Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?*

Benjamin Moll

Princeton University

September 29, 2010

Abstract

Does capital misallocation from financial frictions cause substantial aggregate productivity losses? To explore this question, I propose a highly tractable theory featuring entrepreneurs who are subject to borrowing constraints and idiosyncratic productivity shocks. Productive entrepreneurs cannot raise capital in the market; however, they may self-finance investment in the sense of paying it out of their own savings. Such self-financing can undo capital misallocation if productivity shocks are sufficiently autocorrelated. If so, financial frictions have *no effect* on aggregate productivity. Conversely, productivity losses may be large if autocorrelation is low. My model economy is further isomorphic to an aggregate growth model with the difference that productivity is endogenous and depends on the quality of credit markets. This implies that financial frictions have no direct effect on aggregate output and savings; only an indirect one through aggregate productivity. In an application of the model, I estimate its critical parameters using plant-level panel data from two emerging market economies and calculate that financial frictions can explain aggregate productivity losses of up to twenty percent relative to the US.

*I am extremely grateful to Rob Townsend, Fernando Alvarez, Paco Buera, Bob Lucas and Rob Shimer for many helpful comments and encouragement. I also thank Abhijit Banerjee, Silvia Beltrametti, Jess Benhabib, Nick Bloom, Lorenzo Caliendo, Wendy Carlin, Steve Davis, Steven Durlauf, Jeremy Fox, Veronica Guerrieri, Lars-Peter Hansen, Chang-Tai Hsieh, Erik Hurst, Oleg Itskhoki, Joe Kaboski, Anil Kashyap, Sam Kortum, David Lagakos, Guido Lorenzoni, Virgiliu Midrigan, Ezra Oberfield, Stavros Panageas, Chad Syverson, Nicholas Trachter, Harald Uhlig, Daniel Yi Xu, Luigi Zingales, and seminar participants at the University of Chicago, Northwestern, UCLA, Berkeley, Princeton, Brown, LSE, Columbia GSB, the 2009 SED and the EEA-ESEM meetings for very helpful comments. For sharing their data, I am particularly grateful to Devesh Raval (Chilean data) and Jim Tybout (Colombian data).

Introduction

Underdeveloped countries often have underdeveloped financial markets. This can lead to an inefficient allocation of capital, in turn translating into low productivity and per-capita income. But available theories of this mechanism often ignore the effects of financial frictions on the *accumulation* of capital and wealth. Even if an entrepreneur is not able to acquire capital in the market, he might just accumulate it out of his own savings. The implications of such effects are not well understood. To explore them, this paper proposes a tractable dynamic theory featuring heterogeneous entrepreneurs who are subject to borrowing constraints. It then applies the theory to quantify productivity losses from financial frictions, using plant-level panel data from two emerging market economies.

Consider an entrepreneur who begins with a business idea. In order to develop his idea, he requires some capital and labor. The quality of his idea translates into his productivity in using these resources. He hires workers in a competitive labor market. Access to capital is more difficult, due to borrowing constraints: the entrepreneur is relatively poor and hence lacks the collateral required for taking out a loan. Now consider a country with many such entrepreneurs: some poor, some rich; some with great business ideas, others with ideas not worth implementing. In a country with well-functioning credit markets, only the most productive entrepreneurs would run businesses, while unproductive entrepreneurs would lend their money to the more productive ones. In practice credit markets are imperfect so the equilibrium allocation instead has the features that the marginal product of capital in a good entrepreneur's operation exceeds the marginal product elsewhere. Reallocating capital to him from another entrepreneur with a low marginal product would increase the country's GDP. Failure to reallocate is therefore referred to as a "misallocation" of capital. Such a misallocation of capital shows up in aggregate data as low total factor productivity (TFP). Financial frictions thus have the potential to help explain differences in per-capita income.[1] Of course, resources other than capital can also be misallocated. I focus on the misallocation of capital because there is empirical evidence that this is a particularly acute problem in developing countries.[2]

The argument just laid out has ignored the fact that capital and other assets can be accumulated over time. Importantly, it has therefore also ignored the possibility of *self-financing*: an entrepreneur without access to external funds can still accumulate internal funds over time

[1]See Restuccia and Rogerson (2008) for the argument that resource misallocation shows up as low TFP. See Hsieh and Klenow (2009) for a similar argument and empirical evidence on misallocation in China and India. See Klenow and Rodríguez-Clare (1997) and Hall and Jones (1999) for the argument that cross-country income differences are primarily accounted for by low TFP in developing countries.

[2]I refer the reader to Banerjee and Duflo (2005), Banerjee and Moll (2009) and the references cited therein.

to substitute for the lack of external funds.[3] Moving to a dynamic setting therefore uncovers a counteracting force to the misallocation described in the static setting in the paragraph above. Self-financing has the potential to undo capital misallocation. As I will argue below, the outcome of the tug-of-war between self-financing and capital misallocation depends crucially on the evolution of individual productivities over time, leading me to consider a setting with idiosyncratic productivity shocks.

An equilibrium in this setting looks as follows: entrepreneurs are continually hit by productivity shocks and they try to adjust their capital to their productivity. They can do this either by borrowing and lending in the credit market, or – since this is only possible to a limited extent – by self-financing. At any point in time, the allocation of capital across heterogeneous entrepreneurs then determines aggregate TFP and GDP just as in the static setting. My main theoretical result is that self-financing is an effective substitute for credit access only if productivity shocks are sufficiently correlated over time. Conversely, if shocks are transitory, the ability of entrepreneurs to self-finance is hampered considerably. This is intuitive. While self-financing is a valid substitute to a lack of external funds, it takes time. Only if productivity is sufficiently persistent, do entrepreneurs have enough time to self-finance. The efficacy of self-financing then translates directly into productivity losses from financial frictions: if shocks are sufficiently autocorrelated, financial frictions have no effect on aggregate productivity. Conversely, productivity losses may be large if autocorrelation is low.[4]

The primary contribution of this paper is to make this argument by means of a tractable dynamic theory of entrepreneurship and borrowing constraints. In the model economy, aggregate GDP can be represented by means of an aggregate production function. The key to this result is that individual production technologies feature constant returns to scale in capital and labor. This assumption also implies that knowledge of the *share of wealth held by a given productivity type* is sufficient for assessing TFP losses from financial frictions. TFP turns out to be a simple truncated weighted average of productivities; the weights are given by the wealth

[3]See the survey by Quadrini (2009) for the argument that such self-financing motifs can explain the high concentration of wealth among entrepreneurial households. In the same spirit, Gentry and Hubbard (2004) and Buera (2009*a*) find that entrepreneurial households have higher savings rates and argue that this is due to costly external financing for entrepreneurial investment. Gentry and Hubbard remark that similar ideas go back at least to Klein (1960). In the context of developing countries, Samphantharak and Townsend (2009) find that households in rural Thailand finance a majority of their investment with cash. Pawasutipaisit and Townsend (2009) find that productive households accumulate wealth at a faster rate than unproductive ones. All this is evidence suggestive of self-financing.

[4]This result is discussed informally in Banerjee and Moll (2009) but is (to my knowledge) otherwise new. In a similar framework, Buera and Shin (2009) argue that persistence matters greatly for the welfare costs from market incompleteness. Gourio (2008) shows that also the effect of adjustment costs on aggregate productivity depends crucially on persistence. That productivity is relatively persistent is a consistent finding in much of the industrial organization literature on this topic (Baily, Hulten and Campbell, 1992; Bartelsman and Dhrymes, 1998). It holds in both developed and developing countries.

shares and the truncation is increasing in the quality of credit markets.[5] Because this aggregation result is by nature static, I first present it in a static model. I then extend the model to a dynamic setting like the one described informally earlier in this introduction: productivity is stochastic and savings are chosen optimally. Crucially, the assumption of individual constant returns delivers linear individual savings policies. The economy is then simply isomorphic to an aggregate growth model with the difference that TFP evolves endogenously over time. The evolution of TFP depends only on the evolution of wealth shares. I finally assume that the stochastic process for productivity is given by a mean-reverting diffusion. Wealth shares then obey a simple differential equation and a complete characterization is possible for particular functional forms for the diffusion process.[6]

Another contribution is an application of this theory to quantify TFP losses from financial frictions across 68 countries. Using plant-level panel data for manufacturing in Chile and Colombia, I first estimate the model's micro-parameters such as the all-important autocorrelation. Turning to my macro sample of 73 countries, I then use external finance to GDP ratios to infer the parameter governing the degree of financial development, and compare TFP predicted by the model to that in the data. I calculate that financial frictions can explain aggregate productivity losses of up to twenty-five percent relative to the US.

Related Literature A large theoretical literature studies the role of financial market imperfections in economic development. Early contributions are by Banerjee and Newman (1993), Galor and Zeira (1993), Aghion and Bolton (1997) and Piketty (1997). See Banerjee and Duflo (2005) and Matsuyama (2007) for recent surveys.[7] I contribute to this literature by developing a tractable theory with forward-looking savings, thus emphasizing the possibility of self-financing.

My paper is most closely related and complementary to a series of more recent, quantitative papers relating financial frictions to aggregate productivity (Jeong and Townsend, 2007; Buera and Shin, 2010; Buera, Kaboski and Shin, 2010; Midrigan and Xu, 2009). While there is some agreement that financial frictions can lower aggregate productivity in theory, there remains disagreement on the *size* of resulting productivity losses. For example, Buera, Kaboski and Shin (2010) calibrate a model of entrepreneurship similar to the one in this paper and argue that financial frictions can explain TFP losses of up to 40%. On the other extreme, Midrigan

[5]See Lagos (2006) for another paper providing a "microfoundation" of TFP – there in terms of frictions in the labor rather than credit market.

[6]As in Krebs (2003) and Angeletos (2007), individual linear savings policy functions imply that there is no stationary distribution of wealth. However, the wealth shares just discussed admit a stationary measure, allowing me to sidestep the nonexistence of a stationary wealth distribution.

[7]There is an even larger empirical literature on this topic. A well-known example is by Rajan and Zingales (1998). See Levine (2005) for a survey.

and Xu (2009) calibrate a model of firm dynamics with financial frictions to plant-level panel data from South Korea but conclude that for the specific data set they study, these frictions only account for relatively small TFP losses of around 2.5%.[8] To better explore the sources of such disagreement is the main goal of this paper. By allowing for a transparent analysis of the main forces at play and the parameters behind them, this is also where the tractability of my theory pays off. The tug-of-war between self-financing and capital misallocation is an example of two such forces and the autocorrelation of productivity is the corresponding critical parameter. Analytical tractability thus aids our theoretical understanding of the quantitative work on finance and development.[9]

To deliver such tractability, I build on work by Angeletos (2007) and Kiyotaki and Moore (2008). Their insight is that heterogenous agent economies remain tractable if individual production functions feature constant returns to scale because then individual policy rules are linear in individual wealth. In contrast to the present paper, Angeletos focuses on the role of incomplete markets à la Bewley and does not not examine credit constraints. Kiyotaki and Moore analyze a similar setup with borrowing constraints but focus on aggregate fluctuations. Both papers assume that productivity shocks are iid over time, an assumption I dispense with. A notable exception allowing for persistent shocks is Kiyotaki (1998). His persistence, however, comes in form of a Markov chain with only two states (productive and unproductive) which is considerably less general than in my paper.[10]

Quantitative papers such as the ones cited above usually examine micro data through the lens of a structural model (see also Giné and Townsend, 2004; Jeong and Townsend, 2008; Townsend, 2009). Most papers conduct this exercise using cross-sectional data. Instead, the present paper examines *panel* data (as do Midrigan and Xu (2009), to my knowledge the only other paper). In light of the importance of persistence and the dynamic nature of capital accu-

[8]The authors stress that this is (in their words) "not an impossibility result"; rather that parameterizations that do generate large TFP losses miss important features of the data. Also note that both their paper and Buera, Kaboski and Shin (2010) differ from mine in some modeling choices: Both papers assume decreasing returns in production whereas I assume constant returns. Buera, Kaboski and Shin (2010) feature fixed costs, occupational choice and two sectors of production, all of which are not present in my paper. Midrigan and Xu's is a partial equilibrium model with risk-neutral firms; in contrast, my model is set up in general equilibrium and firms are owned by risk-averse entrepreneurs. I revisit some of these differences – especially fixed costs – later in the paper.

[9]Tractability comes at the cost of some empirical realism. My theory should therefore be viewed as a complement to full-blown quantitative papers like the ones just discussed.

[10]Another similarity between my paper and Kiyotaki (1998) is the characterization of equilibrium in terms of the share of wealth of a given productivity type. Other papers exploiting linear savings policy rules in environments with heterogenous agents are: Banerjee and Newman (2003) who analyze trade and inequality in the presence of capital market imperfections but don't feature productivity shocks as here (implying there are no long-run effects of financial frictions); Azariadis and Kaas (2009) who examine the allocation of capital across industries; Kocherlakota (2009) whose focus is on bubbles in land price; and Krebs's (2003) analysis of human capital risk.

mulation more generally, the use of panel data seems essential for assessing productivity losses from financial frictions. Finally, I contribute to broader work on the macroeconomic effects of micro distortions (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Bartelsman, Haltiwanger and Scarpetta, 2008). Hsieh and Klenow (2009) in particular argue that misallocation of both capital and labor substantially lowers aggregate TFP in India and China. Their analysis makes use of abstract "wedges" between marginal products. In contrast, I formally model one reason for such misallocation: financial frictions resulting in a misallocation of capital.[11]

After developing a simple static model (Section 1) and extending it to a full dynamic setting (Section 2), I report my empirical results (Section 3). Section 4 is a conclusion.

1 Static Model

This section presents a simple static general equilibrium model of heterogenous entrepreneurs that are subject to collateral constraints. The model is a variant of standard static models of entrepreneurship such as Evans and Jovanovic (1989), Holtz-Eakin, Joulfaian and Rosen (1994) and Banerjee and Duflo (2005, section 5). The full dynamic model presented in the next section will embed this same static model in a dynamic framework; that is, it will have the feature that *period by period*, entrepreneurs solve the static problem presented in this section.

1.1 Setup

There is a continuum of entrepreneurs that are indexed by their productivity z and their wealth a.[12] The joint distribution of wealth and productivity is denoted by $g(a, z)$. The corresponding marginal distributions are denoted $\varphi(a)$ and $\psi(z)$. Each entrepreneur owns a private firm which uses k units of capital and l units of labor to produce

$$y = f(z, k, l) = (zk)^{\alpha} l^{1-\alpha} \tag{1}$$

units of output, where $\alpha \in (0, 1)$. There is also a mass L of workers. Each worker is endowed with one efficiency unit of labor which he supplies inelastically. Entrepreneurs hire workers in a competitive labor market at a wage w. They also rent capital from other entrepreneurs

[11] Also related – albeit different at first sight – is the theoretical literature on international financial markets. Relabeling entrepreneurs as countries, the models used there are strikingly similar. Gourinchas and Jeanne (2006) argue that distortions from imperfect capital markets are essentially transitory, the reason being the same self-financing mechanism highlighted in this paper. See Manuelli (2009), Bulow and Rogoff (1989) and Guerrieri, Lorenzoni and Perri (2009) for other examples.

[12] Here, "productivity" is a stand-in term for a variety of factors such as entrepreneurial ability, an idea for a new product, an investment "opportunity", but also demand side factor such as idiosyncratic demand shocks.

in a competitive capital rental market at a rental rate R. This rental rate equals the user cost of capital, that is $R = r + \delta$ where r is the interest rate and δ the depreciation rate.[13] Entrepreneurs face collateral constraints

$$k \leq \lambda a, \quad \lambda \geq 1. \tag{2}$$

This formulation of capital market imperfections is analytically convenient. Moreover, by placing a restriction on an entrepreneur's leverage ratio k/a, it captures the common intuition that the amount of capital available to an entrepreneur is limited by his personal assets. The constraint can also be motivated as arising from a limited enforcement problem.[14] Finally, note that by varying λ, I can trace out all degrees of efficiency of capital markets; $\lambda = \infty$ corresponds to a perfect capital market, and $\lambda = 1$ to the case where it is completely shut down. λ therefore captures the degree of financial development, and one can give it an institutional interpretation.

1.2 Individual Behavior

Entrepreneurs maximize profits. Their profit function is

$$\Pi(a, z) = \max_{k,l} \{ f(z, k, l) - wl - (r + \delta)k \quad \text{s.t.} \quad k \leq \lambda a \}. \tag{3}$$

Note that profits depend on wealth a due to presence of the collateral constraints (2).

Lemma 1 *Profits and factor demands are linear in wealth, and there is a productivity cutoff for being active $\underline{z}$:*

$$k(a, z) = \begin{cases} \lambda a, & z \geq \underline{z} \\ 0, & z < \underline{z} \end{cases}, \quad l(a, z) = \left(\frac{1 - \alpha}{w} \right)^{1/\alpha} z k(a, z)$$

$$\Pi(a, z) = \max\{z\pi - r - \delta, 0\} \lambda a, \quad \pi = \alpha \left(\frac{1 - \alpha}{w} \right)^{(1-\alpha)/\alpha}.$$

The productivity cutoff is defined by $\underline{z}\pi = r + \delta$.

[13] That the rental rate equals the user cost of capital is irrelevant in the static model in this section. Instead, it anticipates the full dynamic model in section 2. I already introduce the notation here to avoid restating Lemma 1 with this notation there.

[14] Consider an entrepreneur with wealth a who rents k units of capital. The entrepreneur can steal a fraction $1/\lambda$ of rented capital. As a punishment, he would lose his wealth. In equilibrium, the financial intermediary will rent capital up to the point where individuals would have an incentive to steal the rented capital, implying a collateral constraint $k/\lambda \leq a$ or $k \leq \lambda a$.

(All proofs are in the Appendix.) Both the linearity and cutoff properties follow directly from the fact that individual technologies (1) display constant returns to scale in capital and labor. Maximizing out over labor in (3), profits are linear in capital, k. It follows that the optimal capital choice is at a corner: it is zero for entrepreneurs with low productivity, and the maximal amount allowed by the collateral constraints, λa, for those with high productivity. The productivity of the marginal entrepreneur is $\underline{z}$. For him, the return on one unit of capital $z\pi$ equals the cost of acquiring that unit $r + \delta$. The linearity of profits and factor demands delivers much of the tractability of my model, particularly when moving to the full dynamic setting in the next section.

1.3 Equilibrium and Aggregation

An *equilibrium* in this economy is defined in the usual way. That is, an equilibrium are prices – interest rate r and wage w – and corresponding quantities, such that (i) individuals solve (3) taking as given equilibrium prices, and (ii) the capital and labor markets clear

$$\int k(a,z)dG(a,z) = \int a dG(a,z), \tag{4}$$

$$\int l(a,z)dG(a,z) = L. \tag{5}$$

The goal of this subsection is to characterize such an equilibrium. The following object will be convenient for this task and throughout the remainder of the paper. Define the *share of wealth held by productivity type* z by

$$\omega(z) \equiv \frac{1}{K}\int_0^\infty a g(a,z) da, \tag{6}$$

where K is the aggregate capital stock.[15] As will become clear momentarily $\omega(z)$ plays the role of a density. It is therefore also useful to define the analogue of the corresponding cumulative distribution function

$$\Omega(z) \equiv \int_0^z \omega(x)dx.$$

Consider the capital market clearing condition (4). Using that $k = \lambda a$, for all active entrepreneurs ($z \geq \underline{z}$), it becomes

$$\lambda(1 - \Omega(\underline{z})) = 1.$$

This equation immediately pins down the threshold $\underline{z}$ as a function of the quality of credit markets λ. Using similar manipulations, we obtain our first main result.

[15] The definition of the aggregate capital stock is the obvious one, $K \equiv \int a dG(a,z)$. See Kiyotaki (1998) and Caselli and Gennaioli (2005) for other papers using wealth shares to characterize aggregates.

Proposition 1 *Given wealth shares $\omega(z)$, aggregate GDP is*

$$Y = ZK^{\alpha}L^{1-\alpha}, \tag{7}$$

where K and L are aggregate capital and labor and

$$Z = \left(\frac{\int_{\underline{z}}^{\infty} z\omega(z)dz}{1-\Omega(\underline{z})}\right)^{\alpha} = \mathbb{E}_{\omega}[z|z \geq \underline{z}]^{\alpha} \tag{8}$$

is measured TFP. The productivity cutoff $\underline{z}$ is defined by

$$\lambda(1-\Omega(\underline{z})) = 1. \tag{9}$$

Factor prices are

$$w = (1-\alpha)ZK^{\alpha}L^{-\alpha} \quad \text{and} \quad R = \alpha\zeta ZK^{\alpha-1}L^{1-\alpha}, \ \text{where } \zeta \equiv \frac{\underline{z}}{\mathbb{E}_{\omega}[z|z \geq \underline{z}]} \in [0,1]. \tag{10}$$

The interpretation of this result is straightforward. In terms of aggregate GDP, this economy is isomorphic to one with an *aggregate production function*, $Y = ZK^{\alpha}L^{1-\alpha}$. The sole difference is that TFP Z is endogenous and as in (8). TFP is simply a weighted average of the productivities of active entrepreneurs (those with productivity $z \geq \underline{z}$). As already discussed, (9) is the capital market clearing condition. Because $\Omega(\cdot)$ is increasing, it can be seen that the productivity threshold for being an active entrepreneur is strictly increasing in the quality of credit markets λ. This implies that, as credit markets improve, the number of active entrepreneurs decreases and their average productivity increases. Because truncated expectations such as (8) are increasing in the point of truncation, it follows that TFP is always increasing in λ.

The wage rate in (10) simply equals the aggregate marginal product of labor. This is to be expected since labor markets are frictionless and hence individual marginal products are equalized among each other and also equal the aggregate marginal product. The same is *not* true for the rental rate R. It equals the aggregate marginal product of capital $\alpha ZK^{\alpha-1}L^{1-\alpha}$ scaled by a constant ζ that is generally smaller than one. ζ only equals one if $\lambda = \infty$ so that only the most productive entrepreneur is active, $\underline{z} = \max\{z\}$, implying that the first-best is achieved.[16] In all other cases, $\zeta < 1$ so that the rental rate is lower than the aggregate marginal product of capital. In the extreme case where capital markets are completely shut down, $\lambda = 1$,

[16]For ζ to equal one, the support of z must also be finite so that $\max\{z\}$ exists.

the rental rate is zero.[17] The rental rate $R = r + \delta$ is also the return on capital faced by a hypothetical investor outside the economy. The observation that rental rates are low, therefore also speaks to the classic question of Lucas (1990): "Why doesn't capital flow from rich to poor countries?" It may be precisely capital market imperfections *within* poor countries that bring down the return on capital thereby limiting capital flows from rich countries.

An instructive special case arises when wealth and productivity are independent. When $g(a, z) = \varphi(a)\psi(z)$, the definition of wealth shares (6) implies $\omega(z) = \psi(z)$. That is, the share of wealth held by a given productivity type coincides with the mass of entrepreneurs with that same productivity. In this case TFP is simply

$$Z = \left(\frac{\int_{\underline{z}}^{\infty} z\psi(z)dz}{1 - \Psi(\underline{z})} \right)^{\alpha} = \mathbb{E}[z|z \geq \underline{z}]^{\alpha}, \tag{11}$$

that is a simple (unweighted) average of productivities. The cutoff is defined by $\lambda(1-\Psi(\underline{z})) = 1$, where $\Psi(\cdot)$ is the cumulative distribution of z.

Finally, the cutoff property is of potential interest from an empirical point of view. A recent empirical literature – as summarized by Bloom and Van Reenen (2009) – examines differences in management practices across firms and countries. One consistent finding is that there is typically a "long left tail" of badly managed firms in developing countries. That is, relatively few firms are badly managed in developed countries like the US. Conversely, there are many such firms in developing countries like India – countries that also often have underdeveloped financial markets. In my theory, better credit markets raise the productivity threshold $\underline{z}$, thereby truncating the left tail of the marginal distribution of productivity $\psi(z)$. To the extent that one can identify management practices as productivity or technology, my theory provides a potential causal link from financial frictions to the "long left tail" of badly managed firms.

1.4 A Pareto Example

The following simple example illustrates some important features of the model. Consider the case in which wealth and productivity are independent so that TFP is given by (11). Since this expression does not impose any restrictions on the productivity distribution $\psi(z)$, one can pick the distribution of one's choice and compute TFP. Hence, let productivity be distributed Pareto on $[1, \infty)$, that is $\Psi(z) = 1 - z^{-\eta}, \eta > 1$. The parameter η is an inverse measure of the

[17] This also requires that $\min\{z\} = 0$ (full support). Also note that for intermediate values, λ has an ambiguous effect on ζ and hence on the rental rate, R. The sign of $\partial\zeta/\partial\lambda = (\partial\zeta/\partial\underline{z})(\partial\underline{z}/\partial\lambda)$ depends on the sign of the elasticity $\partial \log \mathbb{E}[z|z \geq \underline{z}]/\partial \log \underline{z}$. For example, if the wealth shares $\omega(z)$ are Pareto (see section 1.4), then $\partial\zeta/\partial\underline{z} = 0$ and the rental rate, R, does not depend on λ. For other forms of the wealth shares, the effect may be either positive or negative.

thickness of the tail of the distribution (a measure of the variance). Under this assumption, the productivity cutoff is simply $\underline{z} = \lambda^{1/\eta}$, and TFP is therefore

$$Z = \left(\frac{\eta}{\eta - 1} \lambda^{1/\eta} \right)^{\alpha}. \tag{12}$$

As already argued, TFP is strictly increasing in λ. More interesting is how TFP depends on the the productivity distribution $\psi(z)$, particularly the tail parameter η. Note that the elasticity of TFP with respect to the quality of credit markets is

$$\frac{\partial \log Z}{\partial \log \lambda} = \frac{\alpha}{\eta}.$$

Figure 1 plots TFP against the parameter measuring the development of credit markets λ for different values of the tail parameter η. TFP for $\lambda = 10$ is normalized to unity for sake

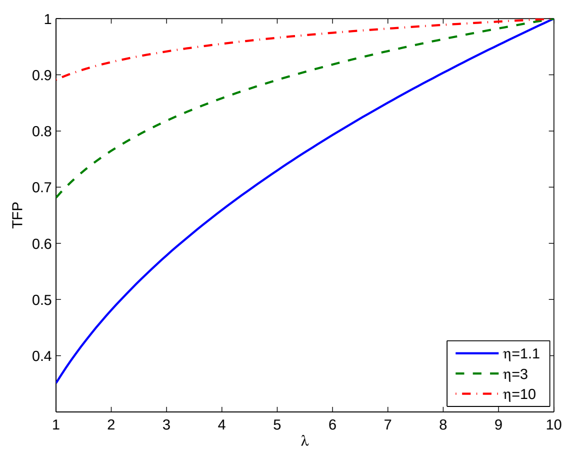

Figure 1: Total Factor Productivity

Note: TFP (12) relative to TFP for $\lambda = 10$. TFP losses are larger, the fatter is the tail of the productivity distribution (the smaller is η). The capital share is given by $\alpha = 0.3$.

of comparison.[18] It can be seen from the Figure and the elasticity of Z with respect to λ that productivity losses from financial frictions are largest if the distribution of idiosyncratic productivities has a thick tail. This is intuitive. A thick tail implies that there are some extremely high-productivity entrepreneurs and that it is highly desirable from the point of view of society to direct capital towards them. With underdeveloped financial markets, this is however not possible so that productivity losses are large. While this example is intended to highlight the qualitative rather than quantitative implications of the model, I remark that the productivity loss from shutting down credit markets, $\lambda = 1$, relative to having good credit markets, $\lambda = 10$, varies considerably. It may be anywhere between ten and more than sixty

[18] As explained in the empirical part, section 3.4, a value of $\lambda = 10$ gives rise to an external finance to GDP ratio roughly equal to that of the US.

percent depending on the value of η.[19]

The Pareto example also delivers a simple expression for the rental rate R. Since

$$\zeta = \frac{\underline{z}}{\mathbb{E}[z|z \geq \underline{z}]} = 1 - \frac{1}{\eta} < 1,$$

we have that

$$R = \alpha \left(1 - \frac{1}{\eta}\right) Z K^{\alpha-1} L^{1-\alpha} < \alpha Z K^{\alpha-1} L^{1-\alpha}.$$

Note again the presence of the tail parameter η. A thicker tail of the productivity distribution (lower η) lowers the rental rate. This is intuitive because a low rental rate is a symptom of badly working credit markets, as discussed above.

2 Full Dynamic Model

The simple static model of entrepreneurship and collateral constraints presented in the previous section took as given the joint distribution $g(a,z)$. In this section I endogenize this joint distribution: for any given entrepreneur, a and z are jointly determined by a stochastic process for z and an optimally chosen time path for wealth a. I have argued in the preceding section that knowledge of the *share of wealth held by a given productivity type*, $\omega(z)$ was sufficient for assessing TFP losses from financial frictions. Together with the aggregate capital stock K and aggregate labor L, TFP then determined aggregate GDP. The same will be true here. Endogenizing the joint distribution of productivity and wealth $g(a,z)$ is therefore equivalent to endogenizing wealth shares and aggregate capital, $\omega(z)$ and K.

2.1 Preferences and Technology

Time is continuous. As in the static model, there is a continuum of entrepreneurs that are indexed by their productivity z and their wealth a. Productivity z follows some Markov process (the exact process is irrelevant for now). I assume a law of large numbers so the share of entrepreneurs experiencing any particular sequence of shocks is deterministic. At each point in time t, the state of the economy is then the joint distribution $g_t(a,z)$. Entrepreneurs have

[19] A natural question is then: what is a reasonable value for η? A large literature has observed that firm sizes – usually measured by employment – follow a Pareto distribution (see for example Simon and Bonini, 1958; Luttmer, 2007; Gabaix, 2009). This observation can be used as a clue. Luttmer finds that the tail of the employment distribution has a tail parameter $\eta_l = 1.06$. In my model, employment is proportional to the product of wealth and productivity, $l \propto az$ (see Lemma 1). Therefore its distribution can only be obtained with knowledge of the wealth distribution. Consider, however the case where wealth is distributed Pareto with parameter η_a. Gabaix (2009) shows that then l is also distributed Pareto with parameter $\eta_l = \min\{\eta, \eta_a\}$. Gabaix (p.275) also states that $\eta_a \approx 1.5$. Therefore, $\eta = \eta_l = 1.06$ would also be value for the tail parameter of the productivity distribution, implying large TFP losses from financial frictions.

preferences

$$\mathbb{E}_0 \int_0^\infty e^{-\rho t} \log c(t) dt, \tag{13}$$

and as in the static model above they own private firms operating constant returns technologies (1). Capital depreciates at the rate δ. Again as above, there is a measure L of workers. Workers have the same preferences as (13) with the exception that they face no uncertainty so the expectation is redundant.

2.2 Budgets

Denote by $a(t)$ an entrepreneur's wealth and by $r(t)$ and $w(t)$ the (endogenous) interest and wage rates. Entrepreneurs can rent capital $k(t)$ in a rental market at a rental rate $R(t) = r(t)+\delta$. Their wealth then evolves according to

$$\dot{a} = f(z,k,l) - wl - (r+\delta)k + ra - c. \tag{14}$$

Savings $\dot{a}$ equal profits – output minus payments to labor and capital – plus interest income minus consumption. The setup with a rental market is chosen solely for simplicity. I show in Appendix B that it is equivalent to a setup in which entrepreneurs own and accumulate capital k and can trade in a risk-free bond. Entrepreneurs face the same collateral constraints as in the static model, $k \leq \lambda a$. Again, this formulation of credit market imperfections, is analytically most tractable and captures the idea that the capital available to an entrepreneur is limited by his wealth.[20] I assume that workers cannot save so that they are in effect hand-to-mouth workers who immediately consume their earnings. Workers can therefore be omitted from the remainder of the analysis.[21]

[20] The constraint can still be motivated in terms of a limited enforcement problem as in footnote 14. See Banerjee and Newman (2003) and Buera and Shin (2010) for a similar motivation of the same form of constraint. Note, however, that the constraint is essentially static because it rules out optimal long term contracts (as in Kehoe and Levine, 2001, for example). On the other hand, as Banerjee and Newman put it "there is no reason to believe that more complex contracts will eliminate the imperfection altogether, nor diminish the importance of current wealth in limiting investment."

[21] A more natural assumption can be made when one is only interested in the economy's long-run equilibrium. Allow workers to save so that their wealth evolves as $\dot{a} = w + ra - c$, but impose that they cannot hold negative wealth, $a(t) \geq 0$ for all t. Workers then face a standard deterministic savings problem so that they decumulate wealth whenever the interest rate is smaller than the rate of time preference, $r < \rho$. It turns out that the steady state equilibrium interest rate always satisfies this inequality (see corollary 1). Together with the constraint that $a(t) \geq 0$, this immediately implies that workers hold zero wealth in the long-run. Therefore, even if I allowed workers to save, in the long-run they would endogenously choose to be hand-to-mouth workers.

2.3 Individual Behavior

Entrepreneurs maximize the present discounted value of utility from consumption (13) subject to their budget constraints (14). Their production and savings/consumption decisions separate in a convenient way. Define the same profit function $\Pi(a, z)$ as in the static model – see (3). Then the budget constraint (14) can be rewritten as

$$\dot{a} = \Pi(a, z) + ra - c.$$

The interpretation is that entrepreneurs solve a static profit maximization problem *period by period*. They then decide to split those profits (plus interest income ra) between consumption and savings. Because the period problem of an entrepreneur is the same as the static problem in the preceding section, all results in Lemma 1 still apply. In particular, factor demands and profits are linear in wealth and there is a productivity cutoff $\underline{z}$ for being active. The profit function is

$$\Pi(a, z) = \max\{z\pi - r - \delta, 0\}\lambda a,$$

implying a law of motion for wealth that is linear in wealth

$$\dot{a} = [\lambda \max\{z\pi - r - \delta, 0\} + r]\, a - c.$$

This linearity allows me to derive a closed form solution for the optimal savings policy function.

Lemma 2 *The optimal savings policy function is linear in wealth*

$$\dot{a} = s(z)a, \quad \textit{where} \quad s(z) = \lambda \max\{z\pi - r - \delta, 0\} + r - \rho \tag{15}$$

is the savings rate of productivity type z.

Importantly, savings are characterized by a constant savings rate out of wealth. This is a direct consequence of the assumption of log utility combined with the linearity of profits.

2.4 Equilibrium and Aggregate Dynamics

An *equilibrium* in the dynamic setting is the exact analogue of an equilibrium in the static setting. That is, such an equilibrium are time paths for prices $r(t), w(t), t \geq 0$ and corresponding quantities, such that (i) entrepreneurs maximize (13) subject to (14) taking as given equilibrium prices, and (ii) the capital and labor markets clear at each point in time

$$\int k_t(a,z)dG_t(a,z) = \int a dG_t(a,z), \tag{16}$$

$$\int l_t(a,z)dG_t(a,z) = L. \tag{17}$$

Aggregation in the dynamic model is very similar to aggregation in the static model. Define the analogous (time-varying) wealth share of productivity type z:

$$\omega(z,t) \equiv \frac{1}{K(t)} \int_0^\infty a g_t(a,z)da, \tag{18}$$

where $g_t(a,z)$ is the joint distribution of productivity and wealth and $K(t)$ is the aggregate capital stock.

We can derive the law of motion for aggregate capital by integrating (15) over all entrepreneurs. Using the definition of the wealth shares (18), we get

$$\begin{aligned} \dot{K}(t) &= \int_0^\infty s(z,t)\omega(z,t)dzK(t) \\ &= \int_0^\infty \left[\lambda \max\{z\pi(t) - r(t) - \delta, 0\} + r(t) - \rho\right]\omega(z,t)dzK(t). \end{aligned} \tag{19}$$

By further manipulating this expression, we get the following extension of Proposition 1 to a dynamic setting.

Proposition 2 *Given a time path for wealth shares $\omega(z,t), t \geq 0$, aggregate quantities satisfy*

$$Y = ZK^\alpha L^{1-\alpha}, \tag{20}$$

$$\dot{K} = \alpha ZK^\alpha L^{1-\alpha} - (\rho + \delta)K, \tag{21}$$

where K and L are aggregate capital and labor and

$$Z(t) = \left(\frac{\int_{\underline{z}}^\infty z\omega(z,t)dz}{1 - \Omega(\underline{z},t)}\right)^\alpha = \mathbb{E}_{\omega,t}[z|z \geq \underline{z}]^\alpha \tag{22}$$

is measured TFP. The productivity cutoff $\underline{z}$ is defined by

$$\lambda(1 - \Omega(\underline{z},t)) = 1. \tag{23}$$

Factor prices are

$$w = (1-\alpha)ZK^{\alpha}L^{-\alpha} \quad \text{and} \quad r = \alpha\zeta ZK^{\alpha-1}L^{1-\alpha} - \delta, \;\; \text{where } \zeta \equiv \frac{\underline{z}}{\mathbb{E}_{\omega,t}[z|z \geq \underline{z}]} \in [0,1]. \quad (24)$$

The expression for GDP (20) is the same as in the static setting. Condition (21) gives a simple law of motion for the aggregate savings. The key to this aggregation result is that individual savings policy rules are linear as shown in Lemma 1.[22] Again, TFP in (22) is endogenous and given by a weighted average of productivities above a threshold $\underline{z}$ that depends on the quality of credit markets, λ, and is defined by (23). The law of motion (21) deserves special treatment. It can be written as

$$\dot{K} \equiv \hat{s}Y - \hat{\delta}K, \quad \text{where} \quad \hat{s} \equiv \alpha \text{ and } \hat{\delta} \equiv \rho + \delta,$$

are constant savings and depreciation rates. This is the same law of motion as in the classic paper by Solow (1956).[23] What is surprising about this observation is that the starting point of this paper – heterogenous entrepreneurs that are subject to borrowing constraints – is very far from an aggregate growth model such as Solow's. One twist differentiates the model from an aggregate growth model: TFP $Z(t)$ is endogenous. It is determined by the quality of credit markets and the evolution of the distribution of wealth as summarized by the wealth shares $\omega(z,t)$. I show in section 2.6 below that, given a stochastic process for idiosyncratic productivity z, one can construct a time path for the wealth shares $\omega(z,t)$. In turn, a time path for TFP $Z(t)$ is implied. But given this evolution of TFP – says proposition 2 – aggregate capital and output behave as in an aggregate growth model. One immediate implication of interest is that financial frictions as measured by the parameter λ have no *direct* effects on aggregate savings; they only affect savings indirectly through TFP.

A second observation of interest concerns the dynamic behavior of factor prices, particularly the interest rate $r(t)$. The expressions (24) are exactly the same as in the static model. Again, the wage rate equals the aggregate marginal product of labor. In contrast, financial frictions break the link between the interest rate and the aggregate marginal product of capital (see the discussion in section 1.3). Frictions generally lower the interest rate. In a dynamic setting this is of interest for several reasons, one of which I wish to highlight here. King and Rebelo (1993) argue that when one tries to explain sustained growth by transitional dynamics in representa-

[22]The same "trick" is used by Angeletos (2007) and Kiyotaki and Moore (2008).

[23]Alternatively, (21) may be written as $\dot{K} = \tilde{s}(K/Y)Y - \delta K$, where the savings rate $\tilde{s}(K/Y) = \alpha - \rho K/Y$ depends on the capital-output ratio (it is higher, the lower the capital-output ratio). In this way the discount rate ρ (a preference parameter) enters the savings rate instead of the the depreciation rate.

tive agent models like the neoclassical growth model, one generates extremely counterfactual implications for the time path of the interest rate. According to their calculations for example, if the neoclassical growth model were to explain the postwar growth experience of Japan, the interest rate in 1950 should have been around 500 percent. As can be seen from (24), it is theoretically possible that both the capital stock and the interest rate approach the steady state *from below*, offering a way out of the problem raised by King and Rebelo (1993).

2.5 Steady State Equilibrium

A *steady state equilibrium* is a competitive equilibrium satisfying[24]

$$\dot{K}(t) = 0, \quad \omega(z,t) = \omega(z), \quad r(t) = r, \quad w(t) = w \quad \text{for all } t. \tag{25}$$

Imposing these restrictions in Proposition 2 yields the following immediate corollary.

Corollary 1 *Given stationary wealth shares $\omega(z)$, aggregate steady state quantities solve*

$$Y = ZK^{\alpha}L^{1-\alpha} \tag{26}$$

$$\alpha Z K^{\alpha-1} L^{1-\alpha} = \rho + \delta, \tag{27}$$

where K and L are aggregate capital and labor and

$$Z = \left(\frac{\int_{\underline{z}}^{\infty} z\omega(z)dz}{1 - \Omega(\underline{z})} \right)^{\alpha} = \mathbb{E}_{\omega}[z|z \geq \underline{z}]^{\alpha}$$

is measured TFP. The productivity cutoff $\underline{z}$ is defined by $\lambda(1 - \Omega(\underline{z})) = 1$. Factor prices are $w = (1-\alpha)ZK^{\alpha}L^{-\alpha}$ and $r = \alpha\zeta ZK^{\alpha-1}L^{1-\alpha} - \delta = \zeta(\rho+\delta) - \delta$, where $\zeta \equiv \underline{z}/\mathbb{E}_{\omega}[z|z \geq \underline{z}] \in [0,1]$.

Most expressions have exactly the same interpretation as in the dynamic equilibrium above. (27) says that the aggregate steady state capital stock in the economy solves a condition that

[24] Note that, although there is a steady state for aggregates, there is no steady state for the joint distribution of productivity and wealth $g_t(a,z)$. The same phenomenon occurs in the papers by Krebs (2003) and Angeletos (2007). The reason is that the growth rate of wealth $s(z)$ is stochastic and does not depend on wealth itself (the log of wealth therefore follows something resembling a random walk). However, *wealth shares* $\omega(z,t)$ still allow for a stationary measure $\omega(z)$. Stationary wealth shares are then defined by

$$\omega(z) \equiv \frac{1}{K} \int_0^{\infty} a g_t(a,z) da,$$

where the reader should note the t subscript on the joint distribution but not on the wealth shares. See the discussion in section 2.6.

is precisely the same as in a standard neoclassical, namely that the aggregate marginal product of capital equals the sum of the rate of time preference and the depreciation rate.[25]

Condition (27) further implies that the capital-output ratio in this economy is given by

$$\frac{K}{Y} = \frac{\alpha}{\rho + \delta}, \tag{28}$$

which is again the same expression as in a standard neoclassical growth model. The capital-output ratio does not depend on the quality of credit markets, λ, except indirectly through TFP. This is consistent with the finding in the development accounting literature that capital-output ratios are relatively similar across countries (Hall and Jones, 1999).

2.6 The Evolution of Wealth Shares

The description of equilibrium so far has taken as given the evolution of wealth shares $\omega(z,t)$. The statements in Proposition 2 and Corollary 1 were of the form: given a time path for $\omega(z,t), t \geq 0$, statement [...] is true. This section fills in for the missing piece and explains how to characterize the evolution of wealth shares.

Note first that the evolution of wealth shares $\omega(z,t)$ and hence TFP losses from financial frictions depend crucially on the assumptions placed on the *stochastic process* for idiosyncratic productivity z. Consider the extreme example where each entrepreneur's productivity is *fixed* $z(t) = z$ for all t. In this case, financial frictions will have *no effect* on aggregate TFP asymptotically. To see this, consider the optimal savings policy function, $\dot{a}(t) = s(z)a(t)$ (see Lemma 2), and note that the savings rate $s(z)$ is increasing in productivity z. Since productivity is fixed over time, the entrepreneurs with the highest productivity $\max\{z\}$ will always accumulate at a faster pace than others. In the long run (as $t \to \infty$), he will therefore hold all the wealth in the economy, implying that stationary wealth shares are

$$\omega(z) = \begin{cases} 1, & z = \max\{z\} \\ 0, & z < \max\{z\}. \end{cases}$$

It follows immediately that TFP is $Z = \max\{z\}^\alpha$. The equilibrium is first-best regardless of the quality of credit markets, λ. The interpretation of this result is that, asymptotically, self-financing completely undoes all capital misallocation caused by financial frictions.[26]

[25] It is interesting to note that, in steady state, aggregates are observationally equivalent to those in a neoclassical growth model whereas, during the transition, they are not and perhaps even closer to a Solow model (see 2.4). This comes from the combination of constant returns as the individual level (1), and decreasing returns in the aggregate due to a constant labor force L.

[26] See Banerjee and Moll (2009) for a very similar result.

If productivity z follows a non-degenerate stochastic process, this is – in general – no longer true. However, characterizing the evolution of wealth shares is harder. To make some headway for this case, I assume that productivity, z, follows a diffusion which is simply the continuous time version of a Markov process:[27]

$$dz = \mu(z)dt + \sigma(z)dW. \tag{29}$$

$\mu(z)$ is called the drift term and $\sigma(z)$ the diffusion term. In addition, I assume that this diffusion allows for a stationary distribution. I would like to note here that other stochastic processes are also possible. For example, Buera and Moll (2010) analyze a similar model under the assumption that z follows a Poisson process. It should also be feasible to analyze the more general class of Lévy processes which comprise both jump processes such as the Poisson process and diffusion processes such as (29).

The following Proposition is the main tool for characterizing the evolution of wealth shares $\omega(z,t)$.

Proposition 3 *The wealth shares $\omega(z,t)$ obey the second order partial differential equation*

$$\frac{\partial \omega(z,t)}{\partial t} = \left[s(z,t) - \frac{\dot{K}(t)}{K(t)}\right]\omega(z,t) - \frac{\partial}{\partial z}[\mu(z)\omega(z,t)] + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left[\sigma^2(z)\omega(z,t)\right]. \tag{30}$$

The wealth shares must also be non-negative and bounded everywhere, integrate to one for all t $\int_0^\infty \omega(z,t)dz = 1$, and satisfy the initial condition $\omega(0,t) = \omega_0(z)$ for all z.

The stationary *wealth shares $\omega(z)$ obey the second order ordinary differential equation*

$$0 = s(z)\omega(z) - \frac{d}{dz}[\mu(z)\omega(z)] + \frac{1}{2}\frac{d^2}{dz^2}\left[\sigma^2(z)\omega(z)\right]. \tag{31}$$

The stationary wealth shares must also be non-negative and bounded everywhere, and integrate to one, $\int_0^\infty \omega(z)dz = 1$.[28]

This PDE and the related ODE are mathematically similar to the Kolmogorov forward equation used to keep track of cross-sectional distributions of diffusion processes. There is unfortunately

[27] Readers who are unfamiliar with stochastic processes in continuous time may want to read the simple discrete time setup with iid shocks in the online Appendix at `http://www.princeton.edu/~moll/research.htm`. The present setup in continuous time allows me to derive more general results, particularly with regard to the persistence of shocks which is the central theme in this paper.

[28] I here leave open the question of precise boundary conditions. This is because the process (29) is very general so that precise boundary conditions are hard to come by without precise restrictions on the drift and diffusion coefficients. In the example analyzed below boundary conditions are not an issue and wealth shares are always determined uniquely. There, the ODE (31) can be solved analytically. The solution has two branches one of which can be set to zero because it explodes as z tends to infinity, replacing the missing boundary condition.

no straightforward intuition for these equations so that readers who are unfamiliar with the related mathematics will have to take them at face value.[29] Solving for the wealth shares also requires solving for equilibrium prices and aggregate quantities which satisfy the capital and labor market clearing conditions (16) and (17). See Appendix C.1 for a complete statement of all equilibrium conditions.

For the remainder of this paper I consider only steady state equilibria as in Corollary 1. See Appendix C.2, for a general algorithm for computing steady state equilibria. One feature of the model deserves further treatment. The *stationary* wealth shares in Corollary 1 and proposition 3 are defined by

$$\omega(z) \equiv \frac{1}{K}\int_0^\infty a g_t(a,z)da. \tag{32}$$

Note that the joint distribution of productivity and wealth $g_t(a,z)$ carries a t subscript. The reason is that, while aggregates are constant in a steady state equilibrium, there is no steady state for the joint distribution of productivity and wealth $g_t(a,z)$. The same phenomenon occurs in the papers by Krebs (2003) and Angeletos (2007). To understand this, note that the growth rate of wealth, that is the savings rate $s(z)$, depends on (stochastic) productivity z but not on wealth itself. An entrepreneur who starts off with twice the wealth than another, but experiences the same sequence of shocks as the other, will always be twice as rich as the other. That is, there is no feature in the model that would pull their wealth together. Extending this logic to the cross-section, the wealth distribution always "fans out" over time so that it does not admit a stationary distribution. If the model were set up in discrete time, the log of wealth would follow a random walk which is the prototypical example of a process without a stationary distribution.

However, and despite the fact that the joint distribution $g_t(a,z)$ is non-stationary, the *wealth shares* $\omega(z,t)$ still admit a stationary measure $\omega(z)$ defined as in (32). This allows me to completely sidestep the nonexistence of a stationary wealth distribution.[30] The linearity of the optimal savings policy function (Lemma 2) is again crucial for this result.

2.7 Closed Form Solution for $\lambda = 1$

The main purpose of this section is to illustrate the role of the *autocorrelation* of productivity shocks for capital misallocation and implied TFP losses. To do so, I specialize to the extreme case of no capital markets, $\lambda = 1$. The case $\lambda = 1$ is restrictive but carries all intuition for

[29] For readers who are familiar with it: If the function $s(z)$ were identically zero, these equations would coincide with the forward equation for the marginal distribution of productivities $\psi(z,t)$. The term $s(z)$ functions like a Poisson killing rate (however note that $s(z)$ generally takes both positive and negative values).

[30] However, see section 6.1 in Angeletos (2007) for potential extensions introducing a stationary wealth distribution.

the more general case $\lambda \geq 1$. The latter is analyzed numerically in the next section. By specializing the stochastic process (29), I can solve the ODE for stationary wealth shares $\omega(z)$, (31), in closed form. All aggregate variables in the model can then be obtained in closed form as well. The example I present here is for one particular stochastic process (29). Other processes can yield closed form solutions as well, a fact that I underline by presenting a second example in Appendix D.[31]

That being said, the following stochastic process is convenient:

$$\frac{dz}{z} = \left[\nu\left(1-z\right) + \frac{\sigma^2}{2}\right] dt + \sigma dW, \tag{33}$$

where ν and σ are positive. This is just the special case of (29) with a drift term $\mu(z) = z\left[\nu\left(1-z\right) + \sigma^2/2\right]$ and a diffusion term $\sigma(z) = \sigma z$. The term $\sigma^2/2$ in the drift makes algebra easier below.[32] Importantly, this process is mean-reverting and therefore allows for a stationary distribution. The speed of mean reversion is determined by the parameter ν. The stationary distribution is given by

$$\psi(z) \propto e^{-\beta z} z^{\beta-1}, \quad \beta = \frac{2\nu}{\sigma^2}. \tag{34}$$

This is the formula for a Gamma distribution with both parameters equal to β as stated above. The mean and variance are

$$\mathbb{E}[z] = 1, \quad \mathbb{V}[z] = \frac{\sigma^2}{2\nu} = \frac{1}{\beta}. \tag{35}$$

I impose the parameter restriction $\sigma^2 < 2\nu$. As can be seen from (34), this assumption ensures that the stationary distribution has zero density at $z = 0$.

This section is chiefly concerned with the persistence of productivity shocks. Wong (1964) shows that the autocorrelation of z between two dates t and $t+s$, $s \geq 0$ is given by

$$Corr[z(t), z(t+s)] = e^{-\nu s} \in (0, 1]. \tag{36}$$

Two intuitive observations can be made. First, the autocorrelation is smaller the bigger is the distance in time between the two observations, s. Second, the autocorrelation is high if the speed of mean reversion ν is low. Taking the limit as $\nu \to \infty$, we can obtain the case $Corr[z(t), z(t+s)] = 0$. This limit therefore corresponds to the case where productivity shocks

[31] I there analyze a process known as a Feller square root process. These processes are examples of a class of stationary diffusion processes whose stationary distributions are of the so-called Pearson class. Wong (1964) presents other examples of the same class, some of which should also work here.

[32] It is a correction for the fact that (33) is a geometric process, i.e. it has *relative* increments dz/z. Applying Itô's Lemma, $\hat{z} \equiv \log z$ evolves according to $d\hat{z} = \nu(1 - \exp(\hat{z})) + \sigma dW$ which does not involve the term $\sigma^2/2$. See the discussion in Dixit (1993) of what he terms the "Jensen-Itô effect".

are iid over time.[33]

Under the specific functional form (33), one can solve the ODE for wealth shares (31) using a guess-and-verify strategy.

Proposition 4 *Consider an economy with no credit markets $\lambda = 1$, and where productivity follows the stochastic process* (33)*. Then the stationary wealth shares are given by*

$$\omega(z) \propto e^{-\beta z} z^{\gamma-1}, \quad \beta = \frac{2\nu}{\sigma^2}, \quad \gamma = \frac{\beta}{2}\left(1 + \sqrt{1 + \frac{4}{\beta\nu}(\rho+\delta)}\right). \tag{37}$$

This expression can be usefully contrasted with the expression for the stationary productivity distribution (34). I conduct the following experiment: vary the parameter governing the autocorrelation ν while holding constant the distribution $\psi(z)$ and particularly its variance $1/\beta$.[34] Figure 2 plots the wealth shares $\omega(z)$ relative to the distribution $\psi(z)$ for different values of $Corr[z(t), z(t+1)] = \exp(-\nu)$. Two observations can be made: First, the wealth shares $\omega(z)$ generally place more mass on higher productivity types (in the sense of first order stochastic dominance). This is because for any positive autocorrelation, there is some scope for self-financing so that higher productivity types accumulate more wealth. Second, wealth is more concentrated with higher productivity types, the higher is the autocorrelation of productivity shocks. To restate the same point in a slightly different manner, note that

$$\omega(z) \to \psi(z) \quad \text{as} \quad \nu \to \infty \quad (\text{so that} \quad Corr \to 0).$$

Taking the limit as the autocorrelation goes to zero implies that we are in an environment where shocks are iid over time. In this case, wealth and productivity will be independent $g_t(a, z) = \varphi_t(a)\psi(z)$ because iid shocks imply that productivity shocks are unpredictable at the time when savings decisions are made. It follows directly from the definition of the stationary wealth shares in (32) that $\omega(z) = \psi(z)$. As we increase the autocorrelation of productivity shocks above zero, self-financing becomes more and more feasible and wealth becomes more and more concentrated among high productivity types.

[33]A continuous time setup is of course not very amenable to iid shocks. See the simpler discrete time setup with iid shocks in the online Appendix at `http://www.princeton.edu/~moll/research.htm`.

[34]The following example should clarify: suppose instead that (the log of) productivity follows a discrete time AR(1) process, $\log z_t = \rho \log z_{t-1} + \sigma\varepsilon_t$. The stationary distribution of this process is a normal distribution with mean zero, and variance $\mathbb{V} = \frac{\sigma^2}{1-\rho^2}$. The analogous experiment is then to vary ρ while holding constant $\mathbb{V} = \frac{\sigma^2}{1-\rho^2}$.

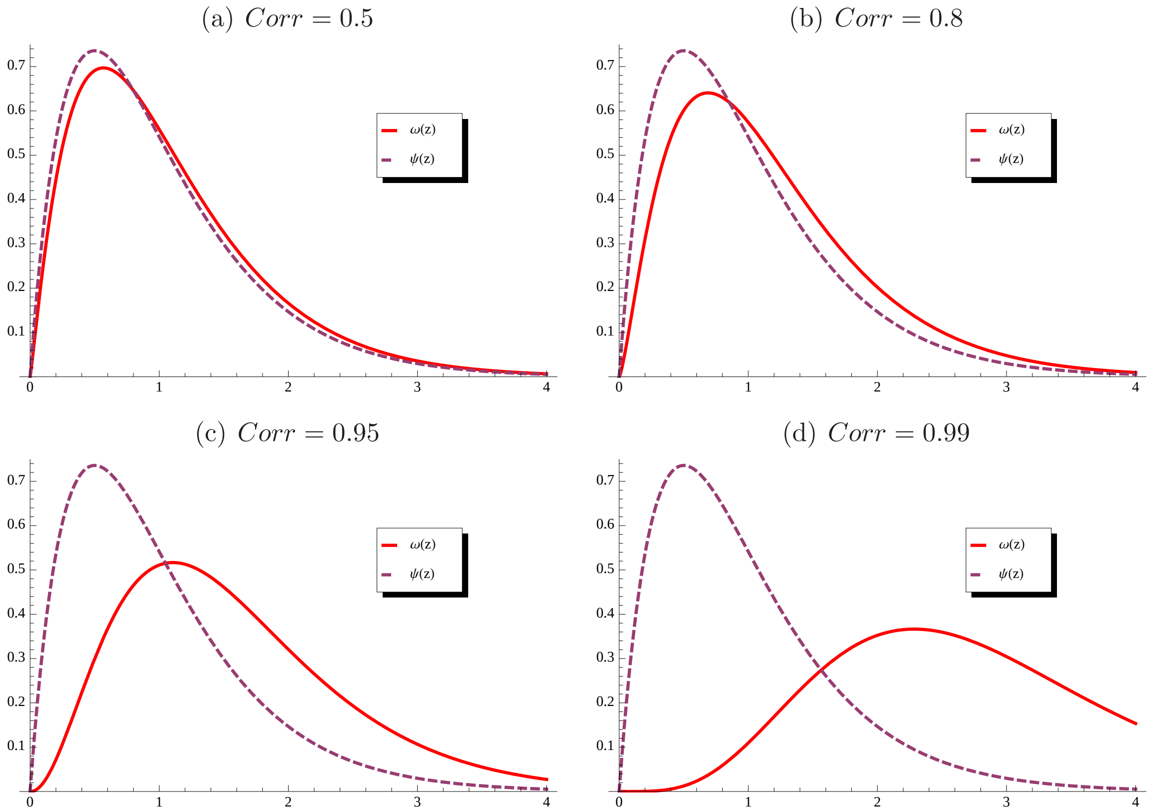

Figure 2: Wealth Shares and Autocorrelation

Note: The dashed lines are the productivity distribution $\psi(z)$ from (34). The solid lines are the wealth shares $\omega(z)$ from (37). As autocorrelation increases (equivalently ν decreases), wealth becomes more concentrated with high productivity entrepreneurs.

It is also instructive to examine the conditional expectation

$$\frac{\mathbb{E}[a|z]}{K} = \frac{\omega(z)}{\psi(z)} \propto z^{\hat{\gamma}}, \quad \hat{\gamma} = \frac{\beta}{2}\left(-1 + \sqrt{1 + \frac{4}{\beta\nu}(\rho+\delta)}\right). \tag{38}$$

This conditional expectation is the *per-capita* wealth share of a given productivity type.[35] Figure 3 plots this per-capita wealth share, again for different values of the autocorrelation of productivity shocks. A higher autocorrelation again implies that high productivity types hold more wealth.

My goal is to assess TFP losses from capital misallocation. Under the specific functional form for the productivity process (33), one can obtain an expression for aggregate TFP. Using that the wealth shares are Gamma and that therefore TFP is $Z = \mathbb{E}_\omega[z]^\alpha = (\gamma/\beta)^\alpha$, we obtain

[35]The first equality follows from

$$\frac{\mathbb{E}[a|z]}{K} = \frac{1}{K}\frac{\int_0^\infty a g_t(a,z)da}{\psi(z)} = \frac{\omega(z)}{\psi(z)}.$$

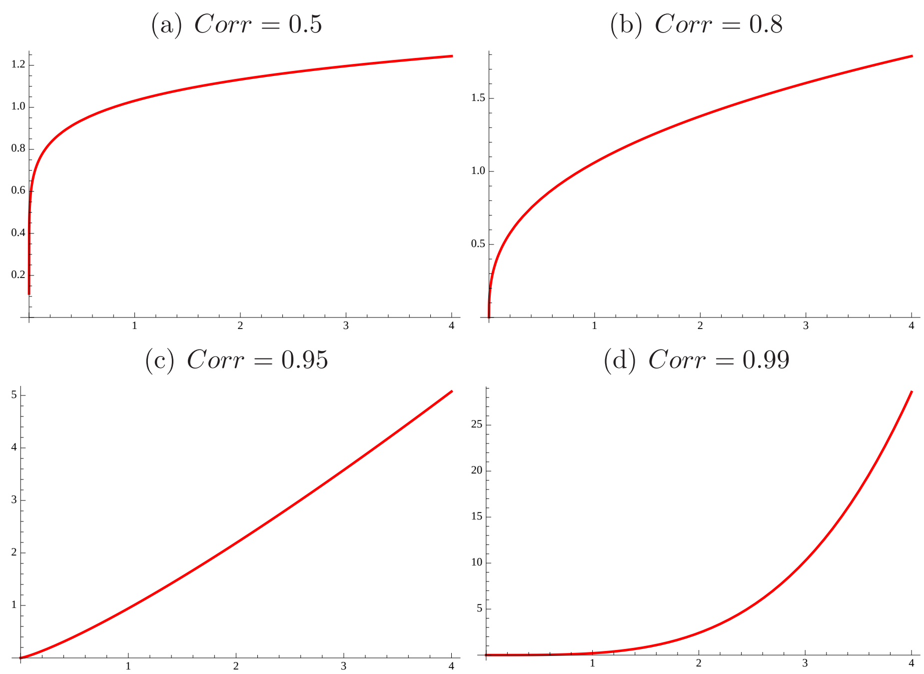

Figure 3: Higher Autocorrelation Implies Better Wealth Allocation

Note: Per-capita wealth $\mathbb{E}[a|z]/K$ from (38). As in Figure 2, wealth is more concentrated with high productivity entrepreneurs if autocorrelation is high.

the expression

$$Z = \left(\frac{1}{2} + \frac{1}{2}\sqrt{1 + \frac{4}{\beta\nu}(\rho + \delta)}\right)^{\alpha}. \tag{39}$$

Figure 4 shows how TFP changes with autocorrelation $Corr[z(t), z(t+1)] = \exp(-\nu)$. As

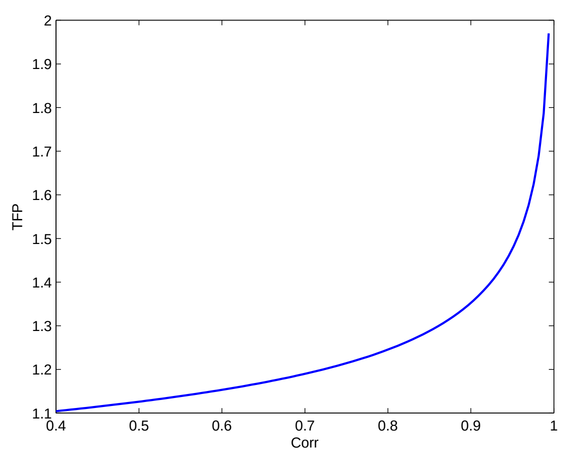

Figure 4: TFP and Autocorrelation.

Note: TFP as in (39) plotted against $Corr = \exp(-\nu)$. Other parameters are $\alpha = 0.3, \beta = 1.5, \rho = \delta = 0.05$.

expected, TFP is higher the more correlated are productivity shocks. This follows immediately from the fact that wealth is more concentrated among high productivity types so that there is less capital misallocation. Two limiting cases are also of interest: first

$$Z \to \mathbb{E}[z]^{\alpha} = 1 \quad \text{as} \quad \nu \to \infty \quad (\text{so that} \quad Corr \to 0).$$

As already discussed, this limit corresponds to the case where productivity shocks are iid over time, also implying that $\omega(z) = \psi(z)$. TFP is then given by the (unweighted) average productivity which here equals unity. Second,

$$Z \to \max\{z\}^{\alpha} = \infty \quad \text{as} \quad \nu \to 0 \quad (\text{so that} \quad Corr \to 1).$$

That is, as autocorrelation tends to one, all wealth is held by the highest productivity type (here $z = \infty$) so that TFP is first-best. A final observation of interest from Figure 4 is that TFP is convex as a function of the autocorrelation of shocks. TFP is not sensitive to the autocorrelation of shocks for low values of the latter whereas it is highly sensitive for high values. This is noteworthy because – as I will argue below – shocks are quite persistent in the data so that the sensitive region is the empirically relevant one.

The main purpose of this section has been to illustrate the role of the autocorrelation of productivity shocks for capital misallocation and hence for TFP losses from financial frictions. Using a closed-form example I have demonstrated that even with no capital markets $\lambda = 1$ the first-best capital allocation is attainable if productivity shocks are sufficiently correlated over time. Conversely, TFP losses can be large if shocks are iid over time or close to that case.

2.8 Numerical Results for $\lambda > 1$

The preceding section presented some closed form results for the case of no capital markets, $\lambda = 1$. The same is unfortunately not possible for the more general case, $\lambda > 1$. Here, I briefly present some numerical results for this more general case to show that the intuition gained in the preceding section carries over in its entirety. The numerical algorithm for solving the steady state equilibrium with a general diffusion and any quality of credit markets, $\lambda \geq 1$, is described in Appendix C.2. Figure 5 graphs TFP against the parameter capturing the quality of credit markets, λ. The left panel displays the level of TFP and the right panel displays TFP losses relative to the case $\lambda = 10$. TFP levels are higher the more correlated are productivity shocks. More importantly, TFP losses relative to good credit markets ($\lambda = 10$) are smaller the higher is autocorrelation. Both findings conform exactly with the intuition built up in the closed-form example of the preceding section. Only relatively persistent shocks allow for wealth accumulation and hence for self-financing to function as a substitute to credit access.

(a) TFP Levels　　　　(b) TFP relative to $\lambda = 10$

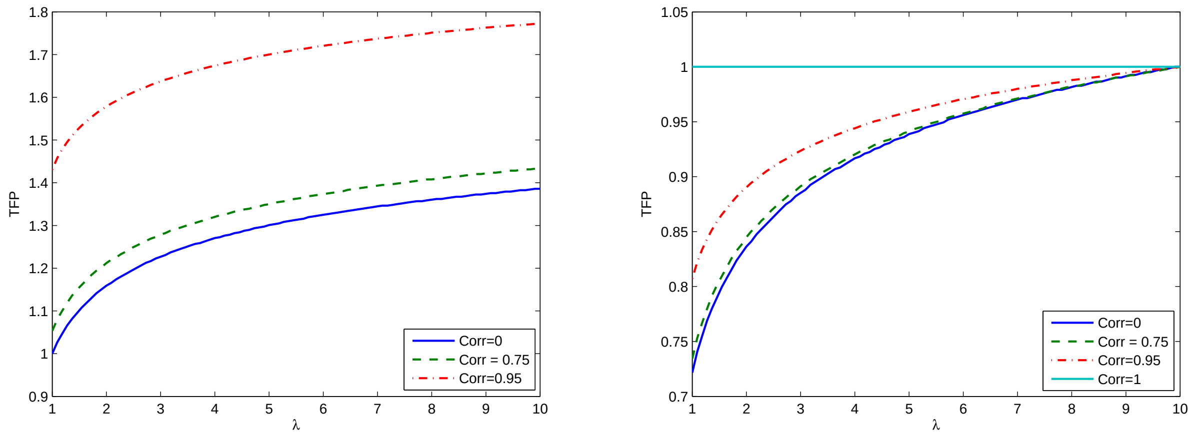

Figure 5: TFP and Autocorrelation.

Note: The left panel plots levels. The right panel TFP losses relative to $\lambda = 10$. For higher autocorrelation, both the level of TFP is higher and losses relative to $\lambda = 10$ are lower. Again, note the sensitivity in the range $Corr = 0.75$ to $Corr = 1$. Parameters are $\alpha = 1/3$, $\rho = \delta = 0.05$, $\beta = 2\nu/\sigma^2 = 1.75$.

Consequently, wealth shares for the case $\lambda > 1$ look exactly like those for $\lambda = 1$ in Figure 2. That is, wealth becomes more concentrated with high productivity entrepreneurs the higher is the autocorrelation of shocks.

3 Empirical Evidence

While the main contribution of this paper is a theoretical one, I wish to argue in this section that despite its simplicity the model is, in fact, quite capable of capturing some salient features of the data. I first describe the datasets that I study and briefly summarize my empirical strategy that I then follow for the remainder of this section.

3.1 Description of Datasets

I use four datasets in my analysis. Two cover micro data and two macro data. I describe all in turn.

Micro data I use manufacturing censuses from two emerging market economies. The first data have been collected by Chile's Instituto Nacional de Estadística (INE) and the second are from Colombia's Departamento Administrativo Nacional de Estadística (DANE). The Chilean data cover the years 1987-1996 and the Colombian data are for 1981-1989. The specific version of the Chilean data is from Greenstreet (2007). The Colombian data are from Tybout and Roberts (1996). These data have been used in numerous other studies, mainly in the industrial organization literature studying the evolution of industry productivity.[36] The two censuses are unbalanced panels and cover all manufacturing plants with more than ten employees. The data include annual measures of output, two types of labor, capital, and intermediate inputs. I observe nominal value added and deflate it by the ISIC-3-digit industry price level to obtain real value added. Labor is hours worked per year and there are two types of labor: blue- and white-collar. For a closer mapping to the model, I weight the two types by their wages and combine them into an aggregate labor variable (see Greenstreet, 2007).[37] To construct capital stocks I follow the method documented in Tybout and Roberts (1996, Chapter 10, pp. 255 - 256) for the Colombian data, and adopt a similar approach for the Chilean data. Data are reported at the plant level so that I do not (and cannot) distinguish between plants and firms, although there are probably multi-plant firms.

Macro data I use data on aggregate variables such as GDP and TFP from Caselli (2005). The data is for the year 1996 and covers 94 countries.[38] I further use data on external finance to GDP ratios from Beck, Demirguc-Kunt and Levine (2000).[39] The data are for the year 1996 and cover 175 countries. I exclude countries in Sub-Saharan Africa which are those with the lowest TFP and GDP. Those countries likely have other, more pressing, problems than underdeveloped credit markets (think Zimbabwe). After further dropping countries with missing data, I am left with data on aggregate variables and external finance to GDP ratios for 68 countries, including Chile and Colombia.

3.2 Empirical Strategy

My empirical strategy can be summarized in the following three steps. First, I estimate the critical micro parameters (mainly the autocorrelation of productivity shocks and the shape of their distribution) using the plant-level panel data from Chile and Colombia. Second, I identify the parameter governing the quality of credit markets λ using external finance to GDP ratios. This is done for all 68 countries in my macro sample. Third, I examine cross-country TFP differences for the 68 countries in my macro sample and compare them to the predictions of

[36] See Liu (1993), Levinsohn and Petrin (2003), Petrin and Levinsohn (2009) and the book by Tybout and Roberts (1996)

[37] I show in the robustness exercises in Appendix E.1 that results are unchaged when I include both types of labor separately.

[38] A previous version of this paper used GDP and TFP from Hall and Jones (1999), which covers 127 countries for the year 1988. The more recent data from Caselli (2005) are advantageous mainly because of the need to merge them with the external finance to GDP ratios from Beck, Demirguc-Kunt and Levine (2000). The latter are much more reliable and complete in 1996 than in 1988.

[39] External finance is defined to be the sum of private credit, private bond market capitalization, and stock market capitalization.

the model. I ask the question: what fraction of cross-country TFP differences can my model explain? To do so, I assume that the micro parameters are the same across countries so that I can use those estimated for Chile and Colombia (step number one). In turn, step number two delivers the degree of financial development λ across countries. I then plug these parameters into the theoretical model and compare them to cross-country TFP and GDP data. Following this strategy, I obtain my number for TFP losses from financial frictions (up to twenty percent relative to the US).[40]

3.3 Estimation of Micro Parameters

Estimation of Plant Productivity I estimate plant-level productivity as the residual from a standard production function estimation. The model assumes a two-factor Cobb-Douglas production functions of the form[41]

$$y = zk^{\alpha}l^{1-\alpha}. \tag{40}$$

I estimate (40) in a variety of ways. My preferred strategy is to set α equal to the capital share at the plant level, averaged over all plants and all years. Appendix E.1 presents the other estimation strategies and argues that results are robust. After having estimated (40), idiosyncratic productivity z can be backed out as the residual. I renormalize all productivities by their industry mean so as to make them comparable across industry and so that the sample mean is one as in the model. I also trim the 1% tails of productivity in each year to make the results robust to outliers.

My notion of idiosyncratic productivity z has so far been that of physical productivity. However, plants in my data do not report physical output. Instead, output is measured as revenue (divided by a common industry-level deflator). Consequently I cannot distinguish between a firm's physical productivity (TFPQ) and its revenue productivity (TFPR). This distinction is highlighted by Foster, Haltiwanger and Syverson (2008) who argue that revenue and physical productivity are generally different from each other due to differences in prices across establishments. The distinction also features prominently in Hsieh and Klenow (2009) in the context of resource misallocation.[42] The lack of data on physical productivity must be kept

[40]A previous version of this paper also examined the micro fit of the model and showed that it can explain some features of the allocation of capital within Chile and Colombia.

[41]In the theoretical part, (1), I wrote $y = (zk)^{\alpha}l^{1-\alpha}$ because it made some of the algebra easier. With a slight abuse of notation, I here adopt the more conventional form (40); expressions from the model used in the empirical exercise are adjusted accordingly.

[42]Hsieh and Klenow do not use physical output data to measure TFPQ. Instead they measure TFPR and obtain TFPQ from a model of monopolistic competition implying that the two are linked by a simple formula. More specifically, the fact that demand curves have a constant elasticity implies that revenues equal output to some power. The same applies to TFPQ and TFPR.

Country	Chile		Colombia	
Specification	(1)	(2)	(1)	(2)
$Corr(z_{it}, z_{i,t-1})$	0.762	0.702	0.825	0.811
	(0.003)	(0.003)	(0.002)	(0.003)
R-squared	0.533	0.534	0.696	0.702
Observations	40055	40055	47511	47511
Years	10	10	9	9
Industry Dummies	No	Yes	No	Yes

Table 1: Autocorrelation

Note: Estimates of autocorrelation from simple AR(1) regression on productivities z_{it} (pooled sample, productivities normalized by industry mean). Specifications (1) and (2) differ in that (2) adds industry dummies. See Appendix E.1 for robustness exercises.

in mind in some parts of my empirical exercise, for example when examining the distribution of productivities across plants.[43] However, it is *not* a problem in other parts: the ability of firms to self-finance as described in the theoretical model of section 2 depends on profits, and hence on revenue not physical productivity. My estimates of autocorrelation in the next paragraph, for instance, therefore use the correct productivity measure.

Autocorrelation I have argued in the theoretical part of this paper that the autocorrelation of productivity shocks is crucial for assessing TFP losses from financial frictions. This autocorrelation is estimated by running a simple AR(1) regression on productivity z_{it}. One major caveat is that the two Chilean and Colombian panels are only ten and nine years long respectively. This implies that estimates of autocorrelation are subject to small sample biases. In particular, autocorrelation is likely biased downwards (Hurwicz, 1950). While I argue below that the estimates of autocorrelation presented here are broadly consistent with those in the literature, one may therefore want to interpret those estimates as a *lower bound* on true autocorrelation.

This caveat notwithstanding, Table 1 summarizes the results for Chile and Colombia. I estimate two specifications. The first runs a simple AR(1) regression on the pooled sample. The second includes industry dummies. Consider first specification (1). Autocorrelation in Chile (0.762) is slightly lower than in Colombia (0.825). Including industry dummies, these autocorrelations change to 0.702 and 0.811 respectively. These estimates are consistent with other estimates in the literature. For example, Gourio (2008, Table 3) estimates an autocorrelation of 0.78, using a much longer panel (35 years) of large US firms (Compustat). See also Hennessy and Whited (2005) and Cooper and Haltiwanger (2006) for two papers in the investment literature finding similar results. Due to the small sample biases just discussed, I

[43]Note that this problem is common to most studies comparing productivities across firms or plants.

pick an autocorrelation that is towards the upper end of the range in Table 1, $Corr = 0.8$, as my parameter input below.

Shape of Productivity Distribution The model assumes that productivity follows the mean reverting diffusion (33), resulting in a stationary productivity distribution $\psi(z)$ that is Gamma (34). I here argue that – despite its simplicity – such a Gamma distribution describes the data rather well. Figure 6 depicts the density of productivities for Chile (productivities are pooled over all years and all industries). Superimposed is a Gamma distribution (34) that represents the best (maximum-likelihood) fit. Its parameter is estimated to be $\beta = 1.67$.[44] It

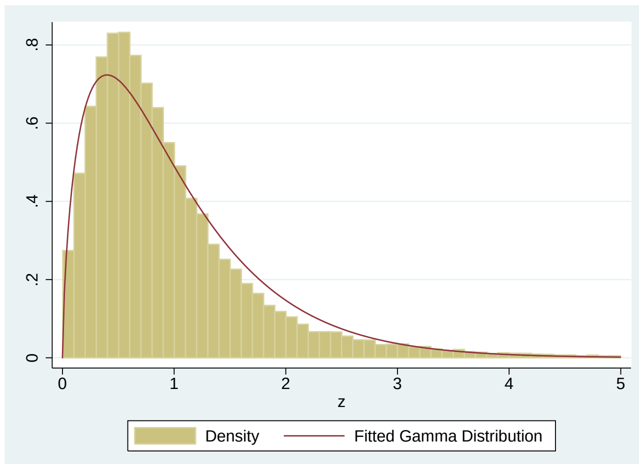

Figure 6: Chile: Density of Productivities

Note: Histogram of productivities z for pooled sample. Production functions are estimated using cost-shares. Productivities normalized by their mean. The Gamma distribution is fitted by maximum likelihood, and its parameter estimated to be $\beta = 1.67$.

can be seen that the parsimoniously specified Gamma distribution (34) represents a good fit of the density of productivities. Despite the model's simplicity, heterogeneity is sufficiently rich to allow for a direct mapping to the data.

One issue arises in relation to the model's prediction that entrepreneurs below a certain productivity threshold $\underline{z}$ will not be active. Taking the model seriously, one should then expect to observe a sharp truncation of the histogram in Figure 6. Such a truncation is apparently absent in the data. I here argue that this might be due to the presence of measurement error.[45]

[44]The sample only covers firms with more than ten employees. This implies that I might be missing some features of the productivity distribution $\psi(z)$. Specifically there could be some highly productive entrepreneurs with very high output but only very few employees. This would lead me to underestimate TFP losses from financial frictions.

[45]Measurement error is ubiquitous, and my data are no exception. I here list three common sources as they apply to my data. First, I construct capital stocks using a perpetual inventory method. Capital stocks are therefore subject to measurement error in the initial capital stock, and the rate of depreciation (which I simply assume constant over time). Second, I have output and investment deflators at the industry level; ideally they would be observed at the firm level, implying an additional source of measurement error. Third, labor input is likely measured with considerable error, especially in the quality dimension.

That is, measurement error implies that, even if such a truncation were in fact present, it likely wouldn't show up in the data. To see this consider Figure 7 which is based on a version of the model that allows for measurement error as explained in Appendix E.2.[46] If productivity is truncated but measured with error, its distribution looks as in Figure 7. Observe in particular

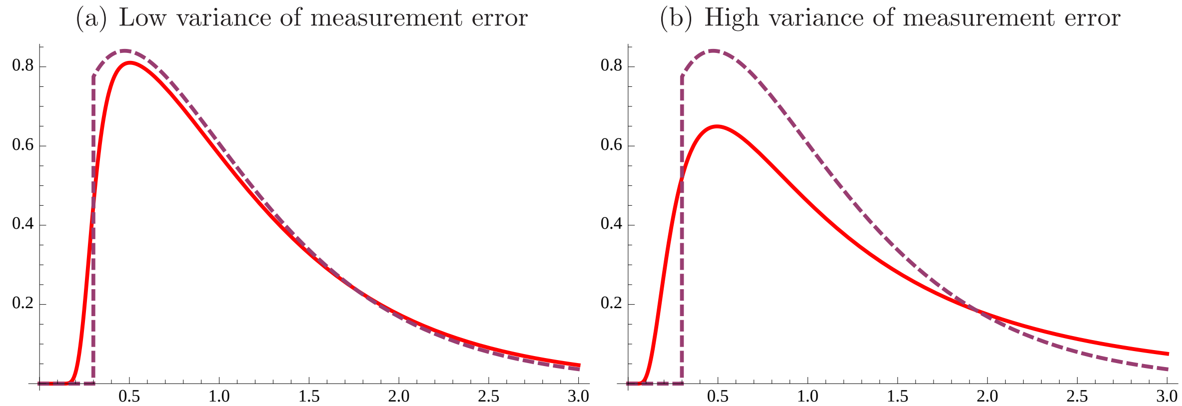

Figure 7: Truncation and Measurement Error

Note: The dashed line plots the distribution of *true* productivity z. The solid line plots the distribution of *observed* productivity $\tilde{z} = z/\varepsilon$ where ε is measurement error. The distributions of both z and ε are Gamma. Panel (a) plots the case $\mathbb{V}(\varepsilon) = 1/20$, panel (b) plots $\mathbb{V}(\varepsilon) = 1/5$. Other parameters are $\underline{z} = 0.3, \beta = 1.67$. See Appendix E.2 for more detail.

that despite *true* productivity being truncated, the distribution of *observed* productivity has full support on the positive real line and is continuous and continuously differentiable at $\underline{z}$; put differently, measurement error dilutes any productivity threshold.[47]

3.4 Identifying λ

I use external finance to GDP ratio from Beck, Demirguc-Kunt and Levine (2000) to identify the parameter λ that governs the degree of financial development across countries. This is possible because these external finance to GDP ratios have a direct counterpart in the model presented above (section 2). The model predicts that the ratio of external finance to *capital* in

[46] I work with a Gamma distribution for productivity. I therefore find it convenient to assume that measurement error is also distributed Gamma (and multiplicative). This is somewhat unconventional, going against the standard assumption that (the log of) measurement error is distributed normally. However, note that for the right choice of parameters, a Gamma distribution closely resembles a log-normal distribution. Also, a remark attributed to physics Nobel laureate Gabriel Lippman applies: "Everyone believes in the [normal] law of errors: the mathematicians, because they think it is an experimental fact; and the experimenters, because they suppose it is a theorem of mathematics".

[47] As already discussed in section 1.3, the literature on management practices across countries Bloom and Van Reenen (2009) does in fact find patterns consistent with a truncation of the productivity distribution. This might be due to better measurement of productivity.

Country	Chile	Colombia	US	India	China
D/Y	1.58	0.52	2.3	0.55	0.20*
λ	2.09	1.21	4.15	1.22	1.07

Table 2: External Finance to GDP ratio (D/Y) in 1996

Note: External finance to GDP ratios in Chile and Colombia are of the same order of magnitude as those for India and China. *The number for China is for 1997 since required data are not available for 1996. λ is calculated from (41), assuming that $\alpha = 1/3, \rho = 0.05, \delta = 0.06$ (implying that $K/Y \approx 3$).

a given economy equals[48]

$$\frac{D}{K} = 1 - \frac{1}{\lambda}.$$

If there are no capital markets, $\lambda = 1$, there is no external finance: $D/K = 0$. If capital markets are perfect, $\lambda = \infty$, the entire capital stock of the economy is financed externally: $D/K = 1$. Together with the expression for the capital output ratio (28) this implies that the external finance to GDP ratio equals

$$\frac{D}{Y} = \frac{D}{K}\frac{K}{Y} = \left(1 - \frac{1}{\lambda}\right)\frac{\alpha}{\rho + \delta}. \tag{41}$$

Table 2 lists external-finance to GDP ratios, D/Y, from Beck, Demirguc-Kunt and Levine (2000) and implies λ's for Chile and Colombia in 1996. For sake of comparison, the table also lists the same variable for the US, India, and China. Chile and Colombia are financially considerably less developed than the United States and on roughly equal footing as India and China. Colombia in particular, is relatively similar to India and China. The same strategy for identifying λ is later made use of for *all countries* in my macro sample (section 3.5). There, I will compute TFP losses from financial frictions relative to the US. The value of $\lambda = 4.15$ for the US should be thought of as relatively close to the first-best value of $\lambda = \infty$.[49]

3.5 Productivity Losses from Financial Frictions Across Countries

In this section, I finally put together all the pieces from the preceding sections and explain how I quantify TFP losses from financial frictions. I argue that financial frictions have a substantial, adverse impact on a country's TFP and in turn on per-capita GDP. A priori, one might expect financial frictions to explain a sizeable fraction of cross-country TFP differences, but surely

[48]To see this note that all active entrepreneurs borrow as much as they can, that is individual borrowing is $d = (\lambda - 1)a$ if $z \geq \underline{z}$ and all inactive entrepreneurs lend. Total borrowing in the economy therefore equals $D = \mathbb{E}[d|d \geq 0] = (\lambda - 1)(1 - \Omega(\underline{z}))K = (\lambda - 1)/\lambda K$ where the last equality uses the market clearing condition $\lambda(1 - \Omega(\underline{z})) = 1$.

[49]The corresponding US external finance to capital ratio is $D/K = 0.76$. With $\lambda = \infty$, it would be one as discussed above.

not all. For example, Hsieh and Klenow (2009, section IV) argue that resource misallocation of both capital and labor can explain roughly 50 percent of the TFP gap between the US and China, and 35 percent of the TFP gap between the US and India. In turn, capital is only one resource that is potentially misallocated; and financial frictions are only one potential reason for why this may be the case. That is, when carrying out the exercise of assessing TFP losses from financial frictions, it is important to keep in mind that we are varying one single factor (financial frictions) across countries.

That being said, Figure 8 plots TFP and per-capita GDP against the external finance to GDP ratio, D/Y. Both TFP and GDP are relative to the US. The model predictions are

(a) TFP (b) GDP

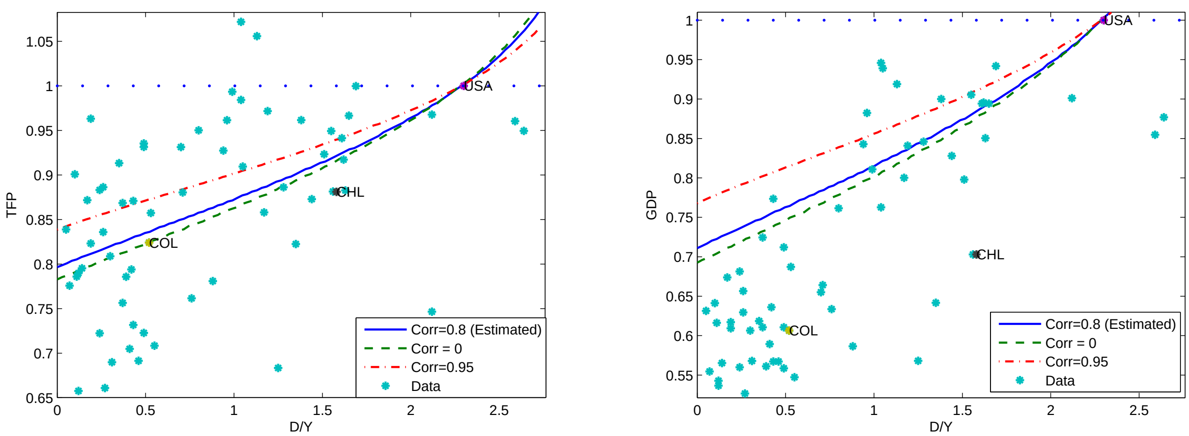

Figure 8: TFP and GDP Losses from Financial Frictions: Data and Model

Note: TFP (panel a) and per-capita GDP (panel b) against external finance to GDP ratio. Solid line: preferred model specification, dashed lines: model prediction with alternative autocorrelation, dots: data (countries) after eliminating 50 percent of the gap relative to the US. That is, TFP is calculated as $TFP = TFP^{DATA} + 0.5(1 - TFP^{DATA})$, and similarly for GDP. A dot on the solid line represents a country for which the model explains 50 percent of the gap relative to the US.

given by the solid and dashed lines.[50] The difference between the solid and dashed lines lies in the value used for the critical autocorrelation parameter. The solid line is my preferred model specification using the estimated autocorrelation of 0.8. The purpose of the dashed lines is to examine how robust results are to different values of the autocorrelation. The dots in the figure represent data (68 countries). For sake of better comparison to studies on resource misallocation, I rescale TFP and GDP in the data so as to eliminate 50 percent of the TFP/GDP gap between a given country and the US.[51] A dot on the solid line is then a country for which

[50]The model's parameters are obtained as explained in the preceding sections. That is, micro-parameters are estimated from micro-data as explained in section 3.3 ($Corr = 0.8$, $\beta = 1.67$), and the parameter governing the quality of credit markets, λ, is identified from external finance to GDP ratios, D/Y, as explained in section 3.4. Other parameters are $\alpha = 1/3$, $\rho = 0.05$, $\delta = 0.06$.

[51]This rescaling takes seriously the statement in the first paragraph of this subsection that there is likely a large variety of determinants of cross-country TFP differences, of which capital misallocation due to financial frictions is only one.

the model explains exactly 50 percent of the gap relative to the US.

Consider first the solid line. In my model, the variation in financial frictions can bring down TFP to 80 percent of the US level (panel a). It also accounts for 47 percent of the Colombia-US gap, 33 percent the Chile-US gap, and 27 percent of both the India-US and China-US gaps. Other countries for which the model can explain roughly 50 percent of the TFP gap to the US (i.e. dots on the solid line) are Paraguay and Papua New Guinea. These TFP differences in turn translate into GDP differences of up to 29 percent (panel b). The model can explain roughly 50 percent of the Argentina-US and 30 percent of the China-US (GDP) gap. Summarizing, in my model, financial frictions lower TFP by up to 25 percent relative to the US. Financial frictions can also account for up to fifty percent of the TFP and GDP gaps between the US and typical countries in my sample.[52]

4 Concluding Remarks

In my framework, self-financing can undo capital misallocation from financial frictions only if idiosyncratic productivity shocks are relatively persistent. The reason is that entrepreneurs accumulate wealth out of past successes so only if high productivity episodes are sufficiently prolonged can they accumulate sufficient internal funds to self-finance their desired investments. As a result, the extent of capital misallocation and therefore TFP losses from financial frictions depend critically on the autocorrelation of idiosyncratic productivity shocks.

I have made this point in a heterogenous-agent growth model with borrowing constraints and forward-looking savings behavior. While featuring sufficiently rich heterogeneity to be mapped directly to micro-data, the model remains highly tractable and TFP is simply a truncated weighted average of individual productivities. In contrast to similar existing theories, the model presented here also allows for persistent productivity shocks. The self-financing mechanism that takes center stage in my paper implies that any empirically serious theory of heterogenous firms and financial frictions must feature such persistence. The tools presented in this paper should therefore also prove useful in other applications.

While developing this theory is the main contribution of this paper, I also apply the theory to quantify TFP losses from financial frictions across countries. I estimate the critical micro-parameters using plant-level panel data from Chile and Colombia and identify the parameter

[52]These estimates are in the same "ballpark" as those found in other studies. For example, Buera and Shin (2010) find output losses of around 30 percent (see their Figure 2 – they don't compare those numbers to data but instead compare the case of no credit markets to that of $\lambda = 5$ (λ is the same parameter as in my model). Buera, Kaboski and Shin (2010) find larger effects: TFP losses of around 40 percent and corresponding GDP losses of 55 percent (see their Figure 4). The main difference is that their model features fixed costs which introduce an additional "misallocation of talent". See also the discussion in the conclusion.

governing the degree of financial development from external finance to GDP ratios. Using these estimates, I calculate that financial frictions can explain productivity losses of up to twenty percent (see Figure 8).

Finally, let me briefly comment on an extension of the present framework that seems both natural and potentially important. This is the idea that it may be fixed costs in production (or non-convexities more generally) that lead financial frictions to have large real effects.[53] One reason for why this may be true is that fixed costs will generally introduce an additional margin of misallocation: an extensive margin resulting in a "misallocation of talent". This is illustrated in Figure 9 which is based on such a model, laid out in Appendix F. The model is static as in section 1 but features fixed costs. Intuitively, rich entrepreneurs find it easier to cover the fixed cost so that some incompetent-but-wealthy entrepreneurs are in business while some productive-but-poor ones are not. While fixed costs can result in such interesting predictions, they are also complicated to analyze except in the static case as here because their presence makes the savings problems of entrepreneurs non-convex. Such challenges are left for future research.

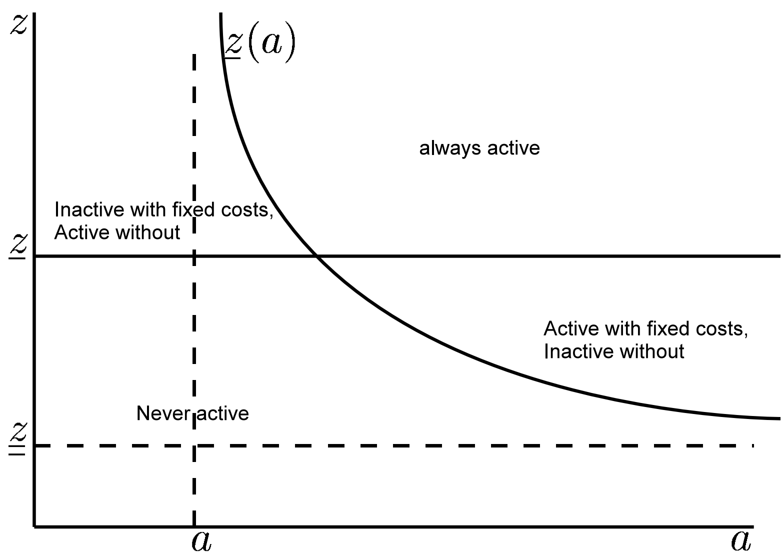

Figure 9: Fixed Costs: Misallocation of Talent on Extensive Margin

Note: Fixed costs introduce an additional margin of misallocation because high productivity, low wealth entrepreneurs cannot cover the fixed cost while low productivity, high wealth entrepreneurs can. Without fixed costs there is only a cutoff in productivity space labelled $\underline{z}$. With fixed costs the cutoff is downward-sloping in wealth-productivity space and given by the line $\underline{z}(a)$.

[53]This idea goes back to Galor and Zeira (1993); a more recent contribution is by Buera, Kaboski and Shin (2010) who make the point that fixed costs introduce a misallocation of talent of the type discussed here. See also Banerjee and Duflo (2005).

Appendix

A Proofs

A.1 Proof of Lemma 1

From the profit maximization problem (3), optimal labor demand is $l = (\pi/\alpha)^{1/(1-\alpha)} zk$, where π is as in the Lemma. Plugging back in, the profit function becomes

$$\Pi(a,z) = \max_k \{z\pi k - (r+\delta)k \quad \text{s.t.} \quad k \leq \lambda a\}.$$

Since this problem is linear, it follows immediately that k is either zero or λa. The cutoff $\underline{z}$ is defined as the value for which entrepreneurs are indifferent between running a firm and being inactive, $\Pi(a, \underline{z}) = 0.\square$

A.2 Proof of Proposition 1

Using the expression for factor demands in Lemma 1, labor demand can be written as

$$l(a,z) = \left(\frac{\pi}{\alpha}\right)^{1/(1-\alpha)} \lambda a z, \quad z \geq \underline{z}, \tag{42}$$

and zero otherwise. It follows that individual output is $y(a,z) = (\pi/\alpha)\lambda a z$, if $z \geq \underline{z}$, and zero otherwise. Aggregate output is then

$$Y = \int_0^\infty \int_0^\infty y(a,z) g(a,z) da dz = \frac{\pi}{\alpha} \lambda X K, \quad \text{where} \quad X = \int_{\underline{z}}^\infty z\omega(z) dz \tag{43}$$

is an auxiliary variable. Next, consider the labor market clearing condition (42). Integrating over all a and z,

$$L = \left(\frac{\pi}{\alpha}\right)^{1/(1-\alpha)} \lambda X K. \tag{44}$$

Substituting into (43), we see that $Y = (\lambda X K)^\alpha L^{1-\alpha}$. Eliminating λ using (9) this is (7) and (8) in the proposition. Substituting the definition of π from Lemma 1 into (44) and rearranging yields the expression for w. Substituting (44) into the cutoff condition $\underline{z}\pi = R$ and rearranging yields the expression for $R.\square$

A.3 Proof of Lemma 2

From Lemma 1, we know that $\dot{a} = A(z)a - c$ where $A(z) = \lambda \max\{z\pi - r - \delta, 0\} + r$. The Bellman equation is then (see Ch.2 in Stokey, 2009),

$$\rho v(a,z) = \max_c \left\{ \log c + \frac{1}{dt} \mathbb{E}[dv(a,z)] \quad \text{s.t.} \quad \dot{a} = A(z) - c \right\}.$$

The proof proceeds with a guess and verify strategy. Guess that the value function takes the form $v(a,z) = V(z) + B \log a$. Using this guess we have that $\mathbb{E}[dv(a,z)] = (B/a)da + \mathbb{E}[dV(z)]$. Rewrite the value function

$$\rho V(z) + \rho B \log a = \max_c \ \log c + \frac{B}{a}[A(z)a - c] + \frac{1}{dt}\mathbb{E}[dV(z)].$$

Take first order condition to obtain $c = a/B$. Substituting back in,

$$\rho V(z) + \rho B \log a = \log a - \log B + A(z)B - 1 + \frac{1}{dt}\mathbb{E}[dV(z)].$$

Collecting the terms involving $\log a$, we see that $B = 1/\rho$ so that $c = \rho a$ and $\dot{a} = [A(z) - \rho]a$ as claimed.$\square$

A.4 Proof of Proposition 2

Throughout this proof, I omit indexing by t for notational simplicity. As in the static model, the capital market clearing condition (16) can be written as

$$\lambda \int_{\underline{z}}^\infty \omega(z) dz = 1 \quad \text{or} \quad \lambda(1 - \Omega(\underline{z})) = 1. \tag{45}$$

Consider next the law of motion for aggregate capital (19). Using that the shares $\omega(z)$ integrate to one, we have that

$$\frac{\dot{K}}{K} = \lambda\pi \int_{\underline{z}}^{\infty} z\omega(z)dz - \lambda(r+\delta)\int_{\underline{z}}^{\infty} \omega(z)dz + r - \rho.$$

Using capital market clearing (45),

$$\frac{\dot{K}}{K} = \lambda\pi X - (\rho+\delta), \quad X \equiv \int_{\underline{z}}^{\infty} z\omega(z)dz. \tag{46}$$

Next, consider the labor market clearing condition. As in the proof of Proposition 1, equation (44), we have that labor market clearing implies $\pi = \alpha(\lambda X)^{\alpha-1}K^{\alpha-1}L^{1-\alpha}$. Substituting into (46) and rearranging, we get

$$\dot{K} = \alpha Z K^{\alpha} L^{1-\alpha} - (\rho+\delta)K, \quad Z = (\lambda X)^{\alpha}.$$

After substituting for λ from (45), this is equation (21) in Proposition 2.$\square$

A.5 Proof of Proposition 3

The law of motion for the joint distribution of wealth is given by the Kolmogorov Forward Equation (see for example Stokey, 2009, p.50).

$$\frac{\partial g(a,z,t)}{\partial t} = -\frac{\partial}{\partial a}\left[g(a,z,t)s(z,t)a\right] - \frac{\partial}{\partial z}\left[g(a,z,t)\mu(z)\right] + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left[\sigma^2(z)g(a,z,t)\right]. \tag{47}$$

Using the definition of $\omega(z,t)$ we have that

$$\frac{\partial\omega(z,t)}{\partial t} = \frac{1}{K(t)}\int_0^{\infty} a\frac{\partial g(a,z,t)}{\partial t}da - \frac{\dot{K}(t)}{K(t)}\omega(z,t). \tag{48}$$

Using an integration by parts

$$-\int_0^{\infty} a\frac{\partial}{\partial a}\left[g(a,z,t)s(z,t)a\right]da = s(z,t)\int_0^{\infty} ag(a,z,t)da. \tag{49}$$

Plugging (47) into (48) and using (49), we obtain the PDE (30). Setting the time derivative equal to zero, one obtains the ODE (31).$\square$

A.6 Proof of Proposition 4

With $\lambda = 1$, $s(z) = z\pi - \rho - \delta$. Using the drift and diffusion in the stochastic process (33), the ODE (31) becomes

$$0 = [z\pi - \rho - \delta]\omega(z) - \frac{d}{dz}\left[\left(\left(\nu + \frac{\sigma^2}{2}\right)z - \nu z^2\right)\omega(z)\right] + \frac{1}{2}\frac{d^2}{dz^2}\left[\sigma^2 z^2\omega(z)\right].$$

There is an additional restriction ensuring that aggregate capital is constant (54). This will be crucial below. Guess a functional form $\omega(z) = e^{-\beta z}z^{\gamma-1}$ (provided that the solution integrates, it can always be scaled so as to integrate to one). Substitute the guess into the ODE and proceed by equating coefficients on three terms:

$$e^{-\beta z}z^{\gamma-1}, \quad e^{-\beta z}z^{\gamma}, \quad e^{-\beta z}z^{\gamma+1}. \tag{50}$$

Consider first the coefficients on the third term. The level term does not contribute to this, the drift term contributes $-\beta\nu$ and the diffusion term $\beta^2\sigma^2/2$. Thus $0 = -\beta\nu + \beta^2(\sigma^2/2)$, or $\beta = 2\nu/\sigma^2$. Consider next the second term of (50). The level term contributes π, the drift term contributes $\beta(\nu+\sigma^2/2)+\nu(\gamma+1)$ and the diffusion term $-\sigma^2\beta(\gamma+1)$ so that $0 = \pi + \beta(\nu+\sigma^2/2) + \nu(\gamma+1) - \sigma^2\beta(\gamma+1)$. Rearrange and use the expression for β,

$$0 = \pi + \nu\beta - \nu\gamma. \tag{51}$$

For the first term in (50), the level term contributes a coefficient of $-\rho-\delta$, the drift term contributes $-\gamma(\nu+\sigma^2/2)$, and the diffusion term contributes $\gamma(\gamma+1)\sigma^2/2$ so that $0 = -\rho - \delta - \gamma\left(\nu + \sigma^2/2\right) + \gamma(\gamma+1)\sigma^2/2$, or again using the expression for β

$$0 = \gamma^2 - \gamma\beta - (\rho+\delta)\frac{\beta}{\nu}. \tag{52}$$

Consider next condition (54). With $\lambda = 1$, $\underline{z} = 0$ so that $\pi\mathbb{E}_\omega[z] = \rho + \delta$. Using the Gamma guess, we have that $\mathbb{E}_\omega(z) = \gamma/\beta$ so that, $\pi(\gamma/\beta) = \rho + \delta$. Substituting into (51), it can be verified that (51) and (52) coincide. γ is then found by solving the quadratic (52) and picking the root that is greater than one (which ensures that $\omega(z)$ doesn't explode as z approaches zero).

B Equivalence between Renting and Owning/Accumulating Capital

I here show that the budget constraint (14) can be derived from a setup in which entrepreneurs own and accumulate capital themselves and trade in risk-free bonds. For sake of clarity, I present the argument for a discrete approximation to the continuous-time framework. Periods are of length Δ. The continuous-time counterparts of the expressions can be obtained by taking Δ to zero. The stock of bonds issued by an entrepreneur, that is his debt, is denoted by d_t. When $d_t < 0$ the entrepreneur is a net lender. In order for there to be an interesting role for credit markets, an entrepreneur's productivity z_t is revealed at the end of period $t - \Delta$, before the entrepreneur issues his debt d_t. That is, entrepreneurs can borrow to finance investment corresponding to their new productivity. The budget constraint and law of motion for capital are

$$0 = \Delta(y_t - w_t l_t - r_t d_t - x_{t+\Delta} - c_t) + d_{t+\Delta} - d_t, \quad k_{t+\Delta} = \Delta x_{t+\Delta} + (1 - \Delta\delta)k_t.$$

where $x_{t+\Delta}$ investment in physical capital. These can be combined as

$$k_{t+\Delta} - d_{t+\Delta} = \Delta(y_t - w_t l_t - \delta k_t - r d_t - c) + k_t - d_t. \tag{BC+}$$

I now argue that by changing slightly the time at which the budget constraint is "recorded", we can derive the budget constraint (14). To this end, define d_t^- and k_t^- as debt and capital *before* investment is made, i.e.

$$d_t^- \equiv d_t - \Delta x_t, \quad k_t^- \equiv k_t - \Delta x_t.$$

The timeline in Figure 10 illustrates this change in the "recording time" of the budget constraint. Using

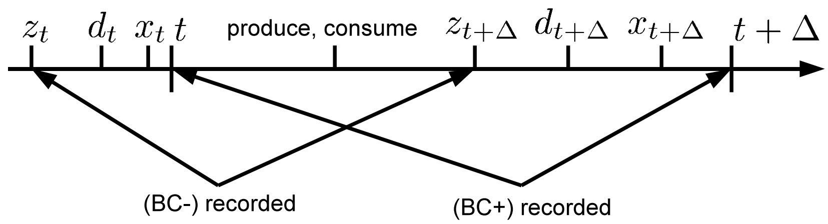

Figure 10: Time Line

Entrepreneurs first observe their productivity z_t, then issue debt d_t to finance investment x_t. The budget constraints (BC+) is recorded *after* investment x is made; (BC-) is recorded *before* investment is made. This amounts to moving the start of time t forward before investment is made. The two are equivalent.

these definitions we can rewrite (BC+) as

$$k_{t+\Delta}^- - d_{t+\Delta}^- = \Delta[y_t - w_t l_t - \delta k_t - r(d_t^- - k_t^- + k_t)] + k_t^- - d_t^-. \tag{BC-}$$

Defining total wealth as $a_t \equiv k_t^- - d_t^-$, and collecting terms we get $a_{t+\Delta} = \Delta[y_t - w_t l_t - (r_t + \delta)k_t + r a_t] + a_t$. Rearranging, dividing by Δ and letting Δ tend to zero we obtain (14).

C Computation of Equilibrium: General Case

C.1 List of Equilibrium Conditions

Due to space restrictions, I only include the conditions for a *stationary* equilibrium as in Corollary 1. Instead of characterizing the equilibrium in terms of the prices w and r, it turns out to be easier to do the analysis in terms of the productivity cutoff $\underline{z}$ and the constant $\pi = \alpha\left[(1-\alpha)/w\right]^{(1-\alpha)/\alpha}$, which was defined in Lemma 1. A steady state equilibrium is a function $\omega : \mathbb{R}^+ \to \mathbb{R}^+$ and constants π and $\underline{z}$ satisfying

$$0 = [\lambda\pi \max\{z - \underline{z}, 0\} + \underline{z}\pi - \rho - \delta]\omega(z) - \frac{d}{dz}[\mu(z)\omega(z)] + \frac{1}{2}\frac{d^2}{dz^2}\left[\sigma^2(z)\omega(z)\right], \quad \int_0^\infty \omega(z)dz = 1, \tag{53}$$

$$\lambda\pi \int_{\underline{z}}^\infty z\omega(z)dz = \rho + \delta, \tag{54}$$

$$\lambda \int_{\underline{z}}^{\infty} \omega(z) dz = 1. \tag{55}$$

Interpretation: (53) is (31) after substituting in for the savings rate $s(z)$ from Lemma 2 and using the productivity cutoff $\underline{z}\pi = r + \delta$. (54) is the condition that aggregate capital is constant $\dot{K} = 0$, and (55) is the capital market clearing condition.

C.2 General Computational Algorithm for Solving Steady State Equilibrium

See the equilibrium conditions in Appendix C.1, (53) to (55). Algorithm:

1. Construct a grid $\pi^i, i = 1, ..., N$.
2. Guess $\underline{z}^{i,j}$ and solve the ODE (53), using a finite difference method. With this method, it does not matter that $s(z)$ is not differentiable at $\underline{z}$.
3. Check the capital market clearing condition (55). If there is excess demand (supply), choose a cutoff $\underline{z}^{i,j+1}$ that is greater (smaller) than $\underline{z}^{i,j}$ (bisection method).
4. Repeat steps (2)-(3) until the capital market clears for each point in the grid $\pi^i, i = 1, ..., N$.
5. Referring to (54), calculate $\Delta(\pi) = |\lambda\pi \int_{\underline{z}}^{\infty} z\omega(z)dz - \rho - \delta|$ for each $\pi^i, i = 1, ..., N$, and choose the π^i for which $\Delta(\pi)$ is smallest.

C.3 A Convenient Approximation

When many equilibria need to be computed for different parameter configurations (e.g. to plot Figures **??** and 8), I use a computationally very inexpensive approximation: ignore the direct effect of λ on the savings rate, and instead use the closed form solutions for the wealth shares, $\omega(z)$, computed in section 2.7 under the assumption $\lambda = 1$. λ then affects TFP only through the truncation $\underline{z}$. Figure 11 plots exact TFP, computed as described in section C.2 and approximate TFP as just described for the same parameter values as in Figure **??**. It can be seen that this approximation works extremely well in practice.

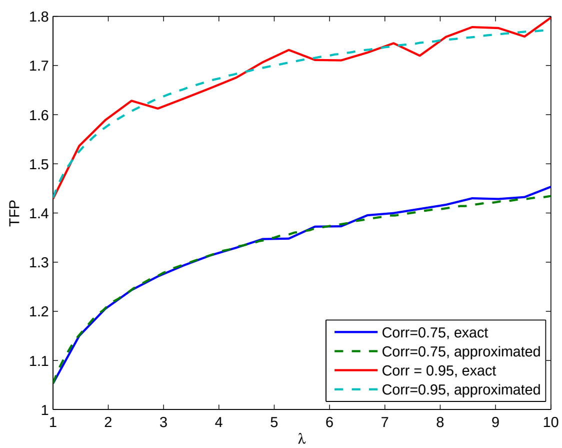

Figure 11: Exact and Approximate TFP

Note: Solid Lines are exact TFP; dashed lines are approximate TFP. Exact TFP is less smooth because it is computed for fewer λ-grid points and suffers from some numerical error, especially for high λ and autocorrelation.

C.4 Continuity and Differentiability of Wealth Shares

A computational difficulty arises because $s(z) = \lambda\pi \max\{z - \underline{z}, 0\} + \underline{z}\pi - \rho$ is generally not differentiable at $z = \underline{z}$. I here prove that despite this fact, $\omega(z)$ is continuous and once differentiable everywhere. The proof uses a discrete approximation. Consider a general diffusion

$$dz = \mu(z)dt + \sigma(z)dt.$$

Under certain regularity conditions this diffusion can be approximated by a binomial tree specified as follows.[54] Divide time into discrete periods of length Δt; start at some z; with probability $p(z)$ the process moves up some

[54] See Nelson and Ramaswamy (1990) who in turn use results from Stroock and Varadhan (1979).

distance Δz and with probability $q(z) = 1 - p(z)$ it moves down. The step size and probabilities are given by (equations 21-23 in Nelson and Ramaswamy)

$$\Delta z = \sqrt{\Delta t}\sigma(z), \quad p(z) = \frac{1}{2}\left(1 + \frac{\mu(z)}{\sigma(z)}\sqrt{\Delta t}\right), \quad q(z) = 1 - p(z) = \frac{1}{2}\left(1 - \frac{\mu(z)}{\sigma(z)}\sqrt{\Delta t}\right).$$

As $\Delta t \to 0$, the resulting binomial process converges to the diffusion above.[55] Using the relationship between the time step and the grid size, the probabilities can also be written as

$$p(z) = \frac{1}{2}\left(1 + \frac{\mu(z)}{\sigma^2(z)}\Delta z\right), \quad q(z) = 1 - p(z) = \frac{1}{2}\left(1 - \frac{\mu(z)}{\sigma^2(z)}\Delta z\right). \tag{56}$$

Next consider the savings behavior of and entrepreneur with productivity z: if he starts out with wealth a_t at time t, he ends up with $a_{t+\Delta t} = s(z)\Delta t a_t + a_t$ at time $t + \Delta t$. Next, consider the point $(z, t + \Delta t)$. This could have arisen in either of two ways: $(z - \Delta z, t)$ followed by an up move, or $(z + \Delta z, t)$ followed by a down move. Combining these two observations, the wealth held by entrepreneurs with ability z at time $t + \Delta t$ must satisfy

$$\omega(z, t + \Delta t) = p(z - \Delta z)[s(z - \Delta z)\Delta t + 1]\omega(z - \Delta z, t) + q(z + \Delta z)[s(z + \Delta z)\Delta t + 1]\omega(z + \Delta z, t). \tag{57}$$

Continuity: Wealth shares at point $(z - \Delta z, t + \Delta t)$ satisfy

$$\omega(z - \Delta z, t + \Delta t) = p(z - 2\Delta z)[s(z - 2\Delta z)\Delta t + 1]\omega(z - 2\Delta z, t) + q(z)[s(z)\Delta t + 1]\omega(z, t).$$

Taking limits as $\Delta z \to 0$ in (57) yields $\lim_{x\uparrow z} \omega(x, t) = \frac{1}{2}[\lim_{x\uparrow z} \omega(x, t) + \omega(z, t)]$, or $\lim_{x\uparrow z} \omega(x, t) = \omega(z, t)$. A symmetric argument around the point $(z+\Delta z, t+\Delta t)$ proves that $\lim_{x\downarrow z} \omega(x, t) = \omega(z, t)$, so that $\lim_{x\uparrow z} \omega(x, t) = \lim_{x\downarrow z} \omega(x, t) = \omega(z, t)$.□

Differentiability: The proof proceeds by taking a first-order approximation in (57) around z. The approximation of $\omega(z - \Delta z)$ is not straightforward. For any points $z - \Delta z$ and $x < z$,

$$\omega(z - \Delta z, t) \approx \omega(x, t) + \omega_z(x, t)[(z - \Delta z) - x].$$

Taking x to z, we have

$$\omega(z - \Delta z, t) \approx \lim_{x\uparrow z} \omega(x, t) - \lim_{x\uparrow z} \omega_z(x, t)\Delta z = \omega(z, t) - \lim_{x\uparrow z} \omega_z(x, t)\Delta z,$$

where the second equality uses continuity from the first part of the proof. A similar approximation holds for $\omega(z + \Delta z, t)$. Take a first-order approximation to all terms in (57) *except* to the terms involving $s(z - \Delta z)$ and $s(z + \Delta z)$:

$$\begin{aligned} \omega(z, t) + \omega_t(z, t)\Delta t = {} & [p(z) - p'(z)\Delta z][s(z - \Delta z)\Delta t + 1][\omega(z, t) - \lim_{x\uparrow z} \omega_z(x, t)\Delta z] \\ & + [q(z) + q'(z)\Delta z][s(z + \Delta z)\Delta t + 1][\omega(z, t) + \lim_{x\downarrow z} \omega_z(x, t)\Delta z] + o(\Delta z). \end{aligned} \tag{58}$$

Dropping all terms that are of order higher than Δz, and rearranging

$$0 = -p(z) \lim_{x\uparrow z} \omega_z(x, t)\Delta z + q(z) \lim_{x\downarrow z} \omega_z(x, t)\Delta z + [-p'(z) + q'(z)]\omega(z, t)\Delta z + o(\Delta z).$$

Importantly, the terms $s(z - \Delta z)\Delta t$ and $s(z + \Delta z)\Delta t$ always drop because they are of order higher than Δz. Dividing by Δz, taking limits as $\Delta z \to 0$, and using that $p(z)$ and $q(z)$ tend to 1/2 while $p'(z)$ and $q'(z)$ tend to zero, we get $0 = \frac{1}{2}[\lim_{x\uparrow z} \omega_z(x, t) - \lim_{x\downarrow z} \omega_z(x, t)]$, which immediately implies the differentiability condition.□

[55] As noted by Nelson and Ramaswamy, the resulting tree does not recombine in the sense that an up move followed by a down move does not bring the process back to the same node. This is a problem for computational approaches but not for the theoretical derivations presented here.

D A Second Example with Closed Form Solution

The following stochastic process which is known as a Feller square root process also yields closed form solutions:[56]

$$dz = \nu(1-z)dt + \sigma\sqrt{z}dW,$$

where ν and σ are positive. The stationary distribution is given by $\psi(z) \propto e^{-\gamma z} z^{\gamma-1}$, $\gamma = 2\nu/\sigma^2$, which is again a Gamma distribution. The mean and variance are $\mathbb{E}(z) = 1$, $\mathbb{V}(z) = \sigma^2/(2\nu) = 1/\gamma$. Again, the assumption $\sigma^2 < 2\nu$ is imposed. As shown in Wong (1964) for example, the autocorrelation is $Corr(z(t), z(t+s)) = e^{-\nu s} \in (0, 1]$. As in Proposition 4, the wealth shares can be found through a guess-and-verify strategy and are

$$\omega(z) \propto e^{-\beta z} z^{\gamma-1}, \quad \beta = \gamma - \frac{\rho+\delta}{\nu}, \quad \gamma = \frac{2\nu}{\sigma^2}.$$

The reader can verify that these behave qualitatively exactly as in Figure 2 for the wealth shares resulting from the geometric mean reverting process in the main text. Similarly, TFP is given by

$$Z = \left(\frac{1}{1-(\rho+\delta)/(\gamma\nu)}\right)^{\alpha}. \tag{59}$$

TFP also behaves qualitatively as (39) in Figure 4 (i.e. it is also convex in $Corr = \exp(-\nu)$).

E Appendix for Empirical Part

E.1 Estimation of Micro Parameters: Robustness

Estimation of Plant Productivity: TO BE (RE)DONE.

Autocorrelation: TO BE (RE)DONE.

Shape of Productivity Distribution: The main text reported only results for Chile. Figure 12 plots the histogram of productivities and a fitted Gamma distribution for Colombia (productivities estimated using cost shares in (40)). The parameter of the Gamma distribution is estimated to be $\beta = 2.32$ which translates

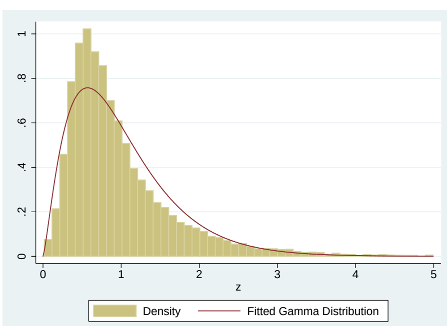

Figure 12: Colombia: Density of Productivities. Note: Histogram of productivities z for pooled sample. Production functions are estimated using cost-shares. Productivities normalized by their mean. The Gamma distribution is fitted by maximum likelihood, and its parameter estimated to be $\beta = 2.32$.

to a variance of $\mathbb{V}(z) = 1/\beta = 1/2.32$. This variance is considerably smaller than the variance for Chile which was 1/1.67. Similarly, I estimate the shape parameter of the distribution if productivity is estimated using the Olley-Pakes methodology described above. The implied variance is 1/1.53 for Chile and 1/2.51 for Colombia which differs again. However, note that none of the results reported in section 3.5 are particularly sensitive to the value of β.

E.2 Adding Measurement Error to the Model

I here show how to add measurement error to the model of section 2. I have shown there that under the stochastic process for productivity (33), productivity z is distributed Gamma, $\psi(z) \propto e^{-\beta z} z^{-\beta-1}$. Suppose instead of observing z, the econometrician observes $\tilde{z} = z/\varepsilon$. For convenience I assume that ε is also distributed Gamma with parameter β_ε, $\psi_\varepsilon(\varepsilon) \propto e^{-\beta_\varepsilon z} z^{-\beta_\varepsilon - 1}$ (ε then has a mean of 1 – see also the discussion in footnote

[56]See Cox, Ingersoll and Ross (1985) for an application in finance. The expressions are actually neater than for the geometric mean reverting process (33) presented in the main text. The main reason for preferring the geometric process are advantages in the empirical implementation, for example that logs can be used for estimating most parameters.

46). It then follows that observed productivity has the distribution[57]

$$\tilde{\psi}(\tilde{z}) = \int_0^\infty \psi(\tilde{z}\varepsilon)\,\psi_\varepsilon(\varepsilon)\varepsilon d\varepsilon. \tag{60}$$

Consider next the case (as in Figure 7) where z is truncated from below at some $\underline{z}$. Using (60), one can then show that the distribution of $\tilde{z}$ is $\tilde{\psi}(\tilde{z}) \propto \tilde{z}^{-\beta_\varepsilon} E\left[1-\beta-\beta_\varepsilon, \underline{z}(\beta+\beta_\varepsilon/\tilde{z})\right]$, where $E(n,z)$ is the exponential integral function. All other results on measurement error are derived in a similar fashion. Note that the stark truncation disappears in general.

Choice of error variance: Measurement error has mean $\mathbb{E}(\varepsilon)=1$ and variance $\mathbb{V}(\varepsilon)=1/\beta_\varepsilon$. It is hard to say what a reasonable value for this variance is. In Figure 7 I choose $\mathbb{V}(\varepsilon)=1/20$ in panel a (low variance) and 1/5 in panel b (high variance). This should be contrasted to the variance of productivities which I estimate to be $\mathbb{V}(z)=1/\beta=1/1.67$ in Chile and 1/2.32 in Colombia.

F Fixed Costs: An Example

This section presents an extension of the static model in section 1 that features fixed costs in production. Consider production functions of the form

$$f(z,k,l) = (z\max\{k-\underline{k},0\})^\alpha l^{1-\alpha}.$$

In order to produce anything, entrepreneurs first have to pay a fixed cost $\underline{k}$ in units of capital. With $\underline{k}=0$ we are back in the environment analyzed in the main text. Without fixed costs, an entrepreneur's decision whether to produce was solely based on his productivity z, specifically whether it was above some threshold $\underline{z}$. With fixed costs this is no longer true. By following the same steps as in Lemma 1, one can show that an entrepreneur's profit is given by

$$\Pi(a,z) = \lambda\max\{z\pi\max\{a-\underline{a},0\}-(r+\delta)a,0\},$$

where $\underline{a}=\underline{k}/\lambda$ is the amount of wealth required to cover the fixed cost. The productivity cutoff for being an active entrepreneur now depends on that entrepreneur's wealth and is given by

$$\underline{z}(a) = \frac{r+\delta}{\pi}\frac{a}{a-\underline{a}} = \underline{z}\frac{a}{a-\underline{a}}, \quad a>\underline{a}.$$

The cutoff is decreasing in an entrepreneur's wealth and it can already be anticipated that it is precisely this feature that generates a misallocation of talent. While this feature is interesting and has possibly important implications for aggregate TFP, it also makes aggregation harder. However, for the special case where productivity is independent of wealth and distributed Pareto, it is still possible to get a nice characterization of aggregate TFP.[58]

Proposition 5 *Consider the static Pareto example in section 1.4, but assume that there are fixed costs in production. Then aggregate GDP is*

$$Y = ZK^\alpha L^{1-\alpha}, \quad \textit{where} \quad Z = \left(\frac{\eta}{\eta-1}\underline{\underline{z}}\right)^\alpha = \mathbb{E}[z|z\geq\underline{\underline{z}}]^\alpha.$$

The cutoff $\underline{\underline{z}}$ is strictly lower than the cutoff $\underline{z}$ in an economy without fixed costs. TFP, Z, is therefore strictly lower as well.

---

[57]In general, given two random variables with joint distribution $f_{XY}(x,y)$, the distribution of the ratio $Z=Y/X$ is (see for example, `http://mathworld.wolfram.com/RatioDistribution.html`)

$$f_Z(z) = \int_{-\infty}^\infty |x| f_{XY}(x,zx)dx.$$

[58]Note that this Proposition is solely about aggregation in a static setting, that is taking as given the distribution of wealth. It doesn't say anything about wealth accumulation. As discussed in Buera (2009*b*) and Banerjee and Moll (2009), fixed costs imply that savings behavior will generally be more complicated and potentially radically different at different wealth levels. There will be some threshold wealth level below which individuals prefer to dissave and simply eat up their wealth until it reaches zero. This also implies that the initial distribution of wealth matters. Such dynamic effects from the combination of fixed costs and financial frictions potentially decrease output and capital further. However, TFP will be as in the Proposition.

This result can best be understood by examining which entrepreneurs are active in an equilibrium with fixed costs and financing frictions. Figure 9 shows whether an entrepreneur with a given wealth/productivity combination is active by subdividing the (a, z) plane accordingly. The downward-sloping line $\underline{z}(a)$ is the threshold for being active. Richer entrepreneurs find it easier to cover the fixed cost and pass that threshold. Without fixed costs, the cutoff is just the vertical line labelled $\underline{z}$ (to see why the two lines intersect as drawn in the graph, recall from Proposition 5 that $\underline{z} > \underline{\underline{z}}$). As can be seen from the graph, some entrepreneurs that would be inactive in an economy without fixed costs, are now active. The reverse is also true. These are exactly the active incompetent-but-wealthy and the inactive talented-but-poor entrepreneurs described by Buera, Kaboski and Shin (2010): in short, a misallocation of talent. Another interesting feature of the economy with fixed costs is that it still aggregates to an aggregate production function that is Cobb-Douglas, despite the fact that individual production functions feature a non-convexity.

Proof of Proposition 5 Using similar steps as in the proof of Proposition 1, we have that individual output is $y(a, z) = (\pi/\alpha)\lambda z(a - \underline{a})$ and therefore aggregate output is

$$Y = \frac{\pi}{\alpha}\int_{\underline{a}}^{\infty}\int_{\underline{z}(a)}^{\infty}\lambda z(a-\underline{a})\varphi(a)\psi(z)dzda = \left(\int_{\underline{a}}^{\infty}\int_{\underline{z}(a)}^{\infty}\lambda z(a-\underline{a})\varphi(a)\psi(z)dzda\right)^{\alpha} L^{1-\alpha}.$$

The credit market clearing condition is

$$\int_{\underline{a}}^{\infty}\int_{\underline{z}(a)}^{\infty}\lambda a\varphi(a)\psi(z)dzda = K \equiv \int_{0}^{\infty} a\varphi(a)da. \tag{61}$$

TFP is therefore

$$Z = \frac{Y}{K^{\alpha}L^{1-\alpha}} = \left(\frac{\int_{\underline{a}}^{\infty}\int_{\underline{z}(a)}^{\infty}\lambda z(a-\underline{a})\varphi(a)\psi(z)dzda}{\int_{\underline{a}}^{\infty}\int_{\underline{z}(a)}^{\infty}\lambda a\varphi(a)\psi(z)dzda}\right)^{\alpha}.$$

With a Pareto distribution

$$Z = \left(\frac{\frac{\eta}{\eta-1}\int_{\underline{a}}^{\infty}\underline{z}(a)^{1-\eta}(a-\underline{a})\varphi(a)da}{\int_{\underline{a}}^{\infty}\underline{z}(a)^{-\eta}a\varphi(a)da}\right)^{\alpha} = \left(\frac{\eta}{\eta-1}\underline{\underline{z}}\right)^{\alpha} = \mathbb{E}[z|z \geq \underline{\underline{z}}]^{\alpha},$$

where the second equality uses that $\underline{z}(a) = \underline{\underline{z}}a/(a - \underline{a})$. This is the first part of the proposition. To see that the presence of fixed costs decreases the threshold $\underline{\underline{z}}$ (and thereby TFP), consider the market clearing condition (61). Again, using the Pareto distribution and that $\underline{z}(a) = \underline{\underline{z}}a/(a - \underline{a})$, the cutoff $\underline{\underline{z}}$ is defined by

$$\lambda\underline{\underline{z}}^{-\eta}\int_{\underline{a}}^{\infty}\left(\frac{a}{a-\underline{a}}\right)^{-\eta} a\varphi(a)da = \int_{0}^{\infty} a\varphi(a)da.$$

Without fixed costs the cutoff would be defined by $\lambda(1 - \Psi(\underline{z})) = \lambda\underline{z}^{-\eta} = 1$. Therefore, to conclude that $\underline{\underline{z}} < \underline{z}$, it suffices to show that

$$\int_{\underline{a}}^{\infty}\left(\frac{a}{a-\underline{a}}\right)^{-\eta} a\varphi(a)da < \int_{0}^{\infty} a\varphi(a)da.$$

This follows since with $\eta > 1$, $[a/(a - \underline{a})]^{-\eta} < 1$ for all a.$\square$

References

Aghion, Philippe, and Patrick Bolton. 1997. "A Theory of Trickle-Down Growth and Development." *The Review of Economic Studies*, 64: 151–172.

Angeletos, George-Marios. 2007. "Uninsured Idiosyncratic Investment Risk and Aggregate Saving." *Review of Economic Dynamics*, 10: 1–30.

Azariadis, Costas, and Leo Kaas. 2009. "Capital Misallocation and Aggregate Factor Productivity." Uni Konstanz Working Paper.

Baily, M., C. Hulten, and D. Campbell. 1992. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity*, 4: 187–267.

Banerjee, Abhijit V., and Andrew F. Newman. 1993. "Occupational Choice and the Process of Development." *The Journal of Political Economy*, 101(2): 274–298.

Banerjee, Abhijit V., and Andrew F. Newman. 2003. "Inequality, Growth and Trade Policy." MIT Working Paper.

Banerjee, Abhijit V., and Benjamin Moll. 2009. "Why Does Misallocation Persist?" *forthcoming, American Economic Journal: Macroeconomics.*

Banerjee, Abhijit V., and Esther Duflo. 2005. "Growth Theory through the Lens of Development Economics." In *Handbook of Economic Growth.* , ed. Philippe Aghion and Steven Durlauf.

Bartelsman, Eric J., and Phoebus J. Dhrymes. 1998. "Productivity Dynamics: U.S. Manufacturing Plants, 19721986." *Journal of Productivity Analysis*, 9: 5–34.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2008. "Cross Country Differences in Productivity: The Role of Allocative Efficiency." University of Maryland Working Paper.

Beck, Thorsten, A. Demirguc-Kunt, and Ross Levine. 2000. "A New Database on the Structure and Development of the Financial Sector." *World Bank Economic Review*, 14: 597605.

Bloom, Nicholas, and John Van Reenen. 2009. "Why Do Management Practices Differ across Firms and Countries?" *forthcoming, Journal of Economic Perspectives.*

Buera, Francisco J. 2009*a*. "A Dynamic Model of Entrepreneurship with Borrowing Constraints: Theory and Evidence." *Annals of Finance*, 5: 443464.

Buera, Francisco J. 2009*b*. "Persistency of Poverty, Financial Frictions, and Entrepreneurship." UCLA Working Paper.

Buera, Francisco J., and Benjamin Moll. 2010. "Do Financial Frictions Distort Aggregate Investment?" UCLA Work in Progess.

Buera, Francisco J., and Yongseok Shin. 2009. "Self-Insurance vs. Self-Financing: A Welfare Analysis of the Persistence of Shocks." Washington University St. Louis Working Paper.

Buera, Francisco J., and Yongseok Shin. 2010. "Financial Frictions and the Persistence of History: A Quantitative Exploration." UCLA Working Paper.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin. 2010. "Finance and Development: A Tale of Two Sectors." UCLA Working Paper.

Bulow, Jeremy, and Kenneth Rogoff. 1989. "Sovereign Debt: Is to Forgive to Forget?" *The American Economic Review*, 79(1): 43–50.

Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences." In *Handbook of Economic Growth.* Vol. 1 of *Handbook of Economic Growth*, , ed. Philippe Aghion and Steven Durlauf, Chapter 9, 679–741. Elsevier.

Caselli, Francesco, and Nicola Gennaioli. 2005. "Dynastic Management." London School of Economics Working Paper.

Cooper, Russell W., and John C. Haltiwanger. 2006. "On the Nature of Capital Adjustment Costs." *Review of Economic Studies*, 73(3): 611–633.

Cox, John C, Jr Ingersoll, Jonathan E, and Stephen A Ross. 1985. "A Theory of the Term Structure of Interest Rates." *Econometrica*, 53(2): 385–407.

Dixit, Avinash. 1993. *The Art of Smooth Pasting.* Fundamentals of Pure and Applied Economics 55, The Routledge.

Evans, David, and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy*, 97(4): 808–827.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98(1): 395–425.

Gabaix, Xavier. 2009. "Power Laws in Economics and Finance." *Annual Review of Economics*, 1(1): 255–293.

Galor, Oded, and Joseph Zeira. 1993. "Income Distribution and Macroeconomics." *The Review of Economic Studies*, 60(1): 35–52.

Gentry, William M, and R Glenn Hubbard. 2004. "Entrepreneurship and Household Saving." *Advances in Economic Analysis & Policy*, 4(1): 1–55.

Giné, Xavier, and Robert Townsend. 2004. "Evaluation of Financial Liberalization: a General Equilibrium Model with Constrained Occupation Choice." *Journal of Development Economics*, 74: 269–307.

Gourinchas, Pierre-Olivier, and Olivier Jeanne. 2006. "The Elusive Gains from International Financial Integration." *Review of Economic Studies*, 73(3): 715–741.

Gourio, Francois. 2008. "Estimating Firm-Level Risk." Boston University Working Paper.

Greenstreet, David. 2007. "Exploiting Sequential Learning to Estimate Establishment-Level Productivity Dynamics and Decision Rules." University of Oxford, Department of Economics Economics Series Working Papers 345.

Guerrieri, Veronica, Guido Lorenzoni, and Fabrizio Perri. 2009. "International Borrowing, Investment and Default." MIT Work in Progess.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *The Quarterly Journal of Economics*, 114(1): 83–116.

Hennessy, Christopher A., and Toni M. Whited. 2005. "Debt Dynamics." *Journal of Finance*, 60(3): 1129–1165.

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S Rosen. 1994. "Sticking It Out: Entrepreneurial Survival and Liquidity Constraints." *Journal of Political Economy*, 102(1): 53–75.

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

Hurwicz, Leonid. 1950. "Least-squares Bias in Time Series." In *Statistical Inference in Dynamic Economic Models.* , ed. T.C. Koopmans, 365–383. New York:Wiley.

Jeong, Hyeok, and Robert Townsend. 2007. "Sources of TFP Growth: Occupational Choice and Financial Deepening." *Economic Theory*, 32: 179–221.

Jeong, Hyeok, and Robert Townsend. 2008. "Growth and Inequality: Model Evaluation Based on an Estimation-Calibration Strategy." *Macroeconomic Dynamics*, , (12): 231284.

Kehoe, Timothy J., and David K. Levine. 2001. "Liquidity Constrained Markets versus Debt Constrained Markets." *Econometrica*, 69(3): 575–598.

King, Robert, and Sergio Rebelo. 1993. "Transitional Dynamics and Economic Growth in the Neoclassical Model." *The American Economic Review*, 83(4): 908–931.

Kiyotaki, Nobuhiro. 1998. "Credit and Business Cycles." *The Japanese Economic Review*, 49(1): 18–35.

Kiyotaki, Nobuhiro, and John Moore. 2008. "Liquidity, Bussiness Cycles and Monetary Policy." Princeton University Working Paper.

Klein, Lawrence R. 1960. "Entrepreneurial Saving." In *Consumption and Saving.* Vol. 2, , ed. Irwin Friend and Robert Jones. Philadelphia:University of Pennsylvania Press.

Klenow, Peter, and Andrés Rodríguez-Clare. 1997. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" In *NBER Macroeconomics Annual 1997, Volume 12.*

Kocherlakota, Narayana R. 2009. "Bursting Bubbles: Consequences and Cures." University of Minnesota Working Paper.

Krebs, Tom. 2003. "Human Capital Risk And Economic Growth." *The Quarterly Journal of Economics*, 118(2): 709–744.

Lagos, Ricardo. 2006. "A Model of TFP." *Review of Economic Studies*, 73(4): 983–1007.

Levine, Ross. 2005. "Finance and Growth: Theory and Evidence." In *Handbook of Economic Growth.* , ed. Philippe Aghion and Steven N. Durlauf. Elsevier B.V.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies*, 70(2): 317–341.

Liu, Lili. 1993. "Entry, Exit, and Learning in the Chilean Manufacturing Sector." *Journal of Development Economics*, 42: 217242.

Lucas, Robert E Jr. 1990. "Why Doesn't Capital Flow from Rich to Poor Countries?" *The American Economic Review*, 80(2): 92–96.

Luttmer, Erzo G. J. 2007. "Selection, Growth, and the Size Distribution of Firms." *The Quarterly Journal of Economics*, 122(3): 1103–1144.

Manuelli, Rodolfo E. 2009. "A General Equilibrium Model of International Credit Markets." Northwestern University Unpublished Manuscript.

Matsuyama, Kiminori. 2007. "Aggregate Implications of Credit Market Imperfections." *NBER Macroeconomics Annual.*

Midrigan, Virgiliu, and Daniel Yi Xu. 2009. "Finance and Misallocation." NYU Working Paper.

Nelson, Daniel B, and Krishna Ramaswamy. 1990. "Simple Binomial Processes as Diffusion Approximations in Financial Models." *Review of Financial Studies*, 3(3): 393–430.

Pawasutipaisit, Anan, and Robert M. Townsend. 2009. "Wealth Accumulation and Factors Accounting for Success." MIT Working Paper.

Petrin, Amil, and James Levinsohn. 2009. "Measuring Aggregate Productivity Growth Using Plant-level Data." University of Minnesota Working Papers.

Piketty, Thomas. 1997. "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing." *The Review of Economic Studies*, 64: 173–189.

Quadrini, Vincenzo. 2009. "Entrepreneurship in Macroeconomics." University of Southern California Working Papers.

Rajan, Raghuram G., and Luigi Zingales. 1998. "Financial Dependence and Growth." *The American Economic Review*, 88(3): 559–586.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Plants." *Review of Economic Dynamics*, 11(4): 707–720.

Samphantharak, Krislert, and Robert M. Townsend. 2009. "Households as Corporate Firms: Constructing Financial Statements from Integrated Household Surveys." Econometric Society Monograph Series.

Simon, Herbert A., and Charles P. Bonini. 1958. "The Size Distribution of Business Firms." *The American Economic Review*, 48(4): 607–617.

Solow, Robert M. 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics*, 70(1): 65–94.

Stokey, Nancy L. 2009. *The Economics of Inaction.* Princeton, NJ:Princeton University Press.

Stroock, D. W., and S. R. Srinivasa Varadhan. 1979. *Multidimensional Diffusion Processes.* Springer, Berlin.

Townsend, Robert M. 2009. "Financial Structure and Economic Welfare: Applied General Equilibrium Development Economics." *forthcoming, Annual Review of Economics.*

Tybout, James R., and Mark Roberts. 1996. *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity and Market Structure.* Oxford University Press.

Wong, Eugene. 1964. "The Construction of a Class of Stationary Markoff Processes." In *Stochastic Processes in Mathematical Physics and Engineering.* , ed. R. Bellman, 264–276. Providence, R.I.:American Mathematical Society.