepreneur or a banker. The income of a talented banker is the probability of finding a talented entrepreneur multiplied by the sum of the two prices that are charged to the investor and the entrepreneur. As long as there are more talented entrepreneurs in the market than bankers, the probability of finding a talented entrepreneur is equal to one. The income of a talented entrepreneur is the share of the surplus received from the match with a capital-abundant investor. The indifference condition is therefore

$$\frac{\min \gamma, 1 - \gamma}{\gamma}(p_i + p_e) = \delta\left(F_{HH} - d^i\right) + (1 - \delta)d^e. \tag{13}$$

Three conditions characterize all decentralized equilibria: the occupational choice condition (13) and two participation constraints in financial services, one for capital-abundant investors (11) and one for talented entrepreneurs (12). For the sake of space, I restrict my attention to the case in which the exogenous parameters are such that the constrained efficient size of the financial sector is *strictly positive* ($\gamma^* = 1/2$). I take the view that the financial sector is essential for the economy. Furthermore, this is the interesting case in which to study policy, because for regions of the parameter space in which the financial sector plays no useful role, policy analysis is trivial. Proposition 2 characterizes the decentralized equilibrium in the $\gamma^* = 1/2$ case in terms of efficiency. A detailed analysis of all possible cases can be found in appendix A.2

Proposition 2. *If it is socially efficient to have a financial sector ($\gamma^* = 1/2$) and a decentralized equilibrium exists,*

  i. *It is unique, $\hat{\gamma}$;*

 ii. *This equilibrium is generically inefficient, $\hat{\gamma} \geq \gamma^*$; and*

iii. *There exists a restriction on the set of exogenous parameters that restores the constrained efficient allocation.*

  Proof. See Appendix A.2.

This restriction can be expressed as $\hat{\delta} = f(\underset{+}{\beta^e}, \underset{+}{z^H/z^L}, \underset{-}{\beta^i}, \underset{+}{k^H/k^L})$. The signs beneath the expression stand for the sign of the derivative of $\hat{\delta}$ with respect to the corresponding variable.

Part (iii) of Proposition 2 might look similar to the Hosios condition in the sense that the condition ensures the externalities cancel out (Hosios, 1990). In the original case of Hosios, efficiency is achieved when the surplus share (bargaining power) between workers and a firm is equal to the matching share (the elasticity of the matching function). In a frictionless environment, there is a particular mechanism, directed search, that restores efficiency. However, in a frictional environment with heterogeneous agents even directed search might not be sufficient. The latter is the case of my model.

The result stated in Proposition 2 has a very intuitive explanation. When talented agents make their occupational choice between finance and entrepreneurship, they do not internalize the externalities that they impose on investors. The more talented agents become bankers, the smaller is the pool of good projects.

The opposite case, in which the set of parameters is such that the constrained efficient size of the financial sector is *zero* ($\gamma^* = 0$) is discussed in appendix A.2. The model of Murphy et al. (1991) can be viewed as a special case of my model under parameter restrictions such that $\gamma^* = 0$.

Proposition 2 states that the decentralized equilibrium is generically inefficient. To put it differently, for a given set of parameters, the solution of the decentralized equilibrium is highly unlikely to be efficient. The question is whether it is possible to restore efficiency. The answer is yes. As discussed, there is a restriction on parameters that restores efficiency. If there is a policy instrument that directly affects one of the exogenous parameters, it is easy to ensure efficiency in the model. For example, if the planner could set the bargaining power of entrepreneurs to the particular value $\hat{\delta}$, it would make the decentralized equilibrium efficient. However, it is not very intuitive to think that such policies exist.

1.7 Taxation

The more interesting question is whether it is possible to restore efficiency using only one tax instrument. Fixing the set of parameters to values such that the decentralized equilibrium exists and is inefficient, I take the tax on the financial sector to be the available tax instrument.

The issue in this economy is that the return to finance is too high in comparison with entrepreneurship. Hence an efficient policy should decrease the return to finance and/or increase the return to entrepreneurship. The former can be done through taxation of the financial sector. The latter can be done through subsidizing entrepreneurship. Taxation of the financial sector has been a hot topic since the Great Recession, especially in the European Union.[5] Subsidies for entrepreneurship are quite common: governments and donors spend billions of dollars subsidizing entrepreneurship training programs around the world (see, for example, Santarelli et al. (2006)).

---

[5]See the discussion of taxation proposals at the European Commission web page: `http://ec.europa.eu/taxation_customs/taxation/other_taxes/financial_sector/index_en.htm`.

I show how a tax τ on bankers' incomes can work. The revenue from this tax is distributed by lump-sum transfers T to balance the government's budget. The last equation of system (14) represents the government's budget constraint. The system below characterizes the equilibrium with taxation:

$$x^e = (1 - \delta)(F_{HH} - d^i(\gamma) - d^e(\gamma)) + T,$$
$$c = (1 - \delta)(F_{HH} - d^i(\gamma) - d^e(\gamma)) - 2(1 - \delta)T - \tau,$$
$$x^e = \frac{1-y}{y}c,$$
$$T = \gamma\beta^e\tau. \tag{14}$$

Given the constrained efficient level $\gamma^* = 1/2$, I impose that $\gamma = \gamma^*$ and calculate the corresponding tax rate. The solution of the system can be represented graphically. Figure 4 plots the tax on banking income in percent as a function of the distortion (inefficiency) $\hat{\gamma} - \gamma^*$. The optimal tax is zero when there is no distortion, and increases with the size of the distortion as expected. The closed-form solution of the system defining the tax on banking income as a function of all exogenous parameters is:

$$\tau = \frac{2\delta(1-\delta)\beta^e F_{HH}}{(2 - \beta^e)} \left[\frac{2(1 - \beta^i)}{\beta^e} \frac{F_{HH} - F_{HL}}{F_{HH}} + \frac{1 - \beta^e}{\beta^e}(1 - 2\delta\frac{F_{HH} - F_{LH}}{F_{HH}} - 1 - \frac{2\delta - \beta^e}{2\delta\beta^e(1 - \delta)} \right].$$
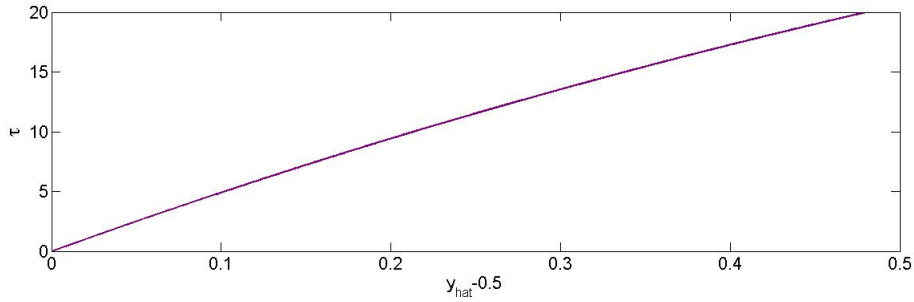

Figure 4: Tax on financial income vs. inefficiency

1.8 Comparative statics

Returning to the solution of the decentralized equilibrium, I analyze the comparative statics of the outcome of the model as exogenous parameters change. The decentralized equilibrium is a function of all exogenous parameters: $\hat{\gamma} = f(\delta, \beta^e, z^H/z^L, \beta^i, k^H/k^L)$. For example, Figure 5 presents the solution $\hat{\gamma}$ as a function of the bargaining power δ. As we can see, the decentralized equilibrium exists only for $\delta \in [0, \hat{\delta}]$; there is no solution for $\delta > \hat{\delta}$. The decentralized equilibrium coincides with the constrained efficient outcome only for one particular value of the bargaining power $\hat{\delta}$.

Figure 6 presents the solution of the decentralized equilibrium as a function of wealth k^H/k^L and talent z^H/z^L dispersion. As we can see, wealth dispersion has a stronger impact on the size of the financial sector. More importantly, the static model predicts that an increase in wealth inequality will be associated with the growth of finance. When the rich get richer, they demand more finance. This is in line with empirical evidence. However, the wealth distribution has
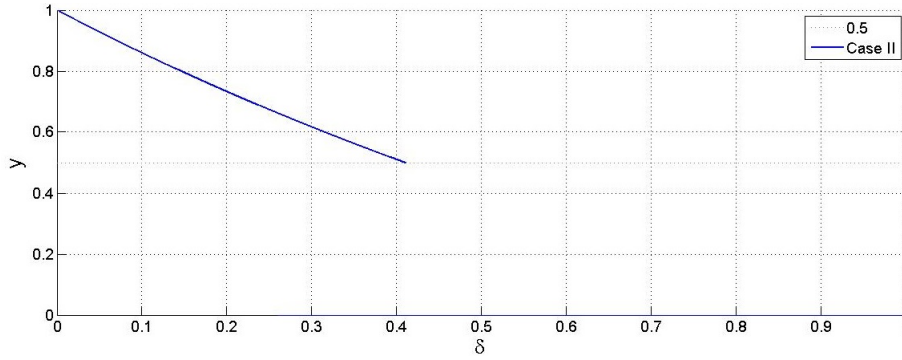
Figure 5: Fraction of bankers vs. bargaining power of entrepreneur
(efficient fraction is 1/2)

been considered completely exogenous up until now. The next section endogenizes the wealth distribution by introducing dynamics into the model.
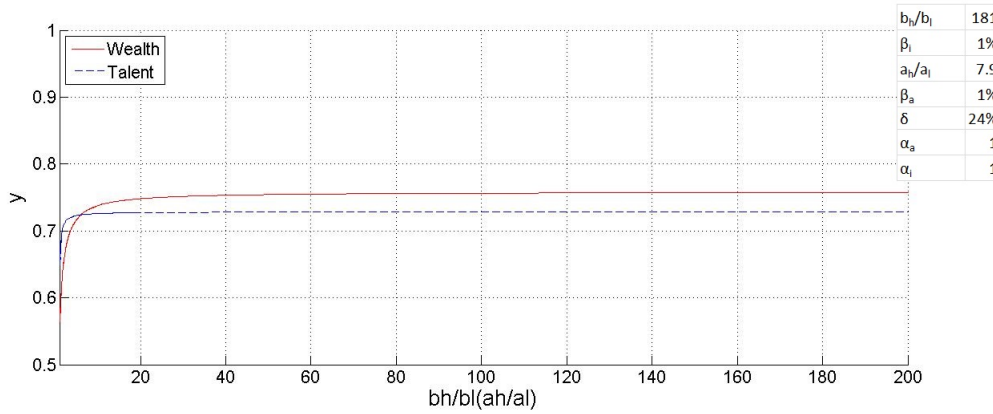

Figure 6: Fraction of bankers vs. dispersion of wealth (talent)
(efficient fraction is 1/2)

2 Dynamic model and quantitative analysis

2.1 Dynamic model

As we saw above, the joint distribution of wealth and talent is an important determinant of the size of the financial sector and the degree of inefficiency. While the distribution of talent is often considered exogenous, it is difficult to think about the wealth distribution as a fully exogenous object. In this subsection, I allow for endogenous wealth accumulation. Endogenous growth of wealth inequality leads to expansion of the financial sector. The rich get richer because they can afford to pay high fees for financial services, which yield a higher return on their savings. The higher are fees, the more talented agents work in finance. Consequently, the growth of finance and the increase in wealth inequality go hand in hand.

17

To introduce simple dynamics, I consider an infinite overlapping generations (OLG) model. The OLG structure seems to be natural for two reasons. First, I study relatively long-term growth of wealth and the size of finance (both have grown for at least the last six decades). Second, the generation structure is well suited to the problem, because agents undergo an interesting life cycle with low-wealth young age and higher-wealth old age. The young make an occupational choice and work in one of the two sectors. The middle-aged invest the wealth they have accumulated while young. The old consume the results of this investment.

I adopt the most basic OLG model. Every individual maximizes lifetime consumption and lives for three periods: youth, middle age and old age. Individuals are born in time t, work at time t, receive their income at $t+1$ and consume at $t+2$. Individuals pass through three stages over the life cycle: working, investment and consumption. The young are endowed with talent z and no wealth. The young make an occupational choice either to be an entrepreneur or a banker. The middle-aged are investors because they have wealth k, which they accumulated while young. The middle-aged have a choice of either being matched randomly or paying to a banker the price p_i in exchange for being matched with certainty with a talented entrepreneur. The middle-aged have no talent, because it fully depreciated within one period. The old consume the result of their investments.[6]

The rest remains as before. Individuals, who are born every period, can be of two types: talented or ordinary. Individuals are assumed to be risk-neutral and not to discount the future. The production function $F(z, k)$ is strictly supermodular and depends on both capital and talent. Financial frictions are two-sided private information and one-to-one matching. The high-type z^h banker can provide intermediated matching, while the low-type z^L banker can provide only random matching.

To keep two types of wealth, I consider a stand-in household that abstracts from the distinction between expected and realized income. Following Lucas and Rapping (1969) and more recent examples (Rogerson, 2008; Gertler and Kiyotaki, 2010), the stand-in household assumption has been a popular tool in macroeconomics to keep models tractable. I introduce the stand-in household in the following way. First, there is income sharing in finance. The realized income that every banker receives is the same as her expected income. Hence, all young talented agents z^H receive the same income, and become capital-abundant investors when they are old. Second, there is pooling of investment funds to ensure that the realized income that an ordinary entrepreneur receives is the same as her expected income. Hence, all ordinary entrepreneurs receive the same amount of capital.[7] These assumptions change nothing from the point of view of expected incomes, but

---

[6]Alternatively, due to risk-neutrality, individuals find it optimal to save their income fully and consume only in the last period. The age-related decline of cognitive abilities is a well-established fact in psychology. There is no consensus regarding the magnitude of the effect or the exact mechanism. The wealth–age profile is also well documented. Wealth grows rapidly over the life cycle and reaches its peak during one's 60s (the end of working age) and flattens or slightly declines afterwards.

[7]We can think of this as an insurance scheme within the financial sector. If agents are slightly risk averse, $u_{t+1}^o = \left(c_{t+1}^o\right)^{1-\epsilon}$, where $\epsilon \approx 0$, all bankers are willing to engage in income sharing, and all investors are willing to engage in fund pooling.

keep the model tractable. If I dropped any of these assumptions, the number of types would grow exponentially with a constant doubling every period.

The simple model produces life-cycle behavior consistent with the data: agents with a given talent level undergo a relatively realistic life cycle with low-income working youth, high-income investment middle age, and retirement with high consumption and zero income. Individuals typically start to accumulate assets for their retirement during middle age, around the age of 40 (Gourinchas and Parker, 2002). Wealth grows rapidly over the life cycle, reaches its peak at the age of 60 and flattens out afterwards. Total individual consumption, including housing and non-housing consumption, mimics individuals' wealth (Yang, 2009).

I keep the distribution of talent constant over time, and assume an initial distribution of wealth parametrized by the share of capital-abundant investors β_0^i and their wealth k_0^H, and the wealth of capital-scarce investors k_0^L. To use the solution of the static model from the previous subsection, I need to define the evolution of the wealth distribution. The system of equations below defines the evolution of the wealth distribution in the model:

$$\beta_t^i = \beta^e, \tag{15}$$

$$k_{t+1}^H = x_t^e = \delta \left(F \left(z^H, k_t^H \right) - d_t^i \right) + (1-\delta)d_t^e, \tag{16}$$

$$k_t^M = \frac{k_t^H(\beta_t^i - \beta^e(1-\hat{\gamma}_t)) + k_t^L * (1-\beta_t^i)}{1 - \beta^e(1-\hat{\gamma}_t)}, \tag{17}$$

$$k_{t+1}^L = \delta F \left(z^L, k_t^M \right). \tag{18}$$

Due to profit sharing, all talented agents receive the same income and become investors next period. Hence the share of capital-abundant investors every period, with the exception of the first one, is equal to $\beta_{t+1}^i = \beta^e$, expression (15). The next-period wealth of capital-abundant investors k_{t+1}^H is defined by expression (16) using the expressions for outside options in the case of intermediated matching (10) and (9). Finally, I define the next-period wealth of capital-scarce investors k_{t+1}^L, expression (18). Due to fund pooling, every entrepreneur who is not matched receives the same amount of funds k_t^M, expression (17), and consequently the same income which becomes her next-period wealth.

The next subsection brings the model to the US data in an attempt to replicate the dynamics of wealth and the financial sector.

2.2 The US experience

This theoretical model has been designed to explain how the role and size of the financial sector is determined and whether this size is efficient. Even though the model is simplistic, the calibrated version of it performs surprisingly well. The goal of the dynamic model is to explain the interrelationship between the growth of the financial sector in terms of employment and the growth of wealth. Therefore, I choose them as data moments to be matched.

In the first calibration exercise, I seek to explain the behavior of the whole financial sector. Then, I recalibrate the model to explain the behavior of one subindustry of the financial sector—

private equity finance. Even though the model can be applied to the financial sector as a whole, private equity finance is an industry for which the assumptions of the model are particularly valid. The first eight parameters described in Table 2 are used to match as closely as possible the share of

Table 2: Parameter values

Parameter		Value
Distribution of talent		
Talented	z^H	6.5
Ordinary	z^L	3.8
Share of talented	β^e	6.7%
Initial distribution of wealth		
Capital-abundant	k_0^H	100
Capital-scarce	k_0^L	95
Share of capital-abundant	β_0^i	3%
Other parameters		
Elasticity of talent	α_z	1
Elasticity of capital	α_k	1.095
Bargaining power of entrepreneur	δ	21%

employment in finance and the ratio of top 5% wealth to median wealth over time in the US. The economy starts initially with an almost egalitarian distribution of wealth ($k^H = 100$ vs. $k^L = 95$); otherwise the share of employment in finance immediately jumps to the steady-state level and the wealth disparity explodes. The distribution of talent remains the same every period: talented agents are assumed to be 1.7 ($z^H/z^L = 6.5/3.8 = 1.7$) times more talented than ordinary ones. Following Romer (1986), the production function exhibits non-decreasing return to scale with respect to capital ($\alpha_k = 1.095$, $\alpha_z = 1$)—it is very similar to the AK production function. While the increasing return on capital generates the growth of aggregate capital, the talent differentials ensure the rise of wealth dispersion. Choosing realistic values for the eight parameters, I recall the definition of $\hat{\delta}$, the maximum entrepreneurial bargaining power consistent with the existence of an equilibrium $\hat{\gamma}$, as a function of other parameters from Proposition 2. (See appendix A.2 for more detail.) The calculated value is $\hat{\delta}_0 = 36.6\%$, and it is growing with wealth dispersion, while the data suggests that the level of entrepreneurial bargaining power is rather small. (Kaplan and Stromberg (2003) report that the average founders' share equals 21.3% of a portfolio company's equity value.) Hence, I set the level of bargaining power to be 21% and keep it constant over time. Since $\delta < \hat{\delta}_t$, according to Proposition 2, the solution of the decentralized equilibrium exists and is inefficient ($\hat{\gamma}_t > \gamma^*$). The inefficiency is growing over time because of increasing wealth dispersion. Table 2 summarizes all parameter values.

Figure 7 shows the comparison between the data and the outcome of the model. On the left-hand side, we can see the share of employment in finance over time. On the right-hand side is the

ratio of top 5% wealth to median wealth over time. As we can see, while the share of employment in finance was growing until the 1980s and then stabilized a little above 5%, the ratio of top 5% wealth to median wealth has been increasing over the whole sample apart from a small drop during the Great Recession. The drop reflects the sharp decrease of asset prices: stocks, housing etc. This business cycle type of consideration is outside the scope of this paper.

The length of the period is a decade. A decade is arguably the shortest scale to study long-term events, such as the accumulation of wealth and structural changes in employment, and possibly the most appropriate one. First, ten years is a standard investment horizon for long-term investors, who are the subjects of this study. Second, a decade is a long enough period to abstract away from business cycle considerations, which are outside the scope of this paper. Furthermore, most of the data is available only for the last 60 years.

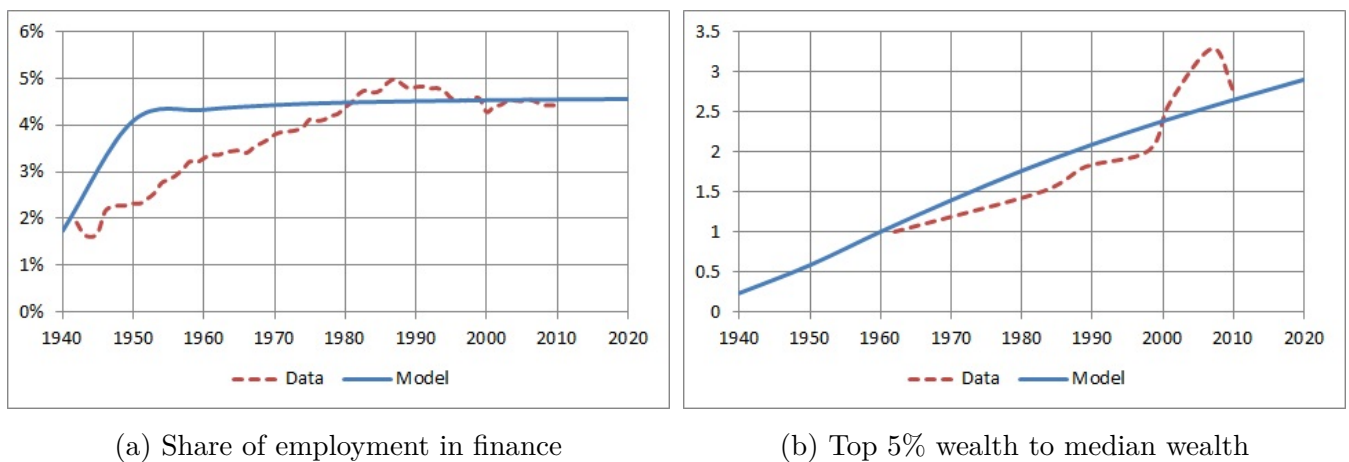(a) Share of employment in finance   (b) Top 5% wealth to median wealth

Figure 7: US data moments

To test the external validity of the model, let us consider some other interesting data trends. First, the share of finance in GDP has steadily increased over the whole period despite the stabilization of the employment share. Figure 8a shows that this is in line with my model. Second, Figure 8b represents the deviation ten-year moving average growth rate of US productivity from the long-term average growth rate for the whole period of observation. It is calculated based on labor productivity per hour worked in 1990 US dollars (converted at Geary–Khamis PPPs). The yearly average for the period after the Second World War is 1.9%. Figure 8b shows that US productivity growth has slowed significantly since 1973, with a minor resurgence of productivity in the 1990s due to the IT revolution. Several explanations of the slowdown have been suggested, but none has been found to be fully satisfactory. My model links the slowdown to a misallocation of talent. As we can see from the data and the model, the biggest increase of the employment share in finance happened in the 1980s. We observe the largest drop in productivity in the US in exactly the same period. This argument is in line with Nordhaus (1982), who argues that a depletion of investment opportunities due to a lack of inventions caused the slowdown. The patent data provides indirect support for this statement. Patent applications and the quality of patents declined significantly in the 1980s despite a constantly increasing number of scientists and researchers. We observe a substantial decline in the share of the US in world patents not only

due to the rise of China, India and Korea, but also in comparison with Scandinavian countries (OECD, 2011). The Scandinavian countries have not experienced any substantial growth of the financial sector. To sum up, I argue that the US productivity slowdown was partly caused by the decreasing number of talented individuals in industry to produce innovations.

(a) Share of finance in GDP   (b) The productivity slowdown   (c) Average return on wealth
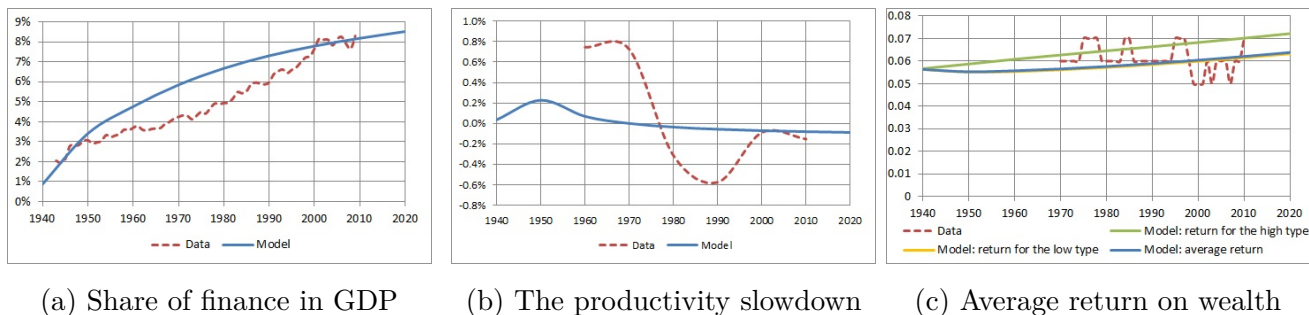
Figure 8: Additional facts about the US

Third, Figure 8c shows the average return on private wealth in the US for the period 1970–2010. The return on wealth (capital) has oscillated around a central value of 5–7% a year. Piketty and Zucman (2014) first noticed that the rate of return on capital is greater than the rate of economic growth over the long term; the result is a growing concentration of wealth. The substantial dispersion in wealth accumulation has been observed despite the minor difference in returns on private wealth between capital-abundant and capital-scarce investors. Given that I do not try to match the average return, it is surprising that the chosen 10-year horizon generates a similar return on capital as in the data.

3 Cross-country data

In this section I provide cross-county evidence to answer the question of how the distribution of wealth and talent affects the size of the financial sector. As predicted by my model, the evidence clearly shows a positive relationship between the size of the financial sector and inequalities of wealth and talent. Even though my model predicts a causal link from the joint distribution of wealth and talent to the equilibrium size of the financial sector, in this section I intend to make no causal statement, only to document correlations.

To test the relationship, I need to have a compatible cross-country measure of moments of talent and wealth distributions, and the size of the financial sector. Unfortunately, data availability limits my choice. For the talent distribution, I employ the scores in the PISA test. The PISA test aims to evaluate education systems worldwide every three years by assessing 15-year-olds' competencies in key subjects: reading, mathematics and science. To date, over 70 countries have participated in PISA. It is a widely used measure for cross-country comparisons of students' performance. Furthermore, it is highly unlikely that this measure of talent suffers from reverse causality. There is no reason why the size of finance today might affect the performance of high-school students today. I used the mean and variance of scores in the PISA test for the years 2003, 2006 and 2009.

I choose the mean and variance of 2009 science scores in the PISA test as a proxy for the moments of talent distributions, because it includes the greatest number of countries. Moreover, this choice hardly affects the results, because PISA scores are highly correlated over time and disciplines: the correlation coefficients exceed 0.97.

To the best of my knowledge, there is no cross-country data on wealth inequality. Therefore, I have to use the income distribution as a proxy for the wealth distribution. Income inequality is a fairly standard proxy for wealth inequality, but possibly underestimates wealth inequality. Income and wealth are not particularly well correlated either at the individual level for a given point (Rodriguez et al. (2002) estimate the correlation between wealth and labor income to be 0.27) or across countries (Fredriksen, 2012). However, if we measure the correlation over time between top income and wealth shares for a particular country, for example the US, we observe that the shares are highly correlated. The more concentrated are the shares, the higher are the correlations between them. In addition, the income shares are more volatile and tend to lead the wealth shares. We can see from Figure 9 that the dynamics of wealth shares closely track the dynamics of income shares for the US.
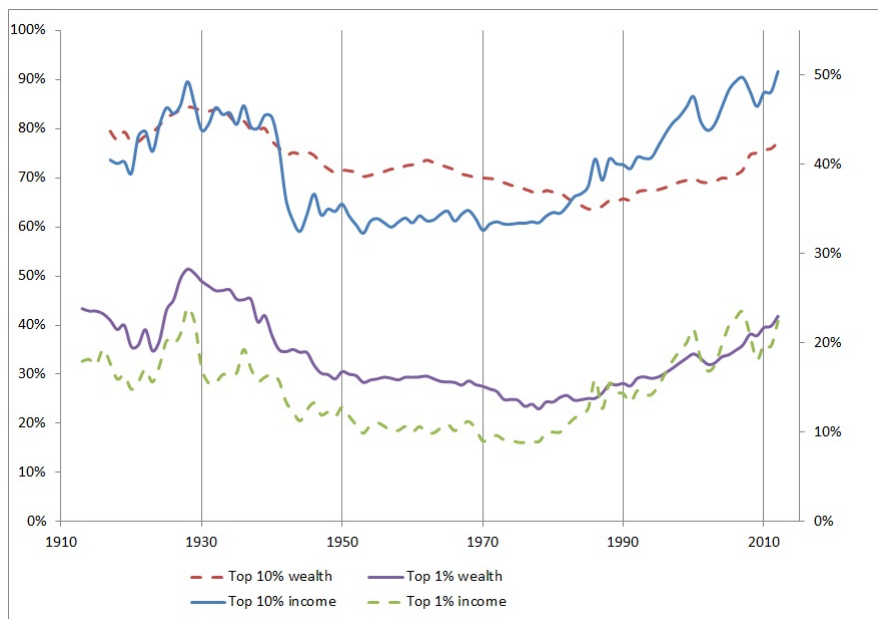

Figure 9: US top income and wealth shares
The correlation for 10% (1%) is 0.52 (0.77)

To measure income inequality, I employ the Gini indexes from the Standardized World Income Inequality Database (SWIID) and top income shares from the World Top Incomes Database. The SWIID provides comparable Gini indexes of gross and net income inequality for 173 countries for as many years as possible from 1960. The World Top Incomes Database includes 45 countries for over a century for some countries.

The last issue is how to measure the size of finance. I construct the share of financial industry employment in total employment using two datasets: the International Labour Organization (ILO) dataset, which contains employment by economic activity for 165 countries starting from 1968, and

the STAN Database for Structural Analysis, which contains industry-level data for employment and output for 15 OECD countries from the 1970s up to the present.

After conducting panel unit root tests, such as the Fisher combination test (Maddala and Wu, 1999) and the Pesaran (2007) panel unit root test, I conclude that real GDP per capita is non-stationary; therefore I compute its growth rate. The specification of the full model is given by

$$Emp_{it} = \gamma_0 + \gamma_1 gGDP_{it} + \gamma_2 II_{it} + \gamma_3 MP_{it} + \gamma_4 VP_{it} + \varepsilon_{it}, \tag{19}$$

where Emp_{it} is the share of employment in the financial sector; $gGDP_{it}$ is the growth rate of real GDP per capita; II_{it} is the measure of income inequality (Gini or top shares); MP_{it} is the mean PISA score; and VP_{it} is the variance of the PISA score.

Table 3: The share of finance (in % of total employment)

	(1)	(2)	(3)	(4)
Top 5% income share	0.00460***	0.00404***		
	(6.43)	(6.00)		
Top 1% income share			0.00466***	
			(5.43)	
Gini index				0.000613*
				(2.41)
GDPPC growth	0.0139	0.00693	-0.0239	0.00738
	(0.24)	(0.11)	(-0.47)	(0.41)
Variance of PISA score		0.00189*	0.00195***	0.00284***
		(2.31)	(3.73)	(5.96)
Mean of PISA score		0.000284	0.000556***	0.000370***
		(1.49)	(5.18)	(4.47)

t statistics in parentheses

* ($p<0.05$), ** ($p<0.01$), *** ($p<0.001$)

The results from Table 3 are consistent with my model. A higher share of financial employment is associated with more unequal income and talent distributions. The estimation results from the biggest sample, the last column with the Gini coefficient, show this clearly. Income inequality has an even stronger effect if I use the top income share instead of the Gini coefficient (columns 2 and 3). The result is not driven by country fixed effects (FE). We can see by comparing column 1 with FE and column 2 that the estimated coefficient of the top 5% share remains positive, significant and almost unchanged.

4   Discussion and conclusion

In this section, I review the impact of the different assumptions on the outcome of the model: the inefficiency result and the inequality result. First, the inefficiency result states that the decentralized equilibrium is generically inefficient. Second, the inequality result states that the endogenous growth of wealth inequality leads to the expansion of the financial sector.

Preferences and technology: First, for simplicity I assume the Cobb–Douglas production function $z^{\alpha_z} k^{\alpha_k}$, which satisfies the supermodularity condition. However, the choice of a production function should not affect the results, because as long as z and k are not fully substitutable, the supermodularity condition holds. (See Topkis (1998) for a comprehensive mathematical treatment of supermodularity, and Milgrom and Roberts (1990) and Vives (2005) for applications in game theory and economics.) According to the Becker–Brock theorem, the supermodularity condition implies that positive assortative matching is the first best allocation of my model. Hence, the results of the model remains unchanged as long as the production function is supermodular. Second, it should not be complicated to include labor as an additional input, but it would not add further insights on the questions addressed in this paper. It should not affect the choice of talented agents, but it might have interesting implications for ordinary agents. Third, if we consider a risk averse utility function instead of a risk neutral one, all agents would like to engage in risk sharing. If profit sharing and fund pooling are available options, the introduction of risk aversion does not change anything, because expected and realized incomes are the same. If these options are not available, the impact of risk aversion is ambiguous. On the one hand, investors are willing to pay a higher price for intermediated matching. The higher is the price, the higher is the income of a banker. On the other hand, due to higher uncertainty with respect to this income, risk aversion makes a banking career a less attractive option.

Distribution of types: First, as long as within each period the wealth distribution is independent from the talent distribution, the investment decision is independent from the occupational choice. This makes the solution of the problem tractable. The consideration of the two-dimensional joint distribution of wealth and talent complicates the analysis enormously without much additional insight for this particular question. Second, the fact that the constrained efficient allocation admits only two values is an artifact of the discrete distribution of talent and the particular type of information advantage for talented agents in finance: a banker with ability z can distinguish between ideas with productivity z and $z' < z$. As long as both assumptions hold, the constrained efficient allocation admits two values (zero and one-half) for each type of talent: the planner would find it optimal either to keep the allocation under random matching or make it as close as possible to the allocation under assortative matching by exhausting fully the opportunities for intermediated matching. The allocation in the case of a continuous talent distribution would strongly depend on the assumption made with respect to the impact of talent on agents' productivity in the two sectors. Third, I can show what happens in the case of a continuous wealth distribution. In my numerical simulations, it does nothing: the constrained efficient solution admits two values; the decentralized solution is inefficient.

Different types of frictions: First, this paper focuses on how the financial sector arises as a result of one type of relevant friction, adverse selection. The financial sector clearly provides other useful functions to the economy: it allocates not only information, but also decision power and risk. On the theoretical side, the financial sector's functions include: screening to mitigate the effect of adverse selection; monitoring to prevent the effects of moral hazard; auditing and punishment to mitigate the effects of opportunistic behavior in the context of costly state verification. (See Freixas and Rochet (2008) for a precise survey of the theoretical literature.) Comparing my model with that of Bolton et al. (2011), who study moral hazard and the financial sector as a liquidity provider, while I consider adverse selection and the financial sector as a classical intermediary, we both obtain a similar result in terms of efficiency, but the mechanisms are substantially different. This suggests that the misallocation result might be a general feature of models with financial frictions. Under the assumption that talent in finance affects the efficiency of monitoring, the inequality result is likely to survive as well. Second, many-to-one matching can be easily introduced into the environment in at least two ways: through diminishing returns on capital and fixed costs of engaging with investors; or through making entrepreneurs' bargaining power depend positively on the number of investors in the market. Ceteris paribus, it is likely that many-to-one matching would lead to more inefficiency in comparison to one-one matching. The more investors can be matched with one entrepreneur, the fewer bankers are needed to restore efficiency. The income of a banker increases with the number investors matched with one entrepreneur. Hence, an even larger fraction of talented agents is attracted to finance. The inefficiency should increase due to both a decline in the constrained efficient fraction of talented agents in finance and a rise in the decentralized one.

Third, the issue of competition has been studied extensively. Monopoly is usually viewed as a bad thing. However, in my framework one monopolistic firm in the financial sector might restore efficiency, because it maximizes the total surplus by pushing all agents to their outside option. The monopolist is always on the short side of the market. It would set the prices for its services to make both entrepreneurs and investors indifferent between paying for the services and being matched, and being randomly matched for free. On top of this, the monopolist can set wages for its workers (bankers) to make them indifferent between the two sectors. Hence, the monopolist could extract the total surplus and would hire the efficient number of bankers. However, this possible advantage of a monopoly in the context of information provision does not overcome common disadvantages of monopoly for a society.

4.1   Conclusion

This paper develops a new model of an economy with a financial sector and heterogeneous agents. The model sheds light on the role of the financial sector and its implication for the allocation of capital between entrepreneurs and the allocation of talent between finance and industry. Talent is important for both industry and the financial sector: more talent in industry means more output is produced, while more talent in finance means capital is allocated more efficiently. The model establishes a link between the growth of the financial sector and the increase in wealth

inequality. It shows that the market overproduces finance, but this inefficiency can be corrected by taxing bankers' income. For the future, it would be interesting to quantitatively assess the size of this inefficiency and the importance of current wealth in limiting investment.

References

ACKERBERG, D., K. CAVES, AND G. FRAZER (2006): "Structural identification of production functions," MPRA Paper 38349, University Library of Munich, Germany.

BAUER, C. AND J. V. R. MORA (2014): "The Joint Determination of TFP and Financial Sector Size," Working paper.

BECKER, G. S. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81, 813–46.

BERKES, E., U. PANIZZA, AND J.-L. ARCAND (2012): "Too Much Finance?" IMF Working Papers 12/161, International Monetary Fund.

BIENZ, C. AND J. HIRSCH (2011): "The Dynamics of Venture Capital Contracts," *Review of Finance*, 16, 157–195.

BOLTON, P., T. SANTOS, AND J. A. SCHEINKMAN (2011): "Cream skimming in financial markets," NBER Working Papers 16804, National Bureau of Economic Research, Inc.

BRUNNERMEIER, M. K., T. M. EISENBACH, AND Y. SANNIKOV (2012): "Macroeconomics with Financial Frictions: A Survey," NBER Working Papers 18102, National Bureau of Economic Research, Inc.

BUERA, F. J. AND Y. SHIN (2013): "Financial Frictions and the Persistence of History: A Quantitative Exploration," *Journal of Political Economy*, 121, 221 – 272.

CAHUC, P. AND E. CHALLE (2012): "Produce Or Speculate? Asset Bubbles, Occupational Choice, And Efficiency," *International Economic Review*, 53, 1105–1131.

COOLEY, T. F., R. MARIMON, AND V. QUADRINI (2013): "Risky Investments with Limited Commitment," NBER Working Papers 19594, National Bureau of Economic Research, Inc.

FREDRIKSEN, K. B. (2012): "Less Income Inequality and More Growth - Are they Compatible? Part 6. The Distribution of Wealth," OECD Economics Department Working Papers No. 929, OECD.

FREIXAS, X. AND J.-C. ROCHET (2008): *Microeconomics of Banking, 2nd Edition*, MIT Press Books, The MIT Press.

GENNAIOLI, N., A. SHLEIFER, AND R. W. VISHNY (2013): "Finance and the Preservation of Wealth," NBER Working Papers 19117, National Bureau of Economic Research, Inc.

GERTLER, M. AND N. KIYOTAKI (2010): "Financial Intermediation and Credit Policy in Business Cycle Analysis," in *Handbook of Monetary Economics*, ed. by B. M. Friedman and M. Woodford, Elsevier, vol. 3 of *Handbook of Monetary Economics*, chap. 11, 547–599.

GOLDIN, C. AND L. F. KATZ (2008): "Transitions: Career and Family Life Cycles of the Educational Elite," *American Economic Review*, 98, 363–69.

GOLDSMITH, R. (1995): *Financial Structure and Development*, vol. 1, Yale University Press.

GORDON, P. (2000): "The Americas," *Global Finance*, 14.

GOURINCHAS, P.-O. AND J. A. PARKER (2002): "Consumption Over the Life Cycle," *Econometrica*, 70, 47–89.

GREENSPAN, A. (2002): "World Finance and Risk Management," http://www.federalreserve.gov/BoardDocs/Speeches/2002/200209253/default.htm.

GREENWOOD, R. AND D. SCHARFSTEIN (2013): "The Growth of Finance," *Journal of Economic Perspectives*, 27, 3–28.

HOCH, I. (1962): "Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data," *Econometrica*, 30, pp. 34–53.

HOSIOS, A. J. (1990): "On the Efficiency of Matching and Related Models of Search and Unemployment," *Review of Economic Studies*, 57, 279–98.

HSIEH, C.-T. AND P. J. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124, 1403–1448.

JOVANOVIC, B. (2014): "Misallocation and Growth," *American Economic Review*, 104, 1149–71.

JOVANOVIC, B. AND B. SZENTES (2013): "On the Market for Venture Capital," *Journal of Political Economy*, 121, 493 – 527.

KAPLAN, S. N. AND P. STROMBERG (2003): "Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts," *Review of Economic Studies*, 70, 281–315.

KNEER, C. (2012): "The Absorption of Talent into Finance: Evidence from U.S. Banking Deregulation," *Job Market Paper of Tilburg University*, 1–28.

LEVINE, R. (2005): "Finance and Growth: Theory and Evidence," in *Handbook of Economic Growth*, ed. by P. Aghion and S. Durlauf, Elsevier, vol. 1 of *Handbook of Economic Growth*, chap. 12, 865–934.

LEVINSOHN, J. AND A. PETRIN (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70, 317–341.

LIZAL, L. AND K. GALUSCAK (2012): "The Impact of Capital Measurement Error Correction on Firm-Level Production Function Estimation," William Davidson Institute Working Papers Series wp1026, William Davidson Institute at the University of Michigan.

LUCAS, ROBERT E, J. AND L. A. RAPPING (1969): "Real Wages, Employment, and Inflation," *Journal of Political Economy*, 77, 721–54.

Lucas, R. J. (1988): "On the mechanics of economic development," *Journal of Monetary Economics*, 22, 3–42.

Maddala, G. S. and S. Wu (1999): " A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test," *Oxford Bulletin of Economics and Statistics*, 61, 631–52.

Mayer-Haug, K., S. Read, J. Brinckmann, N. Dew, and D. Grichnik (2013): "Entrepreneurial talent and venture performance: A meta-analytic investigation of SMEs," *Research Policy*, 1–23.

Merton, R. C. (1995): *The Global Financial System: A Functioning Perspective*, Harvard Business School Press.

MGI (2011): "Big data: The next frontier for innovation, competition, and productivity," The mckinsey global institute report, The McKinsey Global Institute.

Michelacci, C. and J. Suarez (2004): "Business Creation and the Stock Market," *Review of Economic Studies*, 71, 459–481.

Midrigan, V. and D. Y. Xu (2014): "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, 104, 422–58.

Milgrom, P. and J. Roberts (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, 1255–77.

Mundlak, Y. (1961): "Empirical Production Function Free of Management Bias," *Journal of Farm Economics*, XLIII, 44–56.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1991): "The Allocation of Talent : Implications for Growth," *The Quarterly Journal of Economics*, 106, 503–530.

Nash, J. (1950): "The Bargaining Problem," *Econometrica*, 18, 155–162.

——— (1953): "Two-Person Cooperative Games," *Econometrica*, 21, 128–140.

Nordhaus, W. D. (1982): "Economic policy in the face of declining productivity growth," *European Economic Review*, 18, 131–157.

Obama, B. (2009): "The President Explains His Larger Vision on the Economy," http://www.whitehouse.gov/video/The-President-Explains-His-Larger-Vision-on-the-Economy.

OECD (2011): "OECD Science, Technology and Industry Scoreboard," Oecd report, OECD.

Olley, G. S. and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–97.

Pesaran, M. H. (2007): "A simple panel unit root test in the presence of cross-section dependence," *Journal of Applied Econometrics*, 22, 265–312.

PHILIPPON, T. (2010): "Financiers versus Engineers: Should the Financial Sector be Taxed or Subsidized?" *American Economic Journal: Macroeconomics*, 2, 158–182.

PHILIPPON, T. AND A. RESHEF (2012): "Wages and Human Capital in the U.S. Finance Industry: 1909–2006," *The Quarterly Journal of Economics*, 127, 1551–1609.

PIKETTY, T. AND G. ZUCMAN (2014): "Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010," *The Quarterly Journal of Economics*.

RESTUCCIA, D. AND R. ROGERSON (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11, 707–720.

RODRIGUEZ, S. B., J. DIAZ-GIMENEZ, V. QUADRINI, AND J.-V. RIOS-RULL (2002): "Updated facts on the U.S. distributions of earnings, income, and wealth," *Quarterly Review*, 2–35.

ROGERSON, R. (2008): "Structural Transformation and the Deterioration of European Labor Market Outcomes," *Journal of Political Economy*, 116, 235–259.

ROMER, P. M. (1986): "Increasing Returns and Long-Run Growth," *The Journal of Political economy*, 94, 1002–1037.

SAHLMAN, W. A. (1990): "The structure and governance of venture-capital organizations," *Journal of Financial Economics*, 27, 473–521.

SANTARELLI, E., L. KLOMP, AND A. THURIK (2006): "GibratâĂŹs Law: An Overview of the Empirical Literature," in *Entrepreneurship, Growth, and Innovation*, ed. by E. Santarelli, Springer US, vol. 12 of *International Studies in Entrepreneurship*, 41–73.

SCHUMPETER, J. A. (1934): *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interestand the Business Cycle*, Transaction Publishers.

SHU, P. (2012): "Essays on Innovation, Productivity, and Talent Allocation," *MIT Job Market Paper*.

TOBIN, J. (1984): "On the efficiency of the financial system," *Lloyds Bank Review*, 153, 1âĂŞ15.

TOPKIS, D. M. (1998): *Supermodularity and Complementarity*, Princeton University Press, first edition ed.

VIVES, X. (2005): "Complementarities and Games: New Developments," *Journal of Economic Literature*, 43, 437–479.

VOLCKER, P. (2009): "Paul Volcker: Think More Boldly," http://online.wsj.com/articles/SB10001424052748704825504574586330960597134.

WADHWA, V., B. A. RISSING, AND G. GEREFFI (2006): "Industry Trends in Engineering Offshoring," report, Duke University.

WOOLDRIDGE, J. M. (2009): "On estimating firm-level production functions using proxy variables to control for unobservables," *Economics Letters*, 104, 112–114.

YANG, F. (2009): "Consumption over the Life Cycle: How Different is Housing?" *Review of Economic Dynamics*, 12, 423–443.

List of Figures

List of Tables

A   Appendix A

A.1   Proof of Proposition 1

The proof is organized in the following way. First, I show that as long as γ is greater than $1/2$, aggregate output $Y(\gamma)$ decreases with γ. Second, depending on other parameters, $Y(\gamma)$ is either a strictly increasing, or strictly decreasing, or a non-monotonic function of γ for the whole interval $\gamma \in [0, 1/2]$. In the latter case, I prove that the function is a convex function for $\gamma \in [0, 1/2]$.

For $\gamma > 1/2$, aggregate output $Y(\gamma)$ is given by

$$Y(\gamma) = \frac{(\beta^i - (1-\gamma)\beta^e)}{(1-(1-\gamma)\beta^e)}(1-\beta^e)F_{LH} + \frac{(1-\beta^i)}{1-(1-\gamma)\beta^e}(1-\beta^e)F_{LL} + (1-\gamma)\beta^e F_{HH}. \tag{20}$$

To show whether it is increasing or decreasing, we take the derivative of (20) with respect to γ:

$$\frac{\partial Y}{\partial \gamma}\Big|_{\gamma>1/2} = \frac{\beta^e(1-\beta^i)(1-\beta^e)}{(1-(1-\gamma)\beta^e)^2}F_{LH} - \beta^e F_{HH} - \frac{\beta^e(1-\beta^i)}{(1-(1-\gamma)\beta^e)^2}F_{LL}. \tag{21}$$

We need to estimate the sign of expression (21). Since F is a supermodular function, $F_{LL} \geq 0$ and $F_{HH} \geq F_{LH}$. Hence

$$\frac{\partial Y}{\partial \gamma}\Big|_{\gamma>1/2} \leq \frac{\beta^e(1-\beta^i)(1-\beta^e)}{(1-(1-\gamma)\beta^e)^2} - \beta^e. \tag{22}$$

Expressing the two terms on the right-hand side of condition (22) using a common denominator, then replacing γ in the numerator by $1/2$ (this value makes the numerator as small as possible), and finally expanding, we obtain:

$$\frac{\partial Y}{\partial \gamma}\Big|_{\gamma>1/2} \leq \beta^e \frac{(1-\beta^e)(1-\beta^i)-(1-\beta^e+\gamma\beta^e)^2}{(1-(1-\gamma)\beta^e)^2} \leq \beta^e \frac{(1-\beta^e)(1-\beta^i)-(1-\beta^e+0.25(\beta^e)^2)}{(1-(1-\gamma)\beta^e)^2} \leq 0. \tag{23}$$

Here is the end of the proof of the first part. To prove the second part, I follow a similar procedure. I calculate the first derivative and restrict my attention to the case in which the first derivative is neither positive or negative for the whole interval $\gamma \in [0, 1/2]$. Then, I show that in this case the second derivative is positive, i.e. the function is convex.

For $\gamma \in [0, 1/2]$, the aggregate output $Y(\gamma)$ is given by

$$Y(\gamma) = \frac{(\beta^i - \gamma\beta^e)}{(1-\gamma\beta^e)}\left[(1-2\gamma)\beta^e F_{HH} + (1-\beta^e)F_{LH}\right] + \\ \frac{(1-\beta^i)}{(1-\gamma\beta^e)}\left[(1-2\gamma)\beta^e F_{HL} + (1-\beta^e)F_{LL}\right] + \gamma\beta^e F_{HH}. \tag{24}$$

Calculating the first derivative from (24), we obtain

$$\frac{\partial Y}{\partial \gamma}\Big|_{\gamma \in [0,1/2]} = \frac{1}{(1-\gamma\beta^e)^2}\left[\beta^e F_{HH}\left(1+2\gamma\beta^e - (\gamma\beta^e)^2 + \beta^i\beta^e - \beta^e + 2\beta^i\right) - \\ (1-\beta^e)\beta^e(1-\beta^i)(F_{LH} - F_{LL}) - (2-\beta^e)\beta^e(1-\beta^i)F_{HL}\right]. \tag{25}$$

The first derivative is negative for the whole interval $\gamma \in [0, 1/2]$ if

$$\beta^e F_{HH}\left(1+2\gamma\beta^e - (\gamma\beta^e)^2 + \beta^i\beta^e - \beta^e + 2\beta^i\right) - (2-\beta^e)\beta^e(1-\beta^i)F_{HL} < \\ (1-\beta^e)\beta^e(1-\beta^i)(F_{LH} - F_{LL}). \tag{26}$$

The left-hand side of inequality (26) increases with γ, while the right-hand side of inequality (26) is independent of γ. Hence if inequality (26) holds for $\gamma = 1/2$, it holds for any $\gamma \in [0, 1/2]$.

$$
\begin{aligned}
[\beta^e F_{HH} \left(1 - 0.25(\beta^e)^2 + \beta^i \beta^e + 2\beta^i\right) - \\
(1 - \beta^e)\beta^e(1 - \beta^i)(F_{LH} - F_{LL}) - (2 - \beta^e)\beta^e(1 - \beta^i)F_{HL}] < 0
\end{aligned}
\tag{27}
$$

If inequality (27) holds, the first derivative is negative. In the opposite case, the sign of the derivative is unknown. Inequality (27) imposes the restriction on the set of exogenous parameters.

We now calculate the second derivative and check its sign:

$$
\begin{aligned}
\frac{\partial^2 Y}{\partial \gamma^2}\big|_{\gamma \in [0,1/2]} = \frac{2\beta^e}{(1 - \gamma\beta^e)^3} \left[\beta^e F_{HH} \left(1 + \beta^i \beta^e + 2\beta^i\right) - \right. \\
\left. (1 - \beta^e)\beta^e(1 - \beta^i)(F_{LH} - F_{LL}) - (2 - \beta^e)\beta^e(1 - \beta^i)F_{HL}\right].
\end{aligned}
\tag{28}
$$

If the second derivative is positive, the function is convex. I need to show that the right-hand side of (28) is positive. As we have seen, when inequality (27) does not hold, the sign of the first derivative is unknown, but it imposes the restriction on the set of exogenous parameters. This is the case in which we need to know the sign of the second derivative. We can estimate the right-hand side using the complementary inequality to condition (27):

$$
\begin{aligned}
\beta^e F_{HH} \left(1 + \beta^i \beta^e + 2\beta^i\right) - (2 - \beta^e)\beta^e(1 - \beta^i)F_{HL} - (1 - \beta^e)\beta^e(1 - \beta^i)(F_{LH} - F_{LL}) \geq \\
\beta^e F_{HH} 0.25(\beta^e)^2 > 0.
\end{aligned}
\tag{29}
$$

This completes the proof of the first part. We show that the sign of the first derivative is either negative or unknown. In the case in which it is unknown, we prove that the second derivative is strictly positive. Hence, the solution of the planner's problem can be either 0 or 1/2. □

A.2 Proof of Proposition 2

The characterization of a decentralized equilibrium is the following triplet: two prices and the share of talented agents in finance (p^i, p^e, γ). We have:

$$
\begin{aligned}
p^i &\leq (1 - \delta)(F_{HH} - d^i(\gamma) - d^e(\gamma)), \\
p^e &\leq \delta(F_{HH} - d^i(\gamma) - d^e(\gamma)), \\
\frac{\min \gamma, 1 - \gamma}{\gamma}(p^i + p^e) &= \delta\left(F_{HH} - d^i(\gamma)\right) + (1 - \delta)d^e(\gamma).
\end{aligned}
\tag{30}
$$

As a reminder, there are three types of agents who affect the solution of intermediated matching: capital-abundant investors, talented entrepreneurs and bankers. The number of investors is β^i, the equilibrium number of bankers is $\gamma\beta^e$, and the equilibrium number of entrepreneurs is $(1 - \gamma)\beta^e$. There are two markets and consequently two prices that clear them: entrepreneur–banker and investor–banker. System (30) should be solved differently depending on who is on the short side of both markets: capital-abundant investors, talented entrepreneurs or bankers. I show that a solution exists only if capital-abundant investors are on the long side of the investor–banker market. Furthermore, the condition $\gamma^* = 1/2$ imposes an additional restriction on the set of exogenous parameters, and eliminates a possible solution with bankers being on the short side with

respect to talented entrepreneurs on the entrepreneur–banker market. However, the solution does not always exist. I state the existence condition as well.

The system can be solved backwards. First, we need to define who is on the short side of the market: capital-abundant investors, talented entrepreneurs or bankers. Second, I solve the random matching problem for a given size of the financial sector $\gamma\beta^e$ to determine the outside options of capital-abundant investors $d^i(\gamma)$ and talented entrepreneurs $d^e(\gamma)$ in the case in which they decide not to be matched with a high-type counterpart with certainty through a banker. Third, using the solution of random matching as outside options, I solve the intermediated matching problem for capital-abundant investors and talented entrepreneurs.

Capital-abundant investors are on the short side: The number of capital-abundant investors is lower than the number of bankers who provide services for investors $\beta^i < \gamma\beta^e$. Hence, competition among bankers drives the price p^i down to zero. Furthermore, the number of bankers cannot be greater than the number of talented entrepreneurs. Otherwise, bankers' income is zero, and any talented agent strictly prefers to be an entrepreneur. Thus, if capital-abundant investors are on the short side of the investor–banker market, the share of talented agents in finance must be $\gamma \leq 1/2$.

If $\gamma \leq 1/2$ and $p^i = 0$, the system (30) collapses to one condition:

$$d^e(\gamma) = 0. \tag{31}$$

Condition (31) does not hold unless $\delta = 0$. Hence, capital-abundant investors cannot be on the short side in equilibrium.

Capital-abundant investors are on the long side: The number of capital-abundant investors is higher than the number of bankers who provide services for investors $\beta^i \geq \gamma\beta^e$. Hence, bankers push capital-abundant investors to their outside options. The first equation of system (30) becomes an equality. Two cases are possible.

First, if the number of bankers is lower than the number the talented entrepreneurs $\gamma \leq 1/2$, $1 - 2\gamma\beta^e$ talented entrepreneurs are left for random matching. We assume that investors have no access to a storage technology, while entrepreneurs have no opportunity for outside borrowing. Thus, the outside options for a random match—the set D in (5)—are $(0,0)$. The solution of the Nash bargaining problem gives the value of random matching for capital-abundant investors, which is equal to the probability of matching with somebody $\frac{1-2\gamma\beta^e}{1-\gamma\beta^e}$ multiplied by the sum of two terms: the probability of matching with a talented entrepreneur $\frac{(1-2\gamma)\beta^e}{1-2\gamma\beta^e}$ multiplied by the fraction of the project's output received by the investor $(1-\delta)F_{HH}$; and the probability of matching with an ordinary entrepreneur $\frac{1-\beta^e}{1-2\gamma\beta^e}$ multiplied by the fraction of the project's output received by the investor $(1-\delta)F_{LH}$:

$$d^i = \frac{1 - 2\gamma\beta^e}{1 - \gamma\beta^e} \frac{1 - \delta}{1 - 2\gamma\beta^e} \left[(1 - 2\gamma)\beta^e F_{HH} + (1 - \beta^e)F_{LH} \right]. \tag{32}$$

A similar expression can be obtained for the talented entrepreneur. The probability of matching with somebody for a talented entrepreneur is equal to 1, so

$$d^e = \frac{\delta}{1 - \gamma\beta^e} \left[(\beta^i - \gamma\beta^e)F_{HH} + (1 - \beta^i)F_{HL} \right].$$

Due to the supermodularity of the output function, sorting is possible. There exists a separating equilibrium such that the incentive compatibility constraint for the capital-scarce investor holds (the low type has no incentive to mimic the high type). The capital-abundant investor is indifferent between being randomly matched and being matched by a banker, while the capital-scarce investor is strictly better off under random matching. In this case, the system (30) takes the form below:

$$p^i = (1 - \delta)(F_{HH} - d^i(\gamma) - d^e(\gamma)),$$
$$p^e = \delta(F_{HH} - d^i(\gamma) - d^e(\gamma)),$$
$$(p^i + p^e) = \delta(F_{HH} - d^i(\gamma)) + (1 - \delta)d^e(\gamma). \tag{33}$$

Surprisingly, as long as $\gamma \leq 1/2$, the income of a banker is an increasing function of the number of bankers, while the income of an entrepreneur is a decreasing function of the number of bankers. The rise of bargaining power δ has no effect on the banker's income and a positive one on entrepreneurial income. The solution of the system (33) is linear in δ:

$$\tilde{\gamma} = \delta \left[\frac{1}{\beta^e} + \frac{1 - \beta^e}{\beta^e} \frac{F_{HH} - F_{LH}}{F_{HH}} - \frac{2(1 - \beta^i)}{\beta^e} \frac{F_{HH} - F_{HL}}{F_{HH}} \right] - \frac{1 - \beta^e}{\beta^e} \frac{F_{HH} - F_{LH}}{F_{HH}}. \tag{34}$$

There exist two thresholds $\bar{\delta} > 0$, such that $\tilde{\gamma} = 0$, and $\tilde{\delta} > 0$, such that $\tilde{\gamma} = 1/2$:

$$\bar{\delta} = \frac{(1 - \beta^e)(F_{HH} - F_{LH})}{(1 - \beta^e)(F_{HH} - F_{LH}) + (2\beta^i - 1)(F_{HH} - F_{HL}) + F_{HL}}, \tag{35}$$

$$\tilde{\delta} = \frac{(1 - \beta^e/2)F_{HH} - (1 - \beta^e)F_{LH}}{(1 - \beta^e)(F_{HH} - F_{LH}) + (2\beta^i - 1)(F_{HH} - F_{HL}) + F_{HL}}. \tag{36}$$

Depending on parameter values, both $\bar{\delta}$ and $\tilde{\delta}$ can potentially be greater than 1. The solution $\tilde{\gamma}$ exists only for $\delta \in [\min[\bar{\delta}, 1], \min[\tilde{\delta}, 1]]$. The solution $\tilde{\gamma}$ exists as long as $\bar{\delta} \leq 1$. Using expression (35), the latter can be rewritten as follows:

$$(2\beta^i - 1)(F_{HH} - F_{HL}) + F_{HL} \geq 0. \tag{37}$$

Second, the number of bankers is greater than or equal to the number of talented entrepreneurs $\gamma \geq 1/2$. Thus, all talented entrepreneurs are matched by bankers. The number of capital-abundant investors $\beta^e(1 - \gamma)$ are matched by bankers. In this case, the solution of the Nash bargaining problem for random matching is given by

$$d^i = \frac{(1 - \beta^e)}{1 - \beta^e(1 - \gamma)}(1 - \delta)F_{LH}, \tag{38}$$

$$d^e = \frac{\delta}{1 - \beta^e + \gamma\beta^e} \left[(\beta^i - \beta^e + \gamma\beta^e)F_{HH} + (1 - \beta^i)F_{HL} \right]. \tag{39}$$

If the number of bankers is greater than the number of talented entrepreneurs, competition among bankers drives the price p^i down to zero. In this case, system (30) takes the form below:

$$p^i = (1 - \delta)(F_{HH} - d^i(\gamma) - d^e(\gamma)),$$
$$p^e = 0,$$
$$\frac{1 - \gamma}{\gamma} p^i = \delta(F_{HH} - d^i(\gamma)) + (1 - \delta)d^e(\gamma). \tag{40}$$

As we can see, the banker's income is a decreasing function of the bargaining power δ, while entrepreneurial income is an increasing function of δ. Furthermore, the expected income of a banker grows with γ. System (40) can be expressed in the form of a quadratic equation in γ:

$$
\gamma^2 + \left[(1-\delta)\frac{(1-\beta^e)}{\beta^e}\frac{F_{HH}-F_{LH}}{F_{HH}} + \frac{\delta-\beta^e}{\beta^e} + \delta(1-\delta)\right]\gamma-
$$
$$
-(1-\delta)^2\frac{(1-\beta^e)}{\beta^e}\frac{F_{HH}-F_{LH}}{F_{HH}} - (1-\beta^i)(1-\delta)\frac{\delta}{\beta^e}\frac{F_{HH}-F_{HL}}{F_{HH}} = 0. \tag{41}
$$

The solution of quadratic equation (41) contains two roots, but one root is always negative. For the solution to exist, the second root has to be greater than $1/2$. Let $\hat{\gamma}$ be the positive solution of (41). This solution exists as long as

$$
\frac{\delta(1+\beta^e(1-\delta))}{4(1-\delta)} \leq (1-\beta^i)\delta\frac{F_{HH}-F_{HL}}{F_{HH}} + (1-\beta^e)\frac{F_{HH}-F_{LH}}{F_{HH}}(1/2-\delta). \tag{42}
$$

Analyzing condition (42), we can conclude that the condition is likely to be satisfied when: the dispersion of wealth k^H/k^L is high; the share of capital-abundant investors β^i is low; and the bargaining power of entrepreneurs δ is relatively low. Furthermore, if $\delta \leq 1/2$, condition (42) is likely to be satisfied when the dispersion of talent is high and the share of talented agents β^e is low. When condition (42) is satisfied with equality, it can be rewritten as the definition of $\hat{\delta}$ defined in Proposition 2.

The constrained efficient allocation is $\gamma^* = 1/2$: This implies that $\Delta\mathbf{Y}$ given by expression (3) is positive and can be rewritten in the following form:

$$
(2\beta^i - 1)(F_{HH} - F_{HL}) + F_{HL} < -\frac{2(1-\beta^i)(1-\beta^e)}{2-\beta^e}(F_{LH} - F_{LL}). \tag{43}
$$

The right-hand side of inequality (43) is negative, therefore its left-hand side is negative as well. If we compare the left-hand side of (43) with the left-hand side side of expression (37), they are exactly the same. Hence, if $\gamma^* = 1/2$, the solution $\tilde{\gamma}$ does not exist. $\qquad\square$

A.3   The solution of the decentralized equilibrium if $\gamma^* = 0$

In this section, I show the solution of the decentralized equilibrium in the case in which the constrained efficient allocation is 0. The proposition below summarizes the case:

Proposition 3. *If the constrained efficient allocation is $\gamma^* = 0$, then both equilibria with few $\tilde{\gamma}$ and many $\hat{\gamma}$ bankers are possible. In this case, there is a range of $\delta \in [\tilde{\delta}, 1)$, such that the decentralized equilibrium is constrained efficient.*

Proof: As shown in appendix A.2, the solution $\hat{\gamma}$ exists as long as condition (42) holds, and it does not depend on whether the constrained efficient allocation is 0 or $1/2$; the solution $\tilde{\gamma}$ does not exist if $\gamma^* = 1/2$. The solution $\tilde{\gamma}$ exists as long as $\bar{\delta} < 1$, defined by expression (35). The latter is likely to be satisfied if $\gamma^* = 0$. $\qquad\square$

B Appendix B

B.1 Private banking and private equity finance

This subsection introduces an alternative way of calibrating the model. Most global banks, such as Credit Suisse, Barclays, BNP Paribas, Citibank, Deutsche Bank, HSBC, JPMorgan Chase and UBS, have a separate business unit with dedicated teams of client advisors and product specialists exclusively for high-net-worth individuals. They provide a wide range of investment opportunities, including bonds, stocks and more importantly private equity finance.

Private equity is an important channel through which long-term investments are made. It has grown steadily over the past three decades, and today private equity funds worldwide manage over $1 trillion. For some countries, such as Israel, the US and the UK, private equity accounts for more than 5% of total investment (see Table 4 for details).

Table 4: The size of private equity

	% GDP 2010	% GDP 2011	% Investment 2010	% Investment 2011
Israel	0.63	2.09	3.50	10.45
UK	1.13	0.75	7.53	5.00
US	0.9	0.98	5.00	5.44
China	0.16		0.33	
World		0.30		1.58

I would like to convince the reader that the matching assumption holds for private equity. A small private equity fund provides an opportunity to invest in a few companies over a long-term horizon for a small number of wealthy investors. As we can see from Table 5, private equity funds typically employ 12 professionals. These professionals select one or two companies each for the fund to invest in. Investments are large (over $50 million). Investors are wealthy and expected to invest over a long-term horizon. The minimum required commitment rises from a median of $1 million for funds of $100 million or less, up to a median of $10 million for funds of $1 billion or more. There is no active market for private equity positions, making these investments illiquid and difficult to value. Private equity funds typically have horizons of 10–13 years, during which the invested capital cannot be redeemed. Furthermore, based on the sample of firms from the Pricewaterhouse National Venture Capital Survey, Gordon (2000) shows that 71.32% of firms reported they had received more than one offer to invest from venture capitalists. The mean number of offers was 3.18. This means entrepreneurs have a choice about which venture capitalists invest in their companies, but this choice is rather limited and matching plays an important role.

Given the long-term horizon and the high entry costs, the question is why investors are willing to engage in these investments. Investors are compensated well by substantially higher returns. Table 6 shows that the return from an investment in private equity funds is three times higher

Table 5: Private equity funds

	N	Number of			Size ($ mn)	
		professionals	investments	I/P	Fund	Investment
VC	94	9	20	2	225	11.25
Buyout	144	13	12	1	600	50

than in stocks. We can see the comparison with inflation and the returns on other assets: stocks, gold, T-bills etc.

Table 6: US real asset returns

Period	PEF	S&P	TBond	Gold	Inflation
1997-2011	9.2%	3.2%	0.4%	7.0%	2.4%
1975-2011		7.5%	1.3%	4.0%	4.2%

The first important ingredient of the model is the wealth distribution. If we look at Table 7, we can conclude that at most 5% of households in the US can afford to invest at least $1 million for the horizon of ten years. I target 5% as the share of capital-abundant investors.

Table 7: Wealth distribution in the US

	1962	1983	1989	1998	2001	2007	2010
Average	194	284	326	362	468	564	464
Median	52	73	78	81	91	108	57
Top 20%	786	1157	1361	1507	1976	2397	2062
Top 5%	2120	3190	3841	4273	5542	6973	5842
Top 1%	6491	9599	12176	13650	15627	19486	16439
1% to median	125	131	156	168	173	181	288

The second important ingredient of the model is the distribution of talent. The most challenging exercise is to calibrate it. I use three different proxies. First, based on US firm-level data I estimate the moments of the productivity distribution in the US. Second, based on data on different test scores from the National Longitudinal Survey of Youth 1997 (NLSY97), I estimates the moments of the IQ distribution in the US. Finally, since my model contains only two levels of talent, we might think of these as being college graduates and high-school graduates. Therefore, I use data on the share of college graduates over time as a proxy for the increase in the supply of talent (skills).

B.2 Firm-level data estimation of the allocation of talent

I propose a two-step procedure to estimate the allocation of talent. In the first step, I estimate a production function to obtain firm-specific productivity, which I associate with talent. In the second step, I use firm-level productivity to construct a distribution of talent across different sectors and formally test if the distributions are different.

General remarks

I rely on the fast-growing literature on the estimation of firm productivity. In general, econometricians have identified several main issues with productivity estimation. First, there is simultaneity bias caused by the relationship between unobserved productivity shocks and production inputs. Second, there are measurement issues—in particular, the true stock of capital is difficult to measure. Third, even after the above issues are resolved, the question remains how to interpret the obtained measure of firm-level productivity. Is it a measure of true productivity or a measure of the markup due to imperfect competition, which possibly biases estimates of true productivity? Finally, the production function may be misspecified. This section addresses all of these issues.

For simplicity, let us assume the Cobb–Douglas production function augmented with capital, labor and intermediate inputs terms takes the following (stylized) form:

$$Y = AK^{\alpha}L^{\beta}X^{\gamma}, \tag{44}$$

where is A is firm productivity, L is labor, K is capital and X is an intermediate input.

After taking logs, the equation (44) transforms into, for firm i at time t:

$$y_{it} = \alpha k_{it} + \beta l_{it} + \gamma x_{it} + \omega_{it} + u_{it}, \tag{45}$$

where y is the log of output, k is the log of capital input, l is the log of labor input, and x is the log of intermediate input. The last two terms represent the observable and unobservable components of the firm's productivity from the firm's point of view. There are two terms in this equation that are unobservable to the econometrician, ω and u. The distinction between the two is important. The u term is intended to represent shocks to production or productivity that are not observable (or predictable) by firms before making their input decisions at time t. In contrast, the ω term represents shocks that are potentially observed or predictable by firms when they make input decisions. Intuitively, ω might represent variables such as the managerial ability of a firm, the expected markup, expected down-time due to machine breakdown, expected defect rates in a manufacturing process, or the expected rainfall at a farm's location. On the other hand, u is the deviation from the expected level. The shock ω_{it} affects the choice of inputs, leading to a simultaneity problem in the estimation.

The two most commonly used solutions to this endogeneity problem are fixed effect regressions (Mundlak, 1961; Hoch, 1962) and instrumental variables estimation techniques. Under the

assumption that ω is constant over time or contains a common trend for all firms, the fixed effects estimator is consistent and unbiased, but in practice it often generates unrealistically low estimates of β_k. On the one hand, there do exist natural instrumental variables in this situation—input prices, as long as one is willing to assume firms operate in competitive input markets—but they are not available in my datasets. I use fixed effects estimation as my benchmark.

The second type of technique relies on structural estimation. Two approaches are popular in the literature: Olley and Pakes (1996) (OP) and Levinsohn and Petrin (2003) (LP). The main idea is to model the choice of inputs to find a control function. OP employ investment as a function of ω and other observed variables; LP argue that investment is often zero, as is the case for my dataset, so they propose to use intermediate inputs instead of investment to form the control function. However, Ackerberg et al. (2006) shows that the estimation of OP and especially LP suffer from a multicollinearity problem. To correct the problem, Wooldridge (2009) proposes to use a more efficient joint estimator.

Data

I employ two sources of cross-country firm-level data: FactSet and Amadeus. FactSet Fundamentals provides information about income and balance sheets for more than 73,000 mostly publicly listed companies and for more than 200 countries during the period from 1990 to 2012. However, after data cleaning, only 28,840 firms in 77 countries remain. Four countries—the US, Japan, China, and the UK—together represent roughly half of the dataset.

Measurement issues

For robustness, I employ two measures of value added Y. The first uses the accounting approach, such as sales and changes in inventories and new investments subtracting the cost of goods sold. The second uses the economic approach, such as profit, depreciation and salaries combined. The results are not particularly sensitive to the different measures of value added, so I stick with the accounting measure. Labor is defined as the number of employees. As has been shown in literature, the biggest challenge is to correctly measure capital. I use the sum of equity and long-term debt as a measure of capital, and depreciation as an instrument. The last measurement issue is deflation. Two alternative deflators are employed: US inflation and the effective interest rate, which is firm specific.

Table 8 provides summary statistics for the variables of interest. The dataset contains the largest employers, such as WalMart (which employed 2.2 million employees), as well as very small firms. The dataset provides information about the world's biggest companies, such as Fannie Mae (with a book value above \$3,000 billion), as well as tiny companies.

Table 8: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Employment	425862	5055.18	22895.69	0	2200000
Total Assets (TA)	596229	2058140	2.01E+07	10	3.22E+09
Revenue % TA	585376	0.814834	0.668393	0	3.355362
Equity in % TA	581685	0.445831	0.383413	-2.86423	0.991044
Leverage	586456	0.22594	0.222609	0	1.316528

Estimation of the production function

The results of the estimation based on the full sample are reported in Table 9: FE is the fixed effects estimate; OP is the Olley–Pakes (1996) estimate; LP is Levinsohn–Petrin (2003) estimate; Wooldridge uses the methodology of Wooldridge (2009) adjusted for measurement errors. All regression variables are expressed in natural logarithms. The dependent variable is productivity $y = Y/L$. The independent variable is capital intensity $k = K/L$; the control variable for LP and OP is the cost of intermediate input $x = X/L$. The estimates are all remarkably similar. As expected, the FE approach slightly underestimates the share of capital. Examining Table 9, we can conclude that the capital share of production is about 87%, the labor share of production is about 8%. These results are consistent with the literature, but the share of capital is somewhat higher, while the share of labor is slightly lower, than in other papers. For example, Olley and Pakes (1996) report 61% for capital and 17% for labor. Similar results are reported by Lizal and Galuscak (2012).

Table 9: Estimation of the production function

		FE	OP	LP	Wooldridge
k		0.80127***	0.81412***	0.87071***	0.89886***
		(624.44)	(155.85)	(235.84)	(56.47)
l		0.12921***	0.08665***	0.07713***	0.04276***
		(48.13)	(27.80)	(18.60)	(9.50)
R-sqr		0.790			0.873
FE		yes			
Obs		201712	201330	201330	168716
Firms		24006	28839	28839	22393

Estimation of the talent distribution

For each of the firms in 77 countries I obtained the measure of talent z in the previous subsection. Then, I compare whether the distribution of z is different for the financial sector and the rest of the economy. Figure 10 shows the distribution of productivity in the US for financial firms and the rest of the economy. We can see that financial firms are more productive, on average.
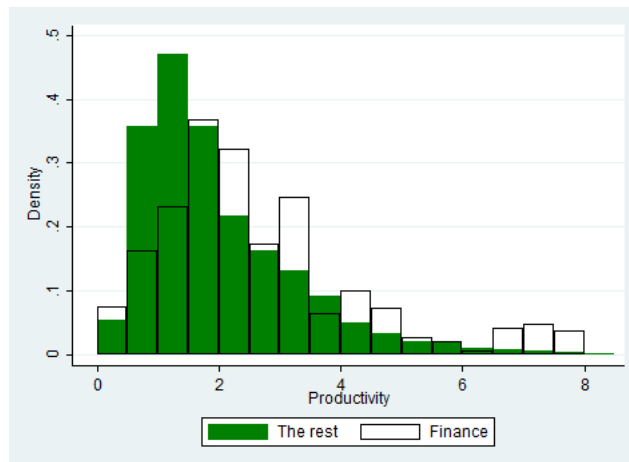
Figure 10: Distribution of productivity

Table 10 reports the descriptive statistics. The cross-country average of the ratio between average financial firms' productivity and average non-financial firms' productivity, as well as the ratio of variances. It appears that the distribution of financial firms' productivity is shifted to the right, and is more concentrated in comparison with that of non-financial firms. This could be a sign of talent misallocation.

Table 10: Firm distribution

	η_i	Mean	Variance
Ratio of financial firms' productivity to non-financial firms' productivity		1.16	0.82

Interpretation of the results

This subsection argues that the obtained measure of firm productivity can be substantially attributed to the talent of a firm's founder. There are certain factors that might affect a firm's productivity A_{it}. Let us focus on the four that are probably the most important: markups, managerial skills, R&D and human capital of labour. First, even though markups vary substantially from industry to industry, cross-county differences appear to be even more important. Markups are attributed mostly to country-sector characteristics rather than to firms. Second, Kaplan and Stromberg (2003) shows that the average founders' share equals 24.3% of a portfolio company's equity value. This means that the founder of a company, who generates the original idea, continues to be actively involved in the decision-making process. In some cases, founders may even remain the manager. Third, based on the small subsample of firms for which R&D data were available, I redo the estimation to confirm the results. Fourth, even though the composition of the labor force (skilled vs. unskilled) can be an important factor that affects a firm's productivity, it should not change the firm's ranking in the firm distribution.