



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Distortions, misallocation and the endogenous determination of the size of the financial sector

Citation for published version:

Bauer, C & Rodríguez Mora, JV 2020, 'Distortions, misallocation and the endogenous determination of the size of the financial sector', *The Economic Journal*, vol. 130, no. 625, pp. 24-49.
<https://doi.org/10.1093/ej/uez031>

Digital Object Identifier (DOI):

[10.1093/ej/uez031](https://doi.org/10.1093/ej/uez031)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

The Economic Journal

Publisher Rights Statement:

This is a pre-copyedited, author-produced version of an article accepted for publication in The Economic Journal following peer review. The version of record (Christian Bauer, José V Rodríguez Mora, "Distortions, Misallocation and the Endogenous Determination of the Size of the Financial Sector", The Economic Journal, Volume 130, Issue 625, January 2020, Pages 24–49) is available online at:
<https://academic.oup.com/ej/article/130/625/24/5524359>

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Distortions, Misallocation and the Endogenous Determination of the Size of the Financial Sector *

Christian Bauer
LMU Munich

José V. Rodríguez Mora
University of Edinburgh, the Alan Turing Institute and CEPR

July 24, 2018

Abstract

We present a model of heterogeneous firms and misallocation where financial frictions are partially overcome if more human resources are devoted to intermediation, at the cost of having less resources employed in directly productive activities. Not only an inefficient financial sector results in an inefficient final good sector, but also an inefficient final good sector results in an inefficient financial sector. Exogenous inefficiencies in the productive sector generate decreased demand for financial services, translating in a smaller and less efficient financial sector, worsening the resource allocation in the productive sector. This direction of causality seems in line with cross-country evidence.

JEL Classification E0, G0, L11, L16

Keywords: Misallocation, Credit search frictions, interaction of product and credit market efficiency..

*This paper was previously circulated (in October 2013) under the title “*The Joint Determination of TFP and Financial Sector Size*”. We thank the Morten O. Ravn and two anonymous referees for very useful and constructive comments.

1 Introduction

Motivation

Although it is not controversial to argue that more efficient financial markets improve the allocation of resources and result in a more efficient final good sector, the issue of reverse causality is seldom studied. Nevertheless, it is quite intuitive that a more efficient final good sector may create incentives to endogenously develop a large financial sector. This may in itself improve the allocation of resources and foster the efficiency of the final good sector even further.

Building on this intuition, this paper contributes to the theoretical literature on misallocation of resources across heterogeneous firms by allowing the degree of capital market imperfections to be endogenously determined. It also examines the empirical cross-country correlation between financial sector size and per capita income under a new light. Our model explains this correlation not because rich countries have inherently more efficient financial sectors, but because they have inherently more efficient final good sectors. This generates the demand to devote a large amount of human resources to intermediation, which itself generates a more efficient firm distribution.

A substantial body of evidence shows that resource misallocation partly explains the differences in total factor productivity (TFP) between rich and poor countries. In poorer countries, inefficient firms use excessive resources at the expense of efficient ones (see, for instance, Hsieh and Klenow (2009), Alfaro, Charlton, and Kanczuk (2008), Bartelsman, Haltiwanger, and Scarpetta (2013), Guner, Ventura, and Xu (2008), Restuccia and Rogerson (2008), Bento and Restuccia (2017) or Bloom and Reenen (2007)). In the large body of literature¹ that builds on Hopenhayn (1992) and Lucas (1978), the explanation for this misallocation is intimately tied to the existence of capital market imperfections that prevent resources from being moved from where they are to where they are needed. Our model builds on this tradition while allowing the degree of capital market imperfections to be endogenously determined.

In our model, society chooses to use human resources toward either financial intermediation or direct production. The trade-off appears because resources devoted to intermediation, albeit not producing output, do facilitate the acquisition of capital by entrepreneurs. Thus, the amount of frictions in the financial sector is determined endogenously by devoting more or fewer resources toward intermediation. Devoting more resources to intermediation (fewer financial frictions) implies that it is less expensive to look for better projects and give up mediocre ones, thereby improving the distribution of the firm's qualities in the productive sector.

We consider 3 exogenous sources that could be conducive to differences in the size of the financial sector across countries: (1) the degree of capital abundance, (2) measures of the *inherent* efficiency of the financial sector and (3) measures of the *inherent* efficiency of the productive sector. The equilibrium degree of efficiency in each of the sectors is a convolution of the inherent (exogenous) efficiency and the decisions on the allocation of resources between sectors that society

¹See Restuccia and Rogerson (2013) and Restuccia and Rogerson (2017) for up to date reviews of the literature.

chooses to make.

The relative abundance of capital turns out to be irrelevant for determining the degree of misallocation in the economy. It affects neither the usage of human resources in financial or productive activities nor the decisions about which firms will be financed (and thus the distribution of the quality of firms).

As in all previous literature, in our model exogenous frictions in the credit market produce a less efficient final good sector by affecting firm selection. The less efficient the financial sector is, the greater its size, since more resources need to be devoted into it.

We also look in a opposite direction of causality. We believe this is a novel and interesting approach: a inherently better (i.e., more efficient and competitive) final good sector results in society devoting more resources to intermediation. Thus, an efficient final good sector results in a larger financial sector and, via this mechanism, in larger and more efficient firms.

In the data, rich countries have larger financial sectors. In the framework of our model this cannot be because rich countries have *ceteris paribus* more capital, because this does not affect financial sector size. Likewise, this correlation cannot be a consequence of richer countries enjoying *inherently* more efficient financial sectors, because in this case their financial sectors would be *smaller*. The only explanation that fits both the data and the model is that the bulk of the differences in the exogenous component of efficiency across countries lies in the productive sector, not in the financial one. Reading the data at the light of the model suggests that rich countries are richer mainly because they have more efficient productive sectors, which in itself endogenously produces larger financial sectors, which generates a distribution of firm productivities with a higher mean and smaller variance.

Implementation and Related Literature

We differentiate the workings of capital markets for deposits from how they work in an investment context. Capital, being a homogeneous good, demands a homogeneous rental price. In our framework, a Walrasian market for deposits serves this purpose. Investments, on the other hand, are heterogeneous in their expected returns. The different rents of capital obtained by investing in heterogeneous projects should be attributed not to capital ownership, but to the ability and effort devoted to finding and identifying quality investments.² We deem as brokers the individuals who endeavor to undertake such an activity. For the broker and the entrepreneur (the original owner of the investment opportunity) finding each other is a time-consuming activity, so we model this activity using search theory.³

² In certain occasions, the owners of capital may well coincide with the brokers, as is the case, for example, for some venture capitalists. This case is thoroughly studied by Silvera and Wright (2007).

³Blanchflower and Oswald (1998) report that raising funds is a principal obstacle to potential entrepreneurs. Further empirical support for this claim is provided by Evans and Jovanovic (1989), Evans and Leighton (1989), Holtz-Eakin, Joulfaian, and Rosen (1994), Gentry and Hubbard (2000), and Guiso, Sapienza, and Zingales (2004).

Although we use search theory to model the degree of imperfections in the financial sector, we believe that our point is more general. Results qualitatively similar to ours would be obtained in any context in which the degree of inefficiencies in the financial sector were a decreasing function of the amount of resources devoted to it. Search seems the more natural way to model this relationship.

Recent years have seen a renewed and increasing interest in the implementation of search environments for modeling the non-Walrasian features of the credit and investment markets (Wasmer and Weil (2004), Duffie, Garleânu, and Pedersen (2005), Silvera and Wright (2010), Besci, Li, and Wang (2005), Haan, Ramey, and Watson (2003), Dell’Ariccia and Garibaldi (2005)).⁴ Following Diamond (1990), this literature highlights, in an encompassing manner, the quantitative importance of information frictions, time usage, and the positive value of establishing creditor-borrower relationships. Implications of these models fit the data well (Dell’Ariccia and Garibaldi (2005), Petersen and Rajan (2002)). In addition, Cui and Radde (2016) present a model where search frictions in the assets markets determine the degree of value of liquid assets.

We differ from the existing literature in that (i) we allow for the aforementioned aggregate resource constraints on capital and human resources, and (ii) we study how final good sector efficiency leads to greater financial sector size and, via this mechanism, to allocative efficiency. We believe that by doing so, we improve on the traditional Mortensen and Pissarides (1994) framework that, when applied to financial intermediation issues, demands an infinitely elastic supply of both firms and capital. In particular, in our context society endogenously determines the severity of frictions. That is, by devoting more human resources into intermediation (i.e., sacrificing them from directly productive activities), finding capital becomes less of an obstacle for entrepreneurs.

A small number of papers study the endogenous share of resources devoted to financial intermediation (such as Philippon (2008) and Buera, Kaboski, and Shin. (2011)). Nevertheless, these papers do not study the bidirectional link between efficiency in product and credit markets.

Our argument is closely related to the misallocation literature. Although misallocation distortions may stem from a wide array of distorted prices faced by individual producers (as in Restuccia and Rogerson (2008)), we highlight credit frictions for prospective entrepreneurs as a specific source of allocative inefficiency. Recognizing credit frictions as a barrier to entry may potentially be useful in guiding empirical work, which depends on inferring misallocation distortions from existing firms (by measuring the residuals in first-order conditions; see Chari, Kehoe, and McGrattan (2007)). Notably, the credit search approach may further be useful in this context to interpret heterogeneity in financing costs because the process of search and bargaining implies that individual companies may face different financing costs, but that these differences need not be thought of as being caused by non-economic factors. The approach may also be useful in theoretical applications where fixed costs play an important allocative role (e.g., Melitz (2003)-type models). Here, these costs are

⁴Unpublished papers include Dell’Ariccia and Garibaldi (1998), Besci, Li, and Wang (2009), Silvera and Wright (2007), and Chiu and Koepl (2011).

explicitly costs for obtaining credit; however, they are endogenous and respond to changes in the environment. Our view is consistent with the broader view that the credit sector, by mobilizing savings, allocating resources, and screening projects, plays a crucial role in shaping the development process of new products, firms, and sectors (Greenwood and Jovanovic (1990), Levine (1997), and Matsuyama (2008)). Midrigan and Xu (2014) show that financial frictions affect the distribution of TFP across firms. Greenwood, Sanchez, and Wang (2010) show how the informational structure of financial firms (how much do they know about creditors) affect missallocation and productive activity, and Greenwood, Sanchez, and Wang (2013) quantify the effect of financial development on missallocation and development.

We contribute to the analysis of aggregate consequences of financial frictions, surveyed by Matsuyama (2008). Li and Sarte (2003) provide evidence that changes in intermediation costs directly affect output. Cooley, Marimon, and Quadrini (2004) show that limited financial contract enforceability amplifies the impact of technological innovations on aggregate output.⁵ More recently, Russ and Valderrama (2009) exploit the relative costs of bank and bond financing to explain how bank lending frictions may affect the firm size distribution through intra-industry reallocations. Here, we instead exploit the value of creditor–borrower relationships.

Finally, we link credit market frictions to recent empirical findings on greater within-industry productivity dispersion and lower aggregate TFP in poorer countries.⁶ These distortions matter: Hsieh and Klenow (2009) conclude that manufacturing TFP would increase by 30-50% in China and 40-60% in India if these gaps were reduced to the observed levels in the United States.⁷ Alfaro, Charlton, and Kanczuk (2008) match each country’s plant size distribution by an appropriate profile of output taxes and subsidies. They suggest interpreting these distortions, among other things, as favorable interest rates on loans based on non-economic factors.

Our paper is organized as follows. Section 2 describes the basic model and solves it for an exogenous capital supply. Sections 3 and 4 analyze the equilibrium conditions and characterization. Section 5 looks at the joint determination of TFP and financial sector size. Finally, section 6 concludes. The appendix provides with several examples with closed form solutions and looks at the results in a growth context.

2 General Environment

The economy is populated by a mass one of agents who live in continuous time and die at an exogenous rate δ . There is a flow of new arrivals that keep the population constant. Agents are

⁵Cagetti and De Nardi (2006) and Quadrini (2000) show that models with financial frictions and entrepreneurship may explain observed distributions of wealth.

⁶Hsieh and Klenow (2009) measure the degree of misallocation by the size of gaps in marginal products of labor and capital across plants within narrowly defined industries.

⁷Cf. Restuccia and Rogerson (2008), Alfaro, Charlton, and Kanczuk (2008), and Bartelsman, Haltiwanger, and Scarpetta (2013).

all identical and at any moment they choose to be either entrepreneurs or brokers.

All production takes place within firms that use only capital and entrepreneurial activity as inputs.⁸ Although all agents (and thus, all entrepreneurs) are ex ante identical, projects differ in their productivities. Let a be an indicator of the productivity of projects, which is drawn from a distribution $G(a)$ (details below). A project with productivity a that uses an amount of capital k generates a stream of income $F(k, a, Y)$, with $\frac{\partial F(k, a, Y)}{\partial k} > 0$, $\frac{\partial F(k, a, Y)}{\partial a} > 0$, $\frac{\partial^2 F(k, a, Y)}{\partial a \partial k}$ and $\frac{\partial^2 F(k, a, Y)}{\partial k \partial k} < 0$.

Y is a measure of aggregate demand or market size. We include it for generality. Notice that we would have a neoclassical model by imposing $\frac{\partial F(k, a, Y)}{\partial Y} = 0$, which is perfectly possible in our context. We allow for a much more general formulation that includes models with monopolistic competition, a common environment in many models with heterogeneous firms.

Each unit of capital will obtain a flow rent r (to be determined in equilibrium) when used in firms. The flow of profits generated by a project is then $\pi(a, r, Y) = \max_k \{F(k, a, Y) - rk\}$, and the capital demanded by a project of productivity a if the return to capital is r is $k^d(a, r)$.

Furthermore, we will assume that $F(k, a, Y)$ is log linear in k , a , and Y . Thus, profits are a log linear function of r , a , and Y , and the ratio of capital income to profits is a constant:

$$\pi(a, r, Y) = (1 - e_k) e_k^{\frac{e_k}{1-e_k}} a^{\frac{e_a}{1-e_k}} r^{-\frac{e_k}{1-e_k}} Y^{\frac{e_y}{1-e_k}} \quad (1)$$

$$\frac{rk^d(a, r, Y)}{\pi(a, r, Y)} = \frac{e_k}{1 - e_k} \quad (2)$$

where e_k , e_a , and e_y are the (constant) elasticities of $F()$ with respect to k , a , and aggregate demand, respectively, and $e_k \in [0, 1]$, $e_y \in [0, 1]$, and $e_a \in \mathfrak{R}^+$. The elasticity of profits to a is an important parameter of our story. In section 5.2 we show that it is linked to efficiency in the product market. For convenience we define $\epsilon = \left(\frac{e_a}{1-e_k}\right)$

Examples of log-linear $F(k; a, Y)$ are as follows:

(1) A Dixit-Stiglitz economy with constant productivity, where $F(k; a, Y) = s \left(\frac{a}{r}\right)^{\sigma-1} Y$ (in this case, ϵ would be the elasticity of substitution between any pair of horizontally differentiated consumer varieties, s would be equal to $(\sigma - 1)^{\sigma-1} \sigma^{-\sigma}$, and a would be the productivity of the individual firm).

(2) A competitive economy with decreasing returns, where $F(k; a) = (ak)^\alpha$ (where a would identify the firms' productivity).

A project generates income while the entrepreneur is alive and disappears when she dies. The capital does not disappear upon the death of the entrepreneur.

Let \bar{k} be the total amount of capital that is (inelastically) supplied by capital owners who, for simplicity, are left out of the model.⁹

⁸It would be trivial to include labor as an additional input, but it would not add further insights.

⁹As an extension, in appendix B we allow k to be accumulated.

2.1 Deposit and Investment Markets

If capital markets were Walrasian, all firms would produce at the maximum possible productivity, all individuals would be entrepreneurs, and the interest rate would be determined to equalize aggregate capital demand to \bar{k} . We depart from this paradigm by introducing search frictions in the allocation of capital to projects. We do so by dividing the capital market into two separate markets, interconnected by brokerage services.¹⁰

On the one hand, there is a Walrasian *deposit market* in which capital is offered by its owners in exchange for a market return r . There is also a *investment market* in which entrepreneurs obtain capital. These markets can be thought as two different rooms that differ along two dimensions.

(1) The rooms have different actors, since entrepreneurs are barred from entering the deposit room, whereas capital owners are barred from the investment room. There is a third type of agents (called brokers) that breach this gap between the markets. Being capable of acting in both rooms, they demand capital in the deposit room and supply capital in the investment room.

(2) The deposit market is frictionless, whereas the investment market is characterized by search frictions.

Thus, in the *deposit room* all agents have instantaneous access to all other participants. This set up ensures that capital will be paid a Walrasian price r per unit of time. The *investment room* is considerably more interesting. It has frictions and, consequently, finding a partner takes time. Once agents meet each other, they need to bargain an allocation of the total surplus.

In this way, we try to capture the dual nature of capital markets. On the one hand, financial capital is a rather homogeneous good from the owner's point of view. Thus, the return of capital cannot differ between them. On the other hand, however, the return of capital might vary in different investment possibilities, and the owners of these investment possibilities do not have immediate access to capital. All capital is used all of the time, but not all investment opportunities are used at any given moment of time. Their owners invest time looking for capital, and a certain number of individuals invest their time in connecting them with the capital left free by projects that died away.

To assume that the investment market is subject to frictions is a metaphor of either the slow search process for partners who are able to attest to the validity of the investment project or the possibility of finding contractual solutions to informational asymmetries or moral hazard problems. As a matter of fact, entrepreneurs spend time looking for finance, and there is a sector of the economy whose task is to intermediate between owners and users of capital.

The extra income that capital generates in a high productivity project (versus a lower productivity one) should not be attributed to a capital rent. It should be attributed to (1) the talent for generating the high productivity, (2) the talent for recognizing that its productivity is high, and (3) the luck involved in being there and being the only one able to access the capital that the en-

¹⁰Another recent paper that divides the capital market between investment and deposit markets, but in a completely different context, is Monnet and Sanches (2015).

trepreneur needs. In our case, brokers obtain rents for intermediating and allowing capital to flow to its uses, and thereby find themselves in a bilateral monopoly position vis-à-vis the entrepreneur in the match.

Thus, society determines the degree of frictions in the investment market by devoting more or fewer resources to intermediation: the larger the number of individuals who opt to become brokers, the faster the existing entrepreneurs will access capital. Our paper focuses on determining how many resources are used in intermediation versus directly in production. We emphasize the human resource constraint that society faces: more resources used for intermediation imply fewer resources in production, but they facilitate production by bringing capital closer to its productive use.

2.2 Agents, Matching and Bargaining

Agents are ex ante identical but choose to be either entrepreneurs or brokers.

Entrepreneurs have investment projects that transform capital into output, but they have neither capital nor access to the deposit room. They obtain capital from brokers in the investment room.

Brokers have no direct role in production, but they have access to both the deposit room (where capital lies) and the investment room (where entrepreneurs look for it).

Thus, the two roles are necessary and complementary in production: entrepreneurs for directly producing, brokers for providing them with access to capital.

We assume that changing professions is costless (thus, imposing that the value must be equal for both professions). Note that this does not mean that the agents can escape from the existence of frictions. If an entrepreneur becomes a broker, she does it without a project. If a broker becomes an entrepreneur, he does it without capital. Two individuals (one of each type) are needed in order to start production. That is what our notion of financial frictions is all about.

The speed at which these agents meet is determined by a matching function with constant returns to scale. The tightness of the capital market (θ) is defined as the ratio of the mass of entrepreneurs who are *searching* for capital to the mass of brokers. As we will see below, brokers always search, but entrepreneurs leave the investment room permanently once they find capital. Thus, the numerator of θ is smaller than the mass of entrepreneurs, whereas the denominator equals the mass of brokers.

Let $p(\theta, \nu)$ denote the function determining the rate at which entrepreneurs meet brokers. The function is decreasing in θ . The parameter ν is a shift parameter denoting the general efficiency of the matching process, with $\frac{\partial p(\theta, \nu)}{\partial \nu} > 0$ and $\lim_{\nu \rightarrow \infty} p(\theta, \nu) = \infty$. A Walrasian investment market would be characterized by $\nu \rightarrow \infty$. The rate at which brokers meet entrepreneurs is $\theta p(\theta, \nu)$, which grows with tightness.

2.2.1 Entrepreneurs

Entrepreneurs are in one of two states.

- (1) They are looking for capital, that is, searching for a broker.
- (2) They are producing in a firm created with the capital that they obtained from a broker.

While entrepreneurs are searching for a broker they have no income. While they are producing, they have income equal to their share of $\pi(a, r, Y)$. This share is determined in a bargaining process to be determined below.

Let a denote the productivity of the investment project, which is drawn from a distribution $G(a)$, common to all agents. We make two simplifying assumptions:

- (1) Entrepreneurs do not need to look for projects; they find them instantaneously.
- (2) Entrepreneurs are unaware of the value of a before they meet the broker, who informs them of its value upon meeting.

Alternatively, we could have assumed that agents need to spend time looking for a project and that when they find one, they receive a signal of its value. This assumption would substantially increase the number of states without adding further insight. Here, we want to capture the idea that bad projects in search of capital produce congestion that has a negative effect on good projects. Having completely uninformed entrepreneurs is the simplest way of modeling this idea.

Once the entrepreneur and the broker meet, assuming they decide to go ahead with the project, the entrepreneur will spend the rest of her life producing. She will get a constant flow of income until the moment of her death, at which time the capital goes back to its owners.

We assume that if the entrepreneur and the broker do not agree on a certain division of the output, both the project and the match get destroyed. That is, not to arrive to an agreement has two consequences:

- (1) The entrepreneur cannot draw another project instantaneously. He needs to search for another partner (i.e., the *match* is destroyed)
- (2) The entrepreneur cannot look for another broker with knowledge of the value of a (i.e., the *project* is destroyed).

Match destruction is consubstantial to having a proper search environment, whereas project destruction simplifies bargaining considerably.

Notice that an obvious consequence of the previous assumptions is that not all values of a are acceptable for production. An entrepreneur that goes ahead with a project is investing the rest of her life in it. Engaging in production means giving up the option value of finding a more productive project. We will show that there exists a unique equilibrium threshold of productivities b such that the project is financed if and only if $a \geq b$.

2.2.2 Brokers

Brokers can have relationships with multiple entrepreneurs without satiation. Their life is as follow. They search in the investment room for investment projects, which they find at a rate $\theta p(\theta, \nu)$. Once they find a partner they bargain over the share of the lifetime output generated by the project. We can think of this as a once-and-for-all payment. At this point, there is no further involvement of the broker in the business of this particular entrepreneur. After the transaction is over, the broker looks for another partner. The continuation value of the broker is no different from what it was before the meeting; they *always* search. They do not care about how many transactions they have completed because they are risk neutral and only look to the future.

2.2.3 Bellman Equations

We assign V_0 and $V_1(a)$ to the values of an entrepreneur looking for capital and one that is producing in a project with productivity a , respectively.

Entrepreneurs looking for a broker have no income, and do not know the value of a . They meet brokers at a rate $p(\theta, \nu)$, and the minimum productivity at which projects are viable is denoted by b . Thus, their value function is simply

$$\delta V_0 = p(\theta, \nu) \int_b^\infty [V_1(a) - V_0] dG(a). \quad (3)$$

At the production stage, entrepreneurs produce, sell, and pay annuities to the broker. Thus, their value is determined by

$$\delta V_1(a, r, Y) = \pi(a, r, Y) - \rho(a, r, Y), \quad (4)$$

where $\rho(a, r)$ is the payment annuity made to the broker (to be determined in equilibrium).

Putting together equations (3) and (4), we have that

$$\delta V_0 = \frac{p(\theta, \nu) [1 - G(b)]}{\delta + p(\theta, \nu) [1 - G(b)]} \int_b^\infty [\pi(a, r, Y) - \rho(a, r, Y)] \frac{dG(a)}{1 - G(b)}. \quad (5)$$

The broker gets paid whenever she meets an entrepreneur with high enough productivity, and he gets paid the PDV of the flow values discounted by the risk of death of the entrepreneur in a perfectly working annuity market.

Thus, the continuation value of being a broker (B) solves $\delta B = \theta p(\theta, \nu) \int_b^\infty \Gamma(a) dG(a)$, with $\Gamma(a) = \frac{\rho(a, r, Y)}{\delta}$.

2.3 Bargaining

Matched entrepreneurs and brokers are in a bilateral monopoly. We assume that they split the income from the project according to Nash bargaining.¹¹ Thus, agents agree on choosing the amount of capital that maximizes the profits generated by the project. The entrepreneur's bargaining weight is $\beta \in (0, 1)$.

The outside options of the broker and the entrepreneur are very different. For the broker the outside option is zero, because she has no satiation with respect to the number of entrepreneurs she can serve, and her continuation value is constant. The entrepreneur, on the other hand, commits his life to their project. Thus, what the entrepreneur gives up by accepting a proposal is the possibility of searching for another project, V_0 .

Denoting by $S(a)$ the total surplus generated by the match, the payments ρ supported by Nash bargaining are such that $\beta S(a) = V_1(a) - V_0$ and $(1 - \beta) S(a) = \Gamma(a)$, giving as payment

$$\rho(a, r, Y) = (1 - \beta) \{ \pi(a, r, Y) - \delta V_0 \}, \quad (6)$$

with the value of an entrepreneur without a match being

$$\delta V_0 = \frac{p(\theta, \nu) [1 - G(b)]}{\delta + p(\theta, \nu) [1 - G(b)]} \times \frac{\frac{\beta}{1-\beta}}{\frac{\beta}{1-\beta} + \frac{\delta}{\delta + p(\theta, \nu) [1 - G(b)]}} \times \int_b^\infty \pi(a, r, Y) \frac{dG(b)}{1 - G(b)}. \quad (7)$$

Equation (7) can be interpreted as follows. The third expression is the average profit generated by a project that gets financed. The second one is the share of the entrepreneur (which is increasing with β and decreasing with the time she spends searching). Finally, the first component determines the (actuarially fair) present discounted value of that income.

The value of a broker can be rewritten (with equivalent interpretation) as:

$$\delta B = \frac{\theta p(\theta, \nu) [1 - G(b)]}{\delta + \theta p(\theta, \nu) [1 - G(b)]} \frac{\frac{1-\beta}{\beta}}{\frac{1-\beta}{\beta} + \frac{\delta}{\delta + \theta p(\theta, \nu) [1 - G(b)]}} \times \int_b^\infty \pi(a, r, Y) \frac{dG(b)}{1 - G(b)}. \quad (8)$$

Finally, the value of an entrepreneur within a match can be re-written as

$$\delta V_1(a) = \delta V_0 + \beta [\pi(a, r, Y) - \delta V_0]. \quad (9)$$

¹¹In appendix A.2 we show that there exists a unique value of β that maximizes income. It is the bargaining power that satisfies the Hosios condition in this context. Our aim in this paper is not to focus on contractual issues. Thus, we can think of β as fixed at the level that satisfies the Hosios condition. Actually, assuming competitive search instead of Nash bargaining yields identical quantitative results.

3 Equilibrium Conditions

Define m as the number of entrepreneurs in the economy. Then, the set of endogenous variables is $\{\theta, m, r, b, Y\}$. Their value has to be such that the following conditions hold.

3.1 Human Resource Constraint: Size of Financial sector

All agents have to be employed, either as brokers (in the denominator of θ), as entrepreneurs without a match (in the numerator of θ), or as entrepreneurs running their business. It is straightforward to prove that in steady state, the share of entrepreneurs who are *not* in a match (and are thus searching) is $\frac{\delta}{\delta + p(\theta, \nu)[1 - G(b)]}$.

Thus, tightness in the investment market (i.e., the ratio of searching entrepreneurs to brokers) is determined both by the number of brokers (i.e., the amount of resources that society devotes to financial activities) and by the threshold of productivity (which would result in more or less rejected projects and a longer or shorter expected search period for entrepreneurs). Specifically, $\theta = \frac{\delta}{\delta + p(\theta)[1 - G(b)]} \frac{m}{1 - m}$, which can be rewritten as

$$1 - m = \frac{\delta}{\theta [\delta + p(\theta, \nu) (1 - G(b))] + \delta} \quad (10)$$

Notice that for any fixed level of tightness, this establishes a positive relationship between the productivity threshold and the size of the financial sector. This positive relationship is the contribution of the financial sector to the workings of the economy. A larger financial sector allows to have more efficient firms, given the degree of liquidity. The reason is as follows. An increase in b implies that the number of rejections would also increase, which means that the share of *searching* entrepreneurs would also increase, which in turn demands an increase in the size of the financial sector in order to keep θ constant.

Thus, a larger financial sector (meaning more a financial sector with more human resources, not necessarily with more capital) enables society to be more selective in the quality of the projects it finances. Good projects are financed faster, and allows people with bad projects to look for better ones. Although the human resources devoted to finance do not produce output directly, they enhance the productivity of firms by reducing the opportunity cost of searching for a better projects.

3.2 No Arbitrage between Professions

Agents, being ex ante homogeneous, need to be indifferent between professions. Otherwise, there would be a possibility of arbitrage. Thus, in equilibrium

$$\delta V_0 = \delta B. \quad (11)$$

This condition has two important consequences:

(1) First, from equations (7) and (8), it follows that the value of both entrepreneurs and brokers is proportional to the expected profit of firms. Thus, profits disappear from equation (11). In order to equalize the values, it does not matter how much is produced on average. What matters is the PDV of the share of production that brokers and entrepreneurs obtain.

(2) Furthermore, θ depends only on the bargaining power of each side. It is immediate to show that arbitrage implies that market tightness equals the ratio of bargaining powers:

$$\theta = \frac{\beta}{1 - \beta}. \quad (12)$$

The more bargaining power entrepreneurs have (β), the more attractive entrepreneurial activities are, and therefore longer search periods without production (i.e., less liquidity) are necessary to equalize value across activities. The decrease in liquidity comes about as a consequence of an increase in the ratio of searching entrepreneurs to brokers.

This expression for the liquidity in investment markets is akin to the one encountered by Wasmer and Weil (2004) in an environment with homogeneous projects. Notice that for bargaining purposes our firms *are also homogeneous*. That is, we assume that if no agreement is reached, it is as though they had never met. The value of the current a has no effect on the outside possibility of any agent. Thus, only the expected future values matter, and they matter in the same manner for both agents. The shares that they obtain from the surplus adjust to make both professions equally attractive. As a consequence, θ does not depend on the degree of frictions in the credit market nor on inefficiencies in the product market, which simplifies the analysis enormously.¹²

3.3 Individually Optimal Search Rule: Productivity Threshold

The productivity threshold b is such that the surplus of a match with this productivity equals zero: $S(b) = 0$. Any match with productivity lower than b would not be worth it, and any match with more than zero would generate a surplus.

Given that the continuation value of the broker is independent of events in this match, the surplus is zero when the entrepreneur is indifferent between going ahead with production or returning to the searching pool. That is,

$$\delta V_1(b) = \delta V_0. \quad (13)$$

Using equation (9), it is clear that projects will be accepted if and only if the stream of profits that they generate is larger than the value of going back into search. From there it follows that b

¹² If agents were not homogeneous, the scenario would be slightly different. In general, θ can be a function of b , but it is *always* independent of expected profit due to the reasons expressed above.

is such that $\pi(b, r, Y) = \delta V_0$. Thus, we can rewrite 13 as

$$\frac{\pi(b, r, Y)}{\int_b^\infty \pi(a, r, Y) \frac{dG(b)}{1-G(b)}} = \frac{p(\theta, \nu) [1 - G(b)]}{\delta + p(\theta, \nu) [1 - G(b)]} \times \frac{\frac{\beta}{1-\beta}}{\frac{\beta}{1-\beta} + \frac{\delta}{\delta + p(\theta, \nu) [1 - G(b)]}}. \quad (14)$$

The right hand side of equation (14) lies in the interval $[0, 1]$ and equals the present discounted value of the share of income that goes to the entrepreneurs. It is increasing in the rate at which they get finance for suitable projects because (i) this brings forward the rewards and (ii) it increases their share by improving their outside option. Thus, it is decreasing in the value of the threshold b and equals zero as it approaches its upper limit.

The left hand side of (14) is the ratio of marginal to average profits. This is a function of b , which is independent of both r and Y . We will deem such a function as $H(a, \epsilon)$:

$$H(b, \epsilon) \equiv \frac{\pi(b, r, Y)}{\int_b^\infty \pi(a, r, Y) \frac{dG(a)}{1-G(b)}} = \frac{(b)^\epsilon}{\int_b^\infty (a)^\epsilon \frac{dG(a)}{1-G(b)}} \in (0, 1), \quad (15)$$

where ϵ is the elasticity of profits with respect to a : $\epsilon = \frac{e_a}{1-e_k}$.

It is intuitive to expect $H(b, \epsilon)$ to be a non-decreasing function of b . If a were bounded by above at value \bar{a} , it is certain that the ratio would not be decreasing in the neighborhood of \bar{a} . Its maximum value is 1, which is achieved as a approaches the limit. Thus, we make the following assumption.

Assumption 1 *The distribution of a is such that the ratio of marginal to average profit (H) is a non-decreasing function of b : $\frac{\partial H(b, \epsilon)}{\partial b} \geq 0$.*

Notice that assumption 1 is a restriction on the distribution of productivities. It is by no means a stringent assumption, it holds for many (if not all) of the commonly used distributions. It holds, for instance, if a is bounded and uniformly distributed. More interestingly, if a is distributed via a Pareto distribution, this takes the peculiar form of $H(b, \epsilon)$ being independent of b , which facilitates algebra enormously. From now on, we will always assume that assumption 1 holds.

3.4 Capital Market Clearing and Determination of r

The Walrasian nature of the deposit market ensures that all capital is used and receives the same remuneration r , which is fixed by the Walrasian auctioneer of the deposit room in order to equal capital supply \bar{k} with capital demand. Demand comes from all active firms. Each of them demands according to productivities $k^d(a, r)$. Thus,

$$\frac{p(\theta) [1 - G(b)]}{\delta + p(\theta) [1 - G(b)]} m \int_b^\infty k^d(a, r) \frac{dG(a)}{1 - G(b)} = \bar{k},$$

where the left-hand side is capital demand: the sum of all capital demanded by projects that are in production (m times the percentage of projects that find capital times their average capital demand). Multiplying both sides by r , and using equation (2), we can rewrite this equation as

$$\frac{p(\theta)[1-G(b)]}{\delta+p(\theta)[1-G(b)]}m\int_b^\infty\frac{e_k}{1-e_k}\pi(a,r,Y)\frac{dG(a)}{1-G(b)}=r\bar{k}. \quad (16)$$

3.5 Output Determination

The determination of output (which, because \bar{k} is constant, equals aggregate consumption expenditures) is as follows. First, output equals the summation of capital income ($r\bar{k}$) and the income generated by the agents. Second, agents discount the future at the death rate, and are equal at birth. Thus, it follows that the average income across all agents needs to equal the expected present discount lifetime income of any of them.

After some manipulation it is easy to see that the value at birth for any agent is

$$\delta V_0 = \frac{\beta p(\theta, \nu)[1-G(b)]}{\delta + \beta p(\theta, \nu)[1-G(b)]} \int_b^\infty \pi(a, r) \frac{dG(a)}{1-G(b)}. \quad (17)$$

Notice that this can be simplified, because b is determined so that average lifetime income equals the profit annuity of the marginal firm: $\delta V_0 = \pi(b, r, Y)$ (see equation (13)).

It follows that GDP equals capital income plus the flow of profits:

$$Y = r\bar{k} + \frac{1}{\delta}\pi(b, r, Y) = r\bar{k} + V_0. \quad (18)$$

4 Equilibrium Characterization and Solution

An equilibrium is a vector $\{\theta, m, r, b, Y\}$ such that **(1)** all labor is used (equation (10)), **(2)** there is no arbitrage possibility between professions ($V_0 = B$), **(3)** the threshold of productivity is chosen to be constrained efficient ($V(b) = V_0$), **(4)** all capital is used, and **(5)** income equals output.

The solution algorithm is extremely simple. As we have seen, arbitrage between professions pins down θ . Given θ , b is determined via equation (14). The size of the financial sector ($1 - m$) is then obtained from the human resource constraint. Finally, r and Y are obtained as residuals equations (16) and (18). These steps yield the following result:¹³

¹³ The only point at which the proof might not be obvious is the solution of (19). As a function of b , the RHS is monotonously decreasing, valued in $(0, 1)$, and it approaches 0 as b approaches its upper limit (or infinity). $H(b, \epsilon)$ approaches 1 as b approaches its upper limit (or infinity). So, there must be at least one solution. If $H(b, \epsilon)$ is monotonously non-decreasing, the solution is obviously unique.

Actually, in order to ensure that the two lines cross, we need to assume that when a is valued at its lowest possible value, the LHS is lower than the RHS: $\frac{a^\epsilon}{\int_a^\infty a^\epsilon dG(a)} < \frac{\beta p(\theta, \nu)}{\delta + \beta p(\theta, \nu)}$, which is obviously the case if, for instance, $\underline{a} = 0$.

Result 1 *A unique solution always exists, and is such that tightness in the credit market is $\theta = \frac{\beta}{1-\beta}$. The threshold of productivity b is the unique solution to*

$$(1 - H(b, \epsilon)) = \frac{p(\theta, \nu) [1 - G(b)]}{\delta + p(\theta, \nu) [1 - G(b)]} \times \frac{\frac{\beta}{1-\beta}}{\frac{\beta}{1-\beta} + \frac{\delta}{\delta + p(\theta, \nu) [1 - G(b)]}}. \quad (19)$$

Given the values of b and θ from above, the number of brokers in the economy is

$$1 - m = (1 - \beta) (1 - H(b, \epsilon)). \quad (20)$$

Given these values, r and Y solve simultaneously:

$$r\bar{k} = \frac{e_k}{1 - e_k} \pi(b, r, Y), \quad Y = \left[\frac{e_k}{1 - e_k} + \frac{1}{\delta} \right] \pi(b, r, Y). \quad (21)$$

4.1 Irrelevance of Capital Abundance

Result 1 yields the following corollary.

Result 2 *The allocative decisions of the economy θ , m , and b are independent of \bar{k} .*

Capital abundance does not affect how many resources the society devotes to brokerage activities, nor does it affect how selective the society is at evaluating projects. Capital abundance does not affect the activities of the agents; it only determines how much capital each of them uses, and the rewards of capital owners. It does not have redistributive effects with regard to who gets to produce or how fast agents find finance.

One way of understanding this result is as follows. Two variables determine the allocation of human resources: the tightness in the labor market and the productivity of the marginal project. The size of the financial sector can be thought as a residual once these variables are determined. Notice that these two variables are perfectly determined by the arbitrage equation ($V_0 = B$) and the definition of b ($V(b) = V_0$), since none of them depends on the interest rate.

Arbitrage establishes a relationship between θ and b that is independent of the interest rate because albeit the value of each profession depends on the interest rate (and thus, capital abundance), the ratio between them is independent of r . The value of θ depends on the share of the deal for each of the parts, but not on how happy (in absolute terms) they are.

The determination of b establishes that the profit generated by the marginal project is proportional to the average profit. Consequently, having more or less capital (and thus the interest rate) does not affect the marginal to average profit ratio ($H(b, \epsilon)$), which makes the allocative decisions independent of r and capital abundance. Any change in r affects in the same manner the average and the marginal entrepreneur, and thus it has no incidence in the productivity that equals them, b . The ratio of marginal to average profit depends only on how fast entrepreneurs find finance

and the share of the deal that they receive in such a case. More abundant capital will make both the average and the marginal firm more profitable (by decreasing the interest rate), but it will not change their relative stand.

We find this result both surprising and useful. It is surprising because it suggests that the positive correlation between financial sectors and income across countries does not flow from the fact that richer countries are wealthier. It is useful because it clarifies the causes of that relationship and the determinants of the degree of capital market imperfections. It states that imperfections in capital markets do not arise from the scarcity of capital, as might be thought, but from deeper, underlying causes that we will explore later. Moreover, the result simplifies the analysis enormously and makes the rest of the results much easier to understand.

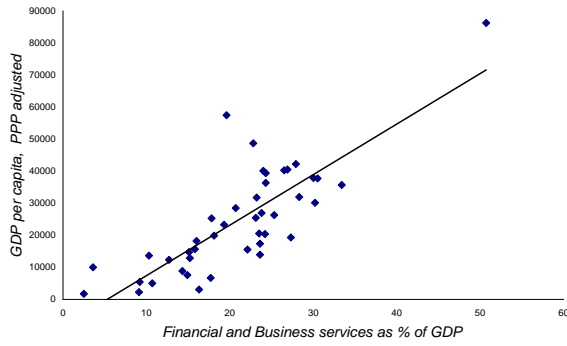
This result is also very robust. It is independent of the pricing mechanism, of the degree of risk aversion of the agents, and of the possibility of self financing. The result comes about whenever there is arbitrage between activities and the interest rate affects proportionally both the marginal and the average firms. Both assumptions seem to us very reasonable.

Finally, independent evidence points in the direction that capital abundance does not determine financial sector size. For instance, Levine and Zervos (1998) show that financial indicators are not associated with private saving rates.

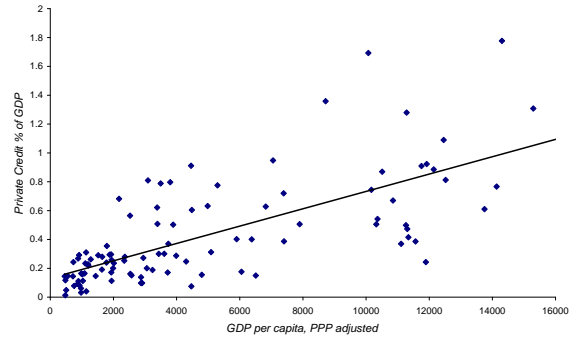
5 The Determination of TFP and Financial Sector Size

GDP and the size of the financial sector are very positively correlated. In figure 1 we plot different measures of the size of the financial sector against GDP per capita for as many economies as we find in the data. Figure 1(a) plots the size of financial and business services for a sample of European and CIS countries. The rest of the graphs plot variables that the literature has usually associated with financial sector breadth and size, following Beck, Demirgüç-Kunt, and Levine (2009). Clearly, richer countries have larger financial sectors.

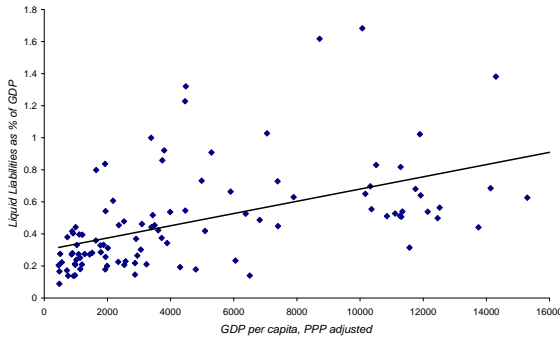
Modern literature tends to explain this correlation via Schumpeterian arguments of economic growth: with more frictions in the capital markets innovative projects do not get financed. Thus, countries with a more efficient financial market become richer. In their seminal work, King and Levine (1993a) and King and Levine (1993b) construct an endogenous growth model where financial sector distortions reduce the rate of economic growth by reducing the rate of innovation. They also provide a wide range of evidence that correlates growth with financial development. Other papers providing theoretical and empirical support for this view are: Levine and Zervos (1998), Levine, Loayza, and Beck (2000), Rousseau and Wachtel (1998), Rousseau and Wachtel (2000), Comin and Nanda (2011) and De Gregorio and Guidotti (1995). Moreover, Neusser and Kugler (1998) find that financial sector GDP is co-integrated for many OECD countries, not so much with manufacturing GDP but mostly with manufacturing total factor productivity. Most of the aforementioned papers relate to growth explicitly: better financial markets, more growth. Overwhelmingly, the direction



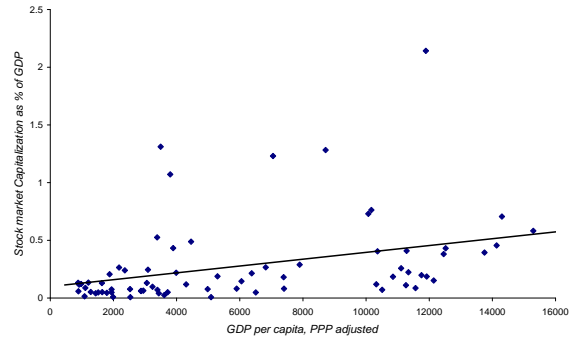
(a) Financial and business services as % of GDP.



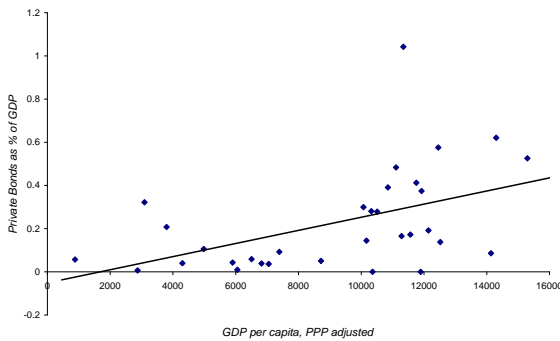
(b) Claims on private sector by deposit money banks and other financial institutions as % of GDP



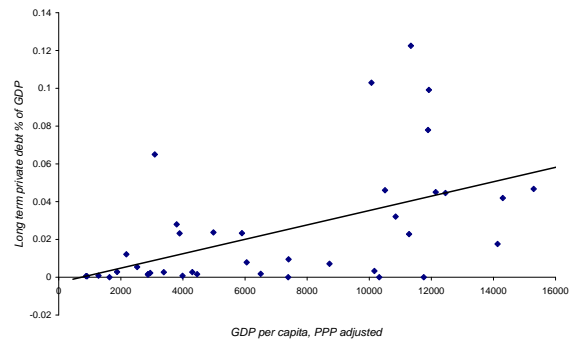
(c) Liquid liabilities as % of GDP



(d) Stock Market capitalization as % of GDP



(e) Outstanding domestic debt securities issued by private domestic entities divided by GDP



(f) Total private long-term debt issues as % of GDP

Figure 1: Different measures of financial sector size plotted against GDP per capita across countries. Sources: For 1(a), UNECE Statistical Division Database (only European and CIS countries); plotted data from 2010. For the rest, Financial Development and Structure Database (World Bank Development Research Group).

of causality that is presumed in the literature goes from financial market efficiency to product market efficiency.

In this section, we show that interpreting the evidence in light of our model suggests that the opposite direction of causality seems qualitatively important in explaining the data: that is, better product markets produce larger financial sectors and more growth.¹⁴

5.1 Effects of Intrinsic Efficiency of the Financial Sector.

An investment sector that is intrinsically more efficient is one where matches are done faster with the same human resource allocation. In our model, this is captured by ν , our characterization of the intrinsic efficiency of the financial sector. Focusing on the steady states, doing comparative statics of this variable is straightforward.

Result 3

(1) *The steady-state values of b and output increase with ν . Furthermore, as ν approaches infinity, the limit of b is its maximum possible value (or infinity if it is unbounded).*

(2) *The steady-state number of brokers, $(1 - m)$, is decreasing with ν .*

Proof: *The RHS of 19 increases with p , and approaches 1 as p approaches infinity. Notice that $p = p(\theta, \nu) = p\left(\frac{\beta}{1-\beta}, \nu\right)$, and ν does not appear anywhere else. Thus, we can think of the effects of ν as a monotonous transformation of the effects of p . From which the result follows. ■*

The first part of the result is very intuitive: the more efficient the matching process is, the better the outcomes that get generated. The reason is that the opportunity cost of going back to search is smaller. It is cheaper to be picky on which projects get financed and thus, both the threshold of quality and output increase. If the market were Walrasian (ν approaching infinity), only the best conceivable projects would get financed.

Likewise, and perhaps more surprisingly, larger values of ν produce a smaller share of GDP devoted to the financial sector. The reason is also straightforward: there are few brokers because they are not needed. Few of them are able to produce fast matches between entrepreneurs and capital, so placing them into productive activities is more efficient. It does not mean that the total GDP generated by the financial sector is smaller, since this increases with ν , only that fewer resources are devoted to intermediation and its *share* of GDP decreases.¹⁵

¹⁴In this section, we do comparative statics on what we think are the more interesting parameters of the model. For the sake of completion, in appendix A we do comparative statics on the remaining parameters.

¹⁵ This result is facilitated by the fact that in our model the match broker-entrepreneur is necessary for production. If entrepreneurs could self-finance via savings the problem would be much more complex, and it would appear a force in the opposite direction, as the “market share” of the financial intermediation sector on total investment would increase with its efficiency (as it would increase the share of intermediation at the expense of the alternative forms of financing). Nevertheless the force stressed in our model would still be there: given a certain share of the sector of financial intermediation, an increase in its productivity should result in a decrease in the amount of human resources devoted to intermediation. In models with different sources of financing, the final effect would result from the convolution of both mechanisms, and a full fledged model would need to be calibrated.

5.2 Effects of the Intrinsic Efficiency of the Productive Sector

One of the main points of our paper is to stress the effects that having a more efficient product market would have on the size of the financial sector. In order to study this direction of causality, in this section we will extend the model. Nevertheless it is worthwhile to notice the following result before we deal with the extension of the model.

Result 4

(1) *In steady state, the minimum productivity threshold b (and consequently Y) are increasing in the elasticity of profits to talent (ϵ), irrespective of the shape of $H(b, \epsilon)$.*

(2) *In steady state, the number of brokers increases with ϵ .*

Proof: *The LHS of (19) decreases with ϵ , whereas the RHS is independent of it and is decreasing in b . Notice that this holds true even if it were the case that $H(b, \epsilon)$ were decreasing in b for some range. ■*

The more important talent is, the more that the society is going to be selective in the projects it chooses to finance. To do so, it is necessary to channel more resources into financial intermediation. The immediate reason is that in order to keep constant the speed at which agents meet, while being more picky on which projects to accept, employing more agents in brokerage is necessary.

While we have not given any interpretation to ϵ yet, there is an intuitive sense in which it is associated to the efficiency and degree of competition in the final good sector.

Intrinsic Efficiency of the Productive Sector

Consider the following tax and transfer scheme. The net profits of a firm are

$$\hat{\pi}(a, r) = \pi(a, r)^{1-\tau} \tilde{\pi}^\tau. \quad (22)$$

where $\tau \in [0, 1)$ measures the degree of progressive redistribution between firms, and $\tilde{\pi}$ is perceived by the agents as lump sum. The tax system does not change the demand of capital of any given firm but, as we will see below, it affects the threshold of quality and, thus, total capital demand.

Clearly, a balanced budget requires:

$$\int_b^\infty \pi(a, r) \frac{dG(a)}{1 - G(b)} = \int_b^\infty \hat{\pi}(a, r) \frac{dG(a)}{1 - G(b)}$$

This follows Benabou (2002). Notice that τ is not the tax rate, but a measure of the progressivity of the tax system. Given any value of τ the effective tax rate is larger the more efficient a firm is, as the tax system is progressive, creating a wedge that potentially distorts activity. Thus, a higher τ implies bigger allocative inefficiencies in the economy. Higher τ transfers more profitability from efficient to inefficient firms, and thus the usage of resources is less efficient. It reflects

a larger degree of inefficiency in the *product market*. It is not very different from other measures of allocative inefficiency that have been proposed in the literature. Thus, we think of $1 - \tau$ as a measure of the degree of competition in the economy.¹⁶ In the context of our model, it has the added advantage of producing simple analytical solutions. Notice that in our context τ simply decreases the elasticity of profits to productivity. Thus, proving the following result is very easy.

Result 5 *A decrease in the allocative inefficiencies of the product sector (i.e., a decrease of τ) produces larger steady-state values of b and Y and a decrease in m .*

In other words, a more efficient product market does a better job at discriminating between bad and good firms. There exists a private interest in investing more in differentiating the treatment of good and bad firms, and the economy does this by generating a larger financial sector: more brokers allow many rejections to take place without decreasing the speed at which entrepreneurs meet brokers.

Notice that causality is the opposite of the one that is usually assumed. It is not that a larger financial sector generates an efficient product sector, but rather the reverse. The large financial sector is large precisely because the productive sector is efficient. If it were inefficient, it would not be very important where capital is placed, since inefficient firms would do relatively well. There would be neither the need nor the demand to devote resources to place capital in a better firm. Human resources would be devoted to production, not to move capital toward more efficient firms.

The fact that the product sector is discriminating and competitive (efficient) is precisely what generates a large financial sector, because it is more important to use capital well. It generates the need and the demand for devoting human resources toward determining the use of capital.

5.3 Discussion

In the model, income, total factor productivity, and the size (and efficiency) of the financial sector are all endogenous objects. They are affected by underlying efficiency considerations in both the financial and productive sectors of the economy. A more efficient economy (larger TFP and larger income) can arise as a consequence of either a larger underlying efficiency in the financial sector or a more competitive product market environment; but these two possible causes of a more efficient economy have very different implications for the size of the financial sector and the average firm.

From result 3 it follows that if differences in TFP were *exclusively* a consequence of underlying differences in financial sector efficiency, we should observe a *negative* correlation between financial sector and TFP (or income). Moreover, rich countries would have more firms, and these firms would be smaller. Repeating the reason is worthwhile: if the financial sector were intrinsically very efficient, it would be small because very few people would be sufficient to run it. It would

¹⁶Subsidizing the more productive firms would also generate inefficiencies, and a bigger financial sector, but this is not captured in our formulation. For this reason our favoured interpretation of τ is as a measure of the degree of competition in the economy.

not demand large resources; in the limit, if the financial sector were Walrasian, the relative size of the financial sector would approach zero. And because a smaller financial sector implies a larger productive sector, more firms would use the available capital. Thus, firms would be smaller.

On the other hand, from result 5 it follows that if the cause of cross-country differences in TFP were *exclusively* a consequence of underlying differences in product market allocative efficiency, we should instead observe a *positive* correlation between the financial sector and TFP. More competitive product markets make it more important to allocate capital to better firms. Competition increases the value of spending resources to ensure that capital flows to the right firms. In response to this demand, a larger fraction of the population opts for brokerage activities, improving the size (and efficiency) of the financial sector. More brokers imply fewer entrepreneurs, each of whom employs a larger share of the capital. The economy ends up with fewer, albeit larger, firms.¹⁷

It follows that if we were to calibrate our model to the correlation income/financial sector size, the bulk of the difference between the fundamentals of rich and poor countries should be in the degree of competition (and allocative efficiency) in product markets, *not in the efficiency of the financial sector*. This is not to suggest that financial sectors of rich countries are intrinsically less efficient, but that the differences in intrinsic efficiency of the financial sector seem to be smaller than the differences in efficiency in the product market. The model interprets the data suggesting that rich countries are richer mainly because they have a more efficient final goods sector, which generates demand for a large financial sector, not *because* their financial sectors are intrinsically more efficient.

Little direct evidence supports proving or disproving this point, since this direction of causality has been mainly unexplored,¹⁸ but it is remarkable that Neusser and Kugler (1998) do find evidence of this direction of causality using the Granger-Lin measure of long-run causality.

Putting everything together, this seems suggestive that the direction of causality going from product market efficiency to financial sector efficiency (and size) has an important role in causing the observed correlations.

6 Conclusions

Our contribution is to present a model of misallocation with endogenous financial frictions. We allow for two resource constraints: one on capital that can be used and one on the number of individuals who can devote their time either to brokerage or to directly productive activities. The more brokers there are, the faster that entrepreneurs will find capital. Our model has three results

¹⁷ It is not only the size of the financial sector. Poschke (2011) shows that average firm size is larger in richer economies. This is something that our model would generate only if the bulk of the difference between the rich and the poor countries lies in the intrinsic efficiency of their final good sectors, and not in the intrinsic efficiency of their financial sector.

¹⁸ An exception is Rajan and Zingales (1998) which shows that better financial sector produces growth even in sectors where the reverse causality could not possibly explain it. They do not test directly for the impact of the reverse causality.

of note:

(i) The way that society uses its human resources is independent of the abundance of capital. Richer economies do not have a larger financial sector *because* they are richer.

(ii) The more efficient the financial sector is, the fewer resources are devoted to it. Intermediation is complementary to entrepreneurial activities in production. Thus, an inefficient financial sector would absorb large amounts of resources relative to what a more efficient capital market would absorb.

(iii) The degree of allocative efficiency of the productive sector has exactly the opposite implication: it generates a larger financial sector. The reason is that there are more incentives to allocate capital in the right projects in a competitive environment. Thus, more resources are devoted in equilibrium to doing so. Moreover, resulting firms would be larger and more homogeneous.

Looking at the data in light of these results suggests that the bulk of the difference between rich and poor countries lies not in how much more efficient the *financial* sector is in rich countries, but on how much more efficient their *final good* sector is.

Our implication is not that in rich countries financial sectors are not better than those in poor countries. Of course they are better.¹⁹ What we want to stress is that productive sectors are also much more efficient in rich countries. In light of the model, this second source of difference between rich and poor countries seems to account for the bulk of the differences in performance across countries.

Of course, many caveats should be kept in mind. We believe that the main one is that by not allowing self-financing we make intermediation a necessity for production. This facilitates result 3, but by no means it imposes it. The mechanism that produces this result (i.e., the economy would not *need* to use many people in finance if they are highly productive) would always be there, although it is conceivable that in a more general model, the relationship between the underlying efficiency of capital markets and the share of intermediation on GDP might not be monotonous.

Moreover, we can see no reasons why allowing for self financing would change our main result. Product market competition will always increase the value of putting inputs into the right firm and diverting input away from the bad one. Such competition will increase the demand for financial services and will tend to increase the size of the financial sector.

In any case, we do not mean to underscore the value of extending the model to allow for self-financing. Actually, we believe that it would open the door to explore fascinating issues, such as the relationship between inequality and financial sector size and efficiency. The project would nevertheless be substantially more involved, since it would need to model the process of capital accumulation and heterogeneity of both capital holdings and investment opportunities. We intend to extend our model in this direction in further research.

¹⁹Indeed a simple extension allowing for endogenous growth (see appendix B.1) captures that better financial markets generate more growth.

References

- ALFARO, L., A. CHARLTON, AND F. KANCZUK (2008): “Plant-Size Distribution and Cross-Country Income Differences,” *NBER International Seminar on Macroeconomics*, 5(1), 243–272.
- BARTELSMAN, E. J., J. C. HALTIWANGER, AND S. SCARPETTA (2013): “Cross-Country Differences in Productivity: The Role of Allocation and Selection,” *American Economic Review*, 103(1), 305–334.
- BECK, T., A. DEMIRGÜÇ-KUNT, AND R. LEVINE (2009): “Financial Institutions and Markets across Countries and over Time - Data and Analysis,” *Policy Research Working Paper Series 4943*, *The World Bank*.
- BENABOU, R. (2002): “Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?,” *Econometrica*, 70 (2), 481–517.
- BENTO, P., AND D. RESTUCCIA (2017): “Misallocation, Establishment Size, and Productivity,” *American Economic Journal: Macroeconomics*, 9(3), 267–303.
- BESCI, Z., V. E. LI, AND P. WANG (2005): “Heterogeneous Borrowers, Liquidity and the Search for Credit,” *Journal of Economic Dynamics and Control*, 29 (8), 1331–1360.
- (2009): “Credit Mismatch and Breakdown,” *Mimeo*, *Washington University in St. Louis*.
- BLANCHFLOWER, D. G., AND A. J. OSWALD (1998): “What Makes an Entrepreneur?,” *Journal of Labor Economics*, 16(1), 26–60.
- BLOOM, N., AND J. V. REENEN (2007): “Measuring and Explaining Management Practices Across Firms and Countries,” *Quarterly Journal of Economics*, 122(4), 1351–1408.
- BUERA, F. J., J. KABOSKI, AND Y. SHIN. (2011): “Finance and Development: A Tale of Two Sectors,” *American Economic Review*, 101(5), 1964–2002.
- CAGETTI, M., AND M. DE NARDI (2006): “Entrepreneurship, Frictions, and Wealth,” *Journal of Political Economy*, 114(5), 835–870.
- CHARI, V. V., P. J. KEHOE, AND E. R. MCGRATTAN (2007): “Business Cycle Accounting,” *Econometrica*, 75 (3), 781–836.
- CHIU, J., AND T. V. KOEPL (2011): “Trading Dynamics with Adverse Selection and Search: Market Freeze, Intervention and Recovery,” *The Review of Economic Studies*, 83(6), 969–1000.
- COMIN, D., AND R. NANDA (2011): “Financial Development and Technology Diffusion,” *Mimeo*, *Harvard Business School*.
- COOLEY, T., R. MARIMON, AND V. QUADRINI (2004): “Aggregate Consequences of Limited Contract Enforceability,” *Journal of Political Economy*, 112(4), 817–847.

- CUI, W., AND S. RADDE (2016): “Search-Based Endogenous Asset Liquidity and the Macroeconomy,” *Working Paper*.
- DE GREGORIO, J., AND P. GUIDOTTI (1995): “Financial Development and Economic Growth,” *World Development*, 23(3), 433 – 448.
- DELL’ARICCIA, G., AND P. GARIBALDI (1998): “Bank Lending and Interest Rate Changes in a Dynamic Matching Model,” *International Monetary Fund Working Paper*, (WP/98/93).
- (2005): “Gross Credit Flows,” *Review of Economic Studies*, 72(3), 665–685.
- DIAMOND, P. (1990): “Pairwise Credit in Search Equilibrium,” *Quarterly Journal of Economics*, 105(2), 285–319.
- DUFFIE, D., N. GARLEÂNÚ, AND L. H. PEDERSEN (2005): “Over-the-Counter Markets,” *Econometrica*, 73(6), 1815–1847.
- EVANS, D. S., AND B. JOVANOVIĆ (1989): “An Estimated Model of Entrepreneurial Choice under Liquidity Constraints,” *Journal of Political Economy*, 97(4), 808–827.
- EVANS, D. S., AND L. S. LEIGHTON (1989): “Some Empirical Aspects of Entrepreneurship,” *American Economic Review*, 79(3), 519–535.
- GENTRY, W. M., AND R. G. HUBBARD (2000): “Tax Policy and Entrepreneurial Entry,” *American Economic Review*, 90(2), 283–287.
- GREENWOOD, J., AND B. JOVANOVIĆ (1990): “Financial Development, Growth, and the Distribution of Income,” *Journal of Political Economy*, 98(5), 1076–1107.
- GREENWOOD, J., J. M. SANCHEZ, AND C. WANG (2010): “Financing Development: The Role of Information Costs,” *American Economic Review*, 100(4), 1875–91.
- (2013): “Quantifying the Impact of Financial Development on Economic Development,” *Review of Economic Dynamics*, 16(1), 194–215.
- GUIO, L., P. SAPIENZA, AND L. ZINGALES (2004): “Does Local Financial Development Matter?,” *Quarterly Journal of Economics*, 119(3), 929–969.
- GUNER, N., G. VENTURA, AND Y. XU (2008): “Macroeconomic Implications of Size-Dependent Policies,” *Review of Economic Dynamics*, 11(4), 721–744.
- HAAN, W. J. D., G. RAMEY, AND J. WATSON (2003): “Liquidity Flows and Fragility of Business Enterprises,” *Journal of Monetary Economics*, 50(6), 1215–1241.
- HOLTZ-EAKIN, D., D. JOULFAIAN, AND H. S. ROSEN (1994): “Entrepreneurial Decisions and Liquidity Constraints,” *RAND Journal of Economics*, 25(2), 334–347.

- HOPENHAYN, H. A. (1992): “Entry, Exit, and Firm Dynamics in Long Run Equilibrium,” *Econometrica*, 60(5), 1127–1150.
- HSIEH, C.-T., AND P. J. KLENOW (2009): “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 124(4), 1403–1448.
- KING, R. G., AND R. LEVINE (1993a): “Finance and Growth: Schumpeter Might Be Right,” *Quarterly Journal of Economics*, 108(3), 717–737.
- KING, R. G., AND R. LEVINE (1993b): “Finance, Entrepreneurship, and Growth,” *Journal of Monetary Economics*, 32 (3), 513–542.
- LEVINE, R. (1997): “Financial Development and Economic Growth: Views and Agenda,” *Journal of Economic Literature*, 35(2), 688–726.
- LEVINE, R., N. LOAYZA, AND T. BECK (2000): “Financial intermediation and growth: Causality and causes,” *Journal of Monetary Economics*, 46 (1), 31–77.
- LEVINE, R., AND S. ZERVOS (1998): “Stock Markets, Banks, and Economic Growth,” *American Economic Review*, 88(3), 537–558.
- LI, W., AND P.-D. G. SARTE (2003): “Credit Market Frictions and Their Direct Effects on U.S. Manufacturing Fluctuations,” *Journal of Economic Dynamics and Control*, 28(3), 419–443.
- LUCAS, R. E. (1978): “On the Size Distribution of Business Firms,” *Bell Journal of Economics*, (9 (2)), 508–523.
- MATSUYAMA, K. (2008): “Aggregate Implications of Credit Market Imperfections,” in *NBER Macroeconomics Annual 2007, vol. 22, ed D. Acemoglu, K. Rogoff, and M. Woodford. Chicago. Chicago University Press*, pp. 1–60.
- MELITZ, M. J. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71(6), 1695–1725.
- MIDRIGAN, V., AND D. Y. XU (2014): “Finance and Misallocation: Evidence from Plant-Level Data,” *American Economic Review*, 104(2), 422–58.
- MONNET, C., AND D. SANCHES (2015): “Private Money and Banking Regulation,” *Journal of Money, Credit and Banking*, 47(6), 1031–1062.
- MORTENSEN, D. T., AND C. A. PISSARIDES (1994): “Job Creation and Job Destruction in the Theory of Unemployment,” *Review of Economic Studies*, 61(3), 397–415.
- NEUSSER, K., AND M. KUGLER (1998): “Manufacturing Growth and Financial Development: Evidence from OECD Countries,” *Review of Economics and Statistics*, 80(4), 638–646.

- PETERSEN, M. A., AND R. G. RAJAN (2002): “Does Distance Still Matter? The Information Revolution in Small Business Lending,” *Journal of Finance*, 57(6), 2533–2570.
- PHILIPPON, T. (2008): “The Evolution of the US Financial Industry from 1860 to 2007: Theory and Evidence,” *Working Paper*, NYU.
- POSCHKE, M. (2011): “The Firm Size Distribution Across Countries and Skill-Biased Change in Entrepreneurial Technology,” *Mimeo*, *McGill University*.
- QUADRINI, V. (2000): “Entrepreneurship, Saving, and Social Mobility,” *Review of Economic Dynamics*, 3(1), 1–40.
- RAJAN, R. G., AND L. ZINGALES (1998): “Financial Dependence and Growth,” *The American Economic Review*, 88(3), 559–586.
- RESTUCCIA, D., AND R. ROGERSON (2008): “Policy Distortions and Aggregate Productivity with Heterogeneous Establishments,” *Review of Economic Dynamics*, 11(4), 707–720.
- RESTUCCIA, D., AND R. ROGERSON (2013): “Misallocation and productivity,” *Review of Economic Dynamics*, 16(1), 1–10.
- (2017): “The Causes and Costs of Misallocation,” *Journal of Economic Perspectives*, 31(3), 151–74.
- ROUSSEAU, P. L., AND P. WACHTEL (1998): “Financial Intermediation and Economic Performance: Historical Evidence from Five Industrialized Countries,” *Journal of Money, Credit and Banking*, 30(4), 657–678.
- ROUSSEAU, P. L., AND P. WACHTEL (2000): “Equity Markets and Growth: Cross-Country Evidence on Timing and Outcomes, 1980-1995,” *Journal of Banking & Finance*, 24 (12), 1933–1957.
- RUSS, K. N., AND D. VALDERRAMA (2009): “A Theory of Banks, Bonds, and the Distribution of Firm Size,” *NBER Working Paper No. 15454*.
- SILVERA, R., AND R. WRIGHT (2007): “The Venture Capital Cycle,” *Working Paper*, University of Pennsylvania.
- (2010): “Search and the Market for Ideas,” *Journal of Economic Theory*, 145(4), 1550–1573.
- WASMER, E., AND P. WEIL (2004): “The Macroeconomics of Labor and Credit Market Imperfections,” *American Economic Review*, 94(4), 944–963.

A Appendix

A.1 Effects of the Destruction Rate

A faster death rate δ shortens the horizon of agents and makes them less picky.

Result 6 .

b is not increasing in δ and is strictly decreasing in δ if $H(b, \epsilon)$ is strictly increasing in b . The number of entrepreneurs does not decrease with δ and strictly increases in δ if $H(b, \epsilon)$ is strictly increasing in b .

Proof: The RHS of (19) decreases with δ , whereas the LHS is independent of it and increasing if $H(b, \epsilon)$ is increasing in b . The movement of $1 - m$ is implied by (20). ■

The shorter the time that the entrepreneur has left before dying, the less important quality is, and the more important that starting work soon becomes. Consequently, if agents have short lives, they are less picky and finance projects of lesser quality: they should not spend their lives looking for projects.

Notice that the number of brokers would *increase*, in spite of the fall in the productivity threshold. The reason is that although it is easier to get your project financed (thus needing fewer brokers per entrepreneur), you are also replaced faster by somebody who is unmatched. Thus, a larger destruction rate demands a larger financial sector, because by definition many entrepreneurs (the newborns) are in need of capital. As $\delta \rightarrow \infty$, the number of brokers approaches $(1 - \beta)$ in order to keep constant the tightness of the market.

A.2 Effects of the Bargaining Power of Entrepreneurs

Result 7 *There exists a value of β called $\hat{\beta} : 1 - \hat{\beta} = -\frac{\theta}{p(\theta, \nu)} \frac{\partial p(\theta, \nu)}{\partial \theta}$ such that $\hat{\beta}$ maximizes b (and thus, Y). If $\beta < \hat{\beta} \rightarrow \frac{db}{d\beta} > 0$, and if $\beta > \hat{\beta} \rightarrow \frac{db}{d\beta} < 0$.*

In the vicinity of $\hat{\beta}$, any marginal increase of β decreases $1 - m$.

Proof: The proof is straightforward once we realize that the derivative of the RHS of (19) with respect to θ (which is just a monotonous transformation of β) is (23), and that then from (20) it follows (24)

$$\frac{\delta[1 - G(b)]}{(\delta + \beta p(\theta, \nu)[1 - G(b)])^2} \frac{p(\theta, \nu)}{1 - \beta} \left[1 - \beta + \frac{\theta \frac{\partial p(\theta, \nu)}{\partial \theta}}{p(\theta, \nu)} \right]. \quad (23)$$

$$\frac{d(1 - m)}{d\beta} = -[1 - H(b, \epsilon)] - (1 - \beta) \frac{\partial H(b, \epsilon)}{\partial b} \frac{db}{d\beta}. \quad (24)$$

Given that at $\beta = \hat{\beta}$, $\frac{db}{d\beta} = 0$, it is clear that in its vicinity $\frac{d(1-m)}{d\beta} < 0$ ■

Increasing β has two opposite effects on the present discounted value of the entrepreneur: it increases the share of profits that she appropriates, but it also increases tightness (θ), reducing the speed at which the entrepreneur finds brokers.

Notice from equation (19) that b is maximized when β is such that it maximizes the value of being an entrepreneur (V_0) for any given b . This is simply the Hosios condition. More power to entrepreneurs produces a search externality, as there will be more of them, which will decrease their speed of finding capital. Their value (and the marginal productivity) is maximized when the elasticity of the matching function equals their bargaining share.

Notice that a larger β increases the value of being an entrepreneur. Arbitrage demands that this is compensated by increasing tightness (equation (12)), reducing the size of the financial sector.²⁰

B Examples and Growth

In this section, we explicitly solve the model for several functional form assumptions. This exercise is instructive of the points raised above and allows us to place our model in the context of growth.

We first solve explicitly for a model with the following functional assumptions:

- (1) The production function is Cobb-Douglas: $F(a, K, Y) = 2\sqrt{aK}$.
- (2) τ characterizes a progressive redistributive scheme as in section 5.2.
- (3) a follows a Pareto distribution with minimum value \underline{a} and parameter γ .

Then, gross profit function, capital demand, net profit, and $\tilde{\pi}$ are, respectively,

$$\begin{aligned}\pi(a, r) &= \frac{a}{r} \\ k^d(a, r) &= \frac{a}{r^2} \\ \hat{\pi}(a, r) &= \left(\frac{a}{r}\right)^{1-\tau} \tilde{\pi}^\tau \\ \tilde{\pi} &= \left(\frac{\gamma-(1-\tau)}{\gamma-1}\right)^{\frac{1}{\tau}} b\end{aligned}\tag{25}$$

Given that b is bounded from below by \underline{a} , it is straightforward to check that result 1 implies the following result.²¹

Result 8 *There exists a level of taxes $\tilde{\tau} = \frac{1-(\gamma-1)\frac{1}{\beta}\frac{\delta}{p(\frac{\beta}{1-\beta}, \nu)}}{1+\frac{1}{\beta}\frac{\delta}{p(\frac{\beta}{1-\beta}, \nu)}} \in (0, 1)$ such that*

$$\begin{aligned}1 - G(b) &= \begin{cases} 1 + \frac{1}{\beta} \frac{\delta}{p(\frac{\beta}{1-\beta}, \nu)} \frac{\tau+\gamma-1}{1-\tau} & \text{if } \tau \leq \tilde{\tau} \\ 1 & \text{if } \tilde{\tau} \leq \tau \end{cases} \\ b &= \begin{cases} \underline{a} \left[\beta \frac{p(\frac{\beta}{1-\beta}, \nu)}{\delta} \frac{1-\tau}{\tau+\gamma-1} \right]^{\frac{1}{\gamma}} & \text{if } \tau \leq \tilde{\tau} \\ \underline{a} & \text{if } \tilde{\tau} \leq \tau \end{cases} \\ 1 - m &= \begin{cases} (1 - \beta) \frac{1-\tau}{\gamma} & \text{if } \tau \leq \tilde{\tau} \\ (1 - \beta) \frac{1}{1+\beta \frac{p(\frac{\beta}{1-\beta}, \nu)}{\delta}} & \text{if } \tilde{\tau} \leq \tau \end{cases}\end{aligned}\tag{26}$$

The threshold of quality is increasing in ν and decreasing in τ , but it can not be smaller than \underline{a} . Thus, there is a degree of inefficiency in the product sector beyond which all projects are accepted.

The Pareto distribution assumption illustrates very clearly the roles of both ν and τ , since they appear in two different equations. This is because it ensures that the marginal to average profit ratio is a constant that does not depend on b and thus indirectly on ν :

$$H(b, 1 - \tau) = \frac{\gamma - (1 - \tau)}{\gamma}$$

²⁰Another effect complicating matters is that b changes when β does. Nevertheless, this effect is nil when the Hosios condition holds, since there the marginal effect of β on b is negligible

²¹This is under the assumption that $(\gamma - 1)\frac{1}{\beta}\frac{\delta}{p(\frac{\beta}{1-\beta}, \nu)} \leq 1$. Otherwise, the solution is just too obvious.

Consequently, whenever we are in an interior solution ($\tau < \tilde{\tau}$, and not all projects are financed), the size of the financial sector is strictly decreasing in τ (equation 20).

The effects of ν are apparent when all projects are financed ($\tau \geq \tilde{\tau}$). In such a case the more efficient capital markets are, the fewer brokers are needed.

We can then go on to define total factor productivity as

$$A = \begin{cases} b \left(1 + \frac{\tau}{\gamma-1}\right) = \underline{a} \left[\beta^{\frac{p(\frac{\beta}{1-\beta}, \nu)}{\delta}} \frac{1-\tau}{\tau+\gamma-1} \right]^{\frac{1}{\gamma}} \left(1 + \frac{\tau}{\gamma-1}\right) & \text{if } \tau \leq \tilde{\tau} \\ \underline{a} \frac{\gamma}{\gamma-1} \frac{\beta^{\frac{p(\frac{\beta}{1-\beta}, \nu)}{\delta}}}{1+\beta^{\frac{p(\frac{\beta}{1-\beta}, \nu)}{\delta}}} & \text{if } \tilde{\tau} \leq \tau. \end{cases} \quad (27)$$

Using result 1, given any level of capital \bar{k} , we get

$$r = \frac{\sqrt{A}}{\sqrt{\bar{k}}} \quad ; \quad Y = 2\sqrt{A}\sqrt{\bar{k}}. \quad (28)$$

The comparative static exercises in this example are illustrative of some of our main points. They clearly show how each exogenous variable affects total factor productivity, interest rates and output.

A decrease in the amount of frictions in the capital markets (increase of ν) increases TFP, but via two different mechanisms. It generates more efficient firms (via raising b), but also makes them smaller, thus increasing capital productivity. In the specific case of the Pareto distribution, while we are in an interior solution ($\tau < \tilde{\tau}$), only the first of these mechanisms is active (because m is fixed). Nevertheless, when $\tilde{\tau} < \tau$, it is possible to notice the second mechanism: the size of the brokerage sector decreases, which results in more firms being created and thus smaller firm size and larger productivity. There is no increase in the quality of projects (since in this case, all projects are being financed), but there is a decrease in average size, and this increases productivity. In the general case (not assuming a Pareto distribution), both effects would be present: an increase in the quality of firms *and* an increase in the size of human resources devoted to productive activities, which results in smaller (more productive) firms.

A decrease of the efficiency of the final goods market (i.e., a higher τ) has two effects that go in opposite directions. On one hand, it decreases the quality of firms (i.e., a lower b), thus decreasing TFP. On the other hand, more human resources are devoted to productive activities (i.e., m increases). Thus, firms are smaller and each of them uses less capital, thereby increasing TFP (this is the effect of $\frac{\tau}{\gamma-1}$). It is easy to check that in this example, the first effect dominates, but we know from result 5 that it does so for any other distribution of qualities and any profit function. Notice the two forces at work here: **(i)** A more efficient final goods market generates the demand for brokerage services, which in turn generates an improvement in the allocation of talent. **(ii)** This gain is partially offset; by putting fewer human resources into productive activities, the number of firms is necessarily smaller, and albeit they are more efficient, they are also larger and each has lower marginal capital productivity.

B.1 Growth

The irrelevance of capital abundance for the allocation of human resources makes it very easy to solve a dynamic version of the model. Assume that agents save (and invest) at an exogenous rate s and that capital depreciates at a rate d (also exogenous). Assume further that once a match has been made, firms can reevaluate their capital needs as they see it fit, given the prevailing interest

rate. Let k_t denote capital at time t :

$$\dot{k}_t = sY_t - dk_t. \quad (29)$$

It follows that the steady-state level of income is

$$Y^{SS} = \begin{cases} 2\frac{s}{d}A & \text{if } \tau \leq \tilde{\tau} \\ 2\frac{s}{d}a \frac{\gamma}{\gamma-1} \frac{\beta^{\frac{p(\frac{\beta}{1-\beta}, \nu)}}{\delta}}{1+\beta^{\frac{p(\frac{\beta}{1-\beta}, \nu)}}{\delta}}} & \text{if } \tilde{\tau} \leq \tau. \end{cases} \quad (30)$$

The equilibrium size of the financial sector would still be determined by (26). Thus, clearly if rich and poor countries were to differ by how efficient their financial sector is, there would be no correlation between financial sector size and income. If, on the other hand, they were to differ in the efficiency of the final goods sector, then the correlation would be positive.

AK endogenous growth model

Solving a model of endogenous growth with credit search fictions is almost straight-forward.

We maintain the assumptions of the Pareto distribution of a and the existence of an inefficiency in the goods market labeled by τ , but we change the utility and production functions.

We assume utility to be given by the standard Dixit–Stiglitz aggregator. For simplicity, the elasticity of demand is imposed to be $\frac{1}{2}$: $Y = \left(\int_0^N x_i^{\frac{1}{2}} di \right)^{1/\frac{1}{2}}$. Normalizing the aggregate price to one, the implied demand for firm i is $x_i = Yp_i^{-2}$.

Finally, we assume that production is done with constant returns to scale and we use only capital as input. A firm with a productivity quality a produces $x_i = 4ak_i$ units of output if it uses k_i units of capital.

It follows that profits and capital demand are the same as in the previous sections and characterized in equation (25). Thus, the equilibrium productivity threshold and size of the financial sector are the same than as the previous model and determined by (26). Furthermore, again as before, they are independent of capital supply.

Under the maintained assumption that capital accumulates according to (29), total factor productivity is still determined by (27), and interest rate and output are, respectively, $r = A$ and $Y = 2AK_t$, and the growth rate of the economy is $\frac{\dot{k}}{k} = 2sA - 2d$

In this context, more efficient financial and final good markets bring more growth (via higher total factor productivity), but as before, the first one with a smaller financial sector and the second one with a larger one.

Notice that our results are compatible with the strong evidence suggesting that better financial sectors create long run growth. For instance, Levine, Loayza, and Beck (2000) show that the exogenous components of financial development and cross-country variation in legal origin (taken as an IV of development) explain growth. In our model, these variables would affect ν , improving the efficiency of matching in the investment room, and thereby increasing TFP and growth.