Financial Intermediation and Economic Development: A Quantitative Assessment*

Pedro S. Amaral
Southern Methodist University
University of Minnesota

Erwan Quintin
Federal Reserve Bank of Dallas

April 2, 2007

Abstract

We present a model in which the importance of financial intermediation for economic development can be measured. We generate financial differences by varying the degree to which contracts can be enforced. Economies where enforcement is poor direct less capital to the production sector, and employ less efficient technologies. Calibrated simulations reveal that the resulting effect on output is large. Furthermore, the model correctly predicts that the average scale of production should rise with the quality of enforcement. Finally, we find that the importance of financial development rises with the importance of capital in production.

*E-mail: amaral@econ.umn.edu and erwan.quintin@dal.frb.org. We thank Rui Castro, Tim Kehoe, Berthold Herrendorf, Ed Prescott and Richard Rogerson for their comments, as well as seminar participants at the 2004 Annual Meeting of the Society for Economic Dynamics, Arizona State University, the Federal Reserve Bank of Minneapolis, Texas A&M, the University of Pittsburgh, the Universidad Católica Argentina, the University of Western Ontario, Queen's University and the Universidad de Guanajuato. The views expressed herein are those of the authors and may not reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

1 Introduction

Under standard neoclassical assumptions, observed differences in human and physical capital per worker cannot account for differences in output per worker across nations.[1] In the language of development accounting, total factor productivity (TFP) varies greatly across countries. At the same time, financial development and economic development are highly correlated.[2] This strong empirical relationship is often presented as evidence that financial development causes economic development by promoting investment and making the allocation of resources more efficient.

A number of models have established a theoretical, qualitative link between financial development and economic development and growth.[3] Cross-country regressions suggest that financial development could have a large impact on economic development. Our goal is to quantify the importance of financial intermediation for economic development in the context of a dynamic general equilibrium model.[4]

In our model economy, agents live for three periods, but only work in the first two. When young, they supply labor inelastically. In the second period of their life they can choose to manage a technology that transforms capital and labor into the consumption good. As in the span of control model of Lucas (1978), some agents manage resources more efficiently than others. Managers finance part of their capital through their own savings and borrow the rest from an intermediary. The market for loans is imperfect however, as agents can choose to default on the payment they owe the intermediary at an exogenous cost.[5]

[1]See for example Prescott (1998), or Hall and Jones (1999).

[2]See for example Goldsmith (1969), McKinnon (1973), Shaw (1973), or King and Levine (1993). Levine, Loayza, and Beck (2000) use GMM dynamic panel estimators to partially deal with the possibility of reverse causality.

[3]See, among many others, Greenwood and Jovanovic (1990), Bencivenga and Smith (1991), Khan (2001) and Erosa and Hidalgo (2004). Amaral and Quintin (2006) argue that financial constraints could explain why developing nations emphasize informal, labor-intensive activities.

[4]Townsend and Ueda (2005) and Jeong and Townsend (2004) apply methods similar to ours to study the connection between financial deepening, inequality, and growth, and the relationship between financial deepening and TFP growth, respectively. Models that quantify the importance of factors other than finance include Hall and Jones (1999), Acemoglu and Zilibotti (2001), Restuccia and Rogerson (2003), Herrendorf and Teixeira (2004), and Restuccia (2004).

[5]Shleifer and Wolfenzon (2002) build a model economy that bears some similarities to ours but focus on comparing qualitative predictions of their model to corporate finance regularities.

This default cost is the basis for the quantitative exercise we carry out. In all our experiments, we compare economies that differ in one respect only: the degree to which financial contracts can be enforced. By generating financial differences via enforcement frictions, we adopt and formalize the view that the quality of institutions is a key determinant of financial development. The findings of Levine, Loayza, and Beck (2000), for instance, suggest not only that financial development causes economic growth, but also that improvements in contract enforcement promote financial development.

In our model, economies where contract enforcement is poor emphasize self-financing and, therefore, production on a small scale. As a result, less productive technologies (more establishments) need to be operated for labor markets to clear. Output is lower in economies with bad financial markets because less capital is used in production and because capital is not directed to its best uses. In other words, our model incorporates the two channels emphasized by the financial development literature ever since the seminal work of Goldsmith (1969), McKinnon (1973), and Shaw (1973).

To quantify the importance of these effects, we begin by calibrating our model to match relevant features of the U.S. economy. In particular, we calibrate the distribution of managerial talent to match salient features of the organization of production in the United States, and the default cost to approximate the ratio of aggregate financial liabilities to output. Then we vary the default parameter to generate a sequence of economies with different levels of financial development.

We find that finance disrupts the organization of production greatly. As access to finance falls, more agents need to become managers for labor markets to clear (more inefficient technologies need to be activated), and the average size of establishments falls by magnitudes very similar to what one observes in the available cross-country data.

We also find that financial differences have a large impact on output. Our calculations suggest that financial differences can account for a significant part of the gap between nations such as the United States and middle-income nations such as Mexico or Argentina. Nevertheless, under standard technological assumptions the resulting dispersion in output falls short of what one observes in the data. This is not surprising since our model abstracts from many development-relevant characteristics of low-income nations.

Importantly, provided opportunities to invest capital outside of the production sector exist, we obtain sizeable output differences without much variation in capital-output ratios, which accords well with standard neoclassical accounting results. In a version of our economy with a storage option, what makes economies with weak enforcement poor in our model is not that they have low savings rates, but rather that they fail to channel enough resources to the production sector, and that they tend to employ less productive technologies.

These results are robust to even drastic changes in most parameters, with two important exceptions. Raising the capital share greatly increases the effect of finance on output, for obvious reasons. Similarly, reducing the degree to which labor can be substituted for capital magnifies the output effect without changing the impact of finance on the capital-output ratio much. That varying those two parameters can magnify the effect of factor differences on output is well-known. For instance, Caselli (2003) argues that if the elasticity of substitution between capital and labor is small enough, the observed dispersion in factor endowments can account for the dispersion in output across countries. What we find is that financial differences provide a complementary magnification effect. The effect on output of given changes in the elasticity of substitution becomes much larger when one models financial frictions explicitly.

Our experiments complement the existing literature on the links between financial development and economic development by bridging the gap between theoretical models and empirical findings. We show that the channels emphasized by the theoretical literature are in fact quantitatively meaningful.

2 The economy

We consider a discrete-time model in which a mass one of 3-period-lived agents are born each period. Agents are endowed with a level $z \in Z$ of managerial ability, which is public information. The distribution μ of managerial ability is the same across generations and has finite support with $\mu\{0\} < 1$.

In the first period of their life, agents inelastically supply one unit of labor services. In the second period, they can once again supply labor services or, instead, can choose to become managers. Managers of ability $z \in Z$ transform inputs $l > 0$ of labor services and $k > 0$ of physical capital into the unique consumption good according to the following schedule:

$$F(l, k, z) = z \left[\alpha k^{\rho} + (1 - \alpha) l^{\rho}\right]^{\frac{\nu}{\rho}} + (1 - \delta)k,$$

where $\alpha \in (0, 1)$, $\nu < 1$ (to allow for managerial profits even when markets are complete), $\rho \in (-\infty, 1]$ measures the degree to which capital and labor can be substituted for each other, and $\delta \in (0, 1)$ is the rate of depreciation of physical capital.[6] In the last period of their life, agents do not work.

Agents' preferences over non-negative lifetime consumption profiles (c_1, c_2, c_3) are represented by the following utility function:

$$U(c_1, c_2, c_3) = \log c_1 + \beta \log c_2 + \beta^2 \log c_3,$$

where $\beta \in (0, 1)$. These preferences imply that real savings rates are similar in all the economies we compare in this paper, which is consistent with the evidence discussed by Hsieh and Klenow (2004).[7]

Managers can self-finance part of their capital using savings from the first period of their life. We denote by s the capital so invested by managers. They can also borrow $d \geq 0$ from an intermediary that can borrow and lend at gross rate $1 + r > 1 - \delta$. We will consider two scenarios for the determination of the intermediary's cost of capital.

In the first scenario, we assume that the intermediary can store deposits on behalf of agents with exogenous net return $r > -\delta$, and can borrow without bound at that same rate. We then

[6]Even though the plant technology is strictly concave in labor services and capital, aggregate returns are constant in this economy. Doubling population, all else equal, doubles the number of potential workers and managers. If population and the capital stock both double, so does output.

[7]These rates differ somewhat because life-cycle considerations are not the only motive for saving in this environment. Managers also save for the purpose of self-financing production. In our simulations, savings rates differ very little across economies.

think of the resources invested in the storage technology as capital not used in production.

The assumption that the intermediary can borrow (as well as invest) at rate r simplifies the exposition but has no impact on our quantitative results. Indeed, investment in the storage technology is positive in all our simulations given the calibration choices we make in the quantitative section. In fact, some key aspects of our quantitative results are sensitive to how one interprets these investments. When mapping our model to data from National Income and Product Accounts (NIPA), we will treat production activities carried out by managers as the production sector. We will think of savings channeled to the storage technology as investments outside of the production sector, which, in an economy such as the United States, is primarily residential investment. Note that this last interpretation is compatible with our specification of the storage technology since residential capital contributes linearly to GDP via imputed rents.[8]

In the second scenario, the interest rate is determined endogenously so that aggregate savings and capital used in production coincide. That is, there is no storage. In the quantitative section, we argue that these two alternative scenarios make very different predictions for differences in capital-output ratios and capital income shares across economies.

Let $w > 0$ denote the price of labor services. In this paper we will only consider equilibria where this price is constant over time. Given quantity $k = s + d \geq 0$ of physical capital, a manager of ability $z \in Z$ generates net income $\Pi(k, z; w, r)$ where

$$\Pi(k, z; w, r) \equiv \max_{l} \left[F(l, k, z) - wl - k(1 + r) \right].$$

The market for loans is imperfect. Specifically, managers can choose to default on their loan and economize on the payment they owe the intermediary, in which case they incur a default cost equal to fraction $\eta > 0$ of their income. Throughout this paper, we think of a manager together with some capital and some workers as a production establishment. In particular, we do not attempt to model different forms of ownership. In practice, opportunities to abscond

[8] Assuming a common opportunity cost of capital across economies is also appropriate for comparing economies with integrated capital markets, which is a reasonable approximation for many nations in our sample.

with resources are not the same for sole proprietors as for managers who operate establishments on behalf of a corporation. The simplest way to model this would be to introduce some heterogeneity in default costs across agents. We do not do so not only to keep computations manageable, but also because introducing ad-hoc heterogeneity is of little value. One would need to provide a theory for why different types of firms arise and coexist in equilibrium, which is beyond the scope of this paper. We think of the degree η to which contracts can be enforced as summarizing the average quality of enforcement in a given economy.

We further assume that the intermediary behaves competitively, so that among the individually rational loans that cover the opportunity cost r of funds, the most favorable to the manager prevails. Therefore, financial contracts for managers of ability $z \in Z$ with savings $a \geq 0$ solve:

$$\begin{aligned} & \max_{s \leq a, d \geq 0} && \Pi(s+d, z; w, r) \\ \text{s.t.} \quad & \Pi(s+d, z; w, r) + s(1+r) && \geq (1-\eta)\left[\Pi(s+d, z; w, r) + (s+d)(1+r)\right]. \end{aligned}$$

The constraint states that it must be individually rational for the managers to repay their loan. When they abide by the contract, managers receive their net income plus the accrued value of their own investment. When they default, they economize on the payment they owe the intermediary but lose fraction η of the resulting resources.[9]

Note that the simple demographic structure we employ reduces the manager's problem to a static optimal contracting problem. A multi-period version of this problem would greatly complicate computations, unless one resorts to the strong simplifying assumption that preferences are linear, as in, among many other papers, Cooley, Marimon, and Quadrini (2004). In this model, savings decisions are not trivial, and we find that they play an important role in our quantitative findings.

Let $d(a, z; \eta, w, r)$, $s(a, z; \eta, w, r)$ and $l(a, z; \eta, w, r)$ denote the policy functions associated with the manager's problem, and let $V(a, z; \eta, w, r) = \Pi\Big(s(a, z; \eta, w, r) + d(a, z; \eta, w, r), z; w, r\Big)$

[9]This specification assumes without loss of generality that the funds $a-s$ managers choose to invest outside of their establishment are not subject to the default cost. The third item of lemma 1 implies that this assumption does not have any effect on results.

be the resulting net income. When the loan market is perfect ($\eta = 1$), managers of ability z operate with the optimal quantity $k^*(z; w, r) = \arg\max_{k \geq 0} \Pi(k, z; w, r)$ of physical capital, given the prices of capital and labor. When $\eta < 1$, managers are constrained to operate at a sub-optimal scale, unless their savings exceed $a^*(z; \eta, w, r) = \inf\{a \geq 0 : s(a, z; \eta, w, r) + d(a, z; \eta, w, r) = k^*(z; w, r)\}$. In particular, $V(a, z; \eta, w, r) = \Pi(k^*(z; w, r), z; w, r)$ for all $a \geq a^*(z; \eta, w, r)$.

The lemma below states that under threshold a^*, the manager's access to outside financing rises with her assets and her managerial ability. This occurs because increases in a or z weaken the individual rationality constraint in the manager's problem. This result also states that it is optimal for borrowing constrained managers to invest all their savings in their establishment. This is because the marginal product of physical capital exceeds its opportunity cost $(1 + r)$ in establishments operated at a sub-optimal scale.

Lemma 1. *Given $w > 0$, $r > -\delta$ and $\eta > 0$,*

(i) $V(\cdot, z; \eta, w, r)$ *is strictly increasing and concave on* $[0, a^*(z; \eta, w, r))$ *for all* $z \in Z$*;*

(ii) $V(a, \cdot; \eta, w, r)$ *rises strictly for all* $a > 0$*;*

(iii) $s(a, z; \eta, w, r) = a$ *on* $[0, a^*(z; \eta, w, r))$ *for all* $z \in Z$*;*

(iv) $d(\cdot, z; \eta, w, r)$ *is strictly increasing and concave on* $[0, a^*(z; \eta, w, r))$ *for all* $z \in Z$*;*

(v) $d(a, \cdot; \eta, w, r)$ *is increasing for all* $a \in [0, a^*(z; \eta, w, r))$*.*

This result is established in the appendix. It implies that limited enforcement disrupts the allocation of resources in potentially two ways. First, establishments are generally operated below their optimal scale. Second, occupational choices depend not only on agents' managerial ability, but also on their wealth. However, we will now argue that even when contractual imperfections are present, occupational choices are monotonic in ability: agents whose ability exceeds a certain threshold become managers in the second period, while agents below that threshold remain workers. To that end, we need to state the problem solved by young agents:

$$\begin{aligned}\max_{a_1,a_2\geq 0} \quad & \log c_1 + \beta \log c_2 + \beta^2 \log c_3 \\ \text{s.t. } c_1 + a_1 &= w \\ c_2 + a_2 &= a_1(1+r) + \max\Big(w, V(a_1, z; \eta, w, r)\Big) \\ c_3 &= a_2(1+r).\end{aligned}$$

Solutions to this maximization problem need not be unique. Indeed, when $w = V(a_1, z; \eta, w, r)$, agents are indifferent between the two possible occupations in the second period. Yet, the level $k(z; \eta, w, r)$ with which managers of ability z operate is unique. To see this, recall that by lemma 1, V is strictly concave in a below $a^*(z; \eta, w, r)$. Therefore, agents of ability z who become managers are either unconstrained, or solve a problem that is strictly concave in a_1. Also, note that agents are indifferent between occupations for at most one ability value. In fact, given $w > 0$, $r > -\delta$ and $\eta \geq 0$, there exists a unique $\underline{z}(\eta, w, r)$ such that agents become managers in the second period of their life when $z > \underline{z}(\eta, w, r)$, remain workers when $z < \underline{z}(\eta, w, r)$, and are indifferent between the two occupations otherwise.

If agents of a given managerial ability find that becoming a manager maximizes their lifetime utility, this remains true for agents of higher ability. Indeed, productivity and access to outside financing both rise with managerial ability, by lemma 1. This simplifying result hinges on our assumption that agents are born identical. Sources of wealth inequality that are not monotonic in managerial ability, such as bequest inequality, would imply a different outcome.[10]

Given $w > 0$, $r > -\delta$ and $\eta > 0$, optimal agent policies are fully described by the threshold managerial ability $\underline{z}(\eta, w, r)$, the quantity $k(z; \eta, w, r)$ of capital employed by managers of

[10]Whether wealth inequality alleviates financial frictions or not depends, in part, on the correlation between first period wealth and managerial ability. In a world where initial wealth is positively correlated with ability, talented managers should be less constrained as they can rely more on self-financing. The opposite is true, of course, of economies where initial wealth and managerial ability are negatively correlated. In the data (see e.g. Quadrini (1999) and Gentry and Hubbard (2004)), entrepreneurial households (a proper subset of what we mean by managers) hold more wealth and have higher savings rates than non entrepreneurial households, which is consistent with our model.

ability $z \geq \underline{z}(\eta, w, r)$, and the quantity

$$l(z; \eta, w, r) = \arg\max_l F[l, k(z; \eta, w, r), z] - wl$$

of labor these agents choose to manage.

In the economy with storage, the market for capital clears trivially since by assumption capital is available perfectly elastically at price $1 + r$. Then, a steady-state equilibrium is a constant wage rate w such that, given the associated policies, the market for labor clears.

To make this formal, note that managers of ability $z > \underline{z}(\eta, w, r)$ have a net excess demand for labor of $l(z; \eta, w, r) - 1$ over their lifetime, since they supply one unit of labor when young. For their part, agents who do not become managers supply 2 units of labor during their lifetime. Agents of ability $\underline{z}(\eta, w, r)$ can be assigned to either occupations in the second period. Each possible fraction $\theta \in [0, 1]$ of those agents assigned to management generates a different value of the aggregate excess demand for labor. These observations lead to the following definition of the excess demand for labor correspondence, for all $w \geq 0$ and $\eta > 0$:

$$ED(w; \eta, r) = \left\{ \int_{z > \underline{z}(\eta, w, r)} [l(z; \eta, w, r) - 1] d\mu - 2 \int_{z < \underline{z}(\eta, w, r)} d\mu \right. \\ \left. + \mu\Big(\underline{z}(\eta, w, r)\Big) \Big[\theta\Big(l\Big(\underline{z}(\eta, w, r); \eta, w\Big) - 1\Big) - 2(1 - \theta)\Big] : \theta \in [0, 1] \right\}$$

A steady state equilibrium is a value for w such that $0 \in ED(w; \eta, r)$. We will call a steady state equilibrium with $w > 0$, hence positive output, *non-degenerate*. The following result provides conditions under which non-degenerate steady-state equilibria exist:

Proposition 1. *Given $r > -\delta$, there exist $\underline{\rho} \leq 0$ and $\underline{\eta} \leq 1$ such that a non-degenerate steady state exists provided:*

1. $\rho \geq \underline{\rho}$, *or,*

2. $\eta \geq \underline{\eta}$.

A non-degenerate equilibrium exists in the economy with storage provided that labor and capital are substitutable enough, or that the degree to which contracts can be enforced is high enough, for any given set of other exogenous parameters. In the appendix, we argue that a non-degenerate steady-state equilibrium exists if $ED(w; \eta, r) \cap I\!R_+ \neq \emptyset$ for w low enough, that is, if excess demand becomes positive when w becomes small. We also argue that this condition is met when $\rho \geq 0$ (which includes the Cobb-Douglas case, $\rho = 0$), but may fail to hold when $\rho < 0$.

Given the calibration choices we make in section 4, we find that the economy only collapses for extremely low values of η and ρ. There, we use numerical simulations to compare economies that differ in one respect only: the degree η to which contracts can be enforced. Intuitively, we should expect steady state wages to rise with the quality of enforcement. Indeed, as η rises more contracts become enforceable and managers' access to outside financing improves. Managers should then operate at a scale closer to the optimal one, with a positive effect on output and the marginal product of labor. One can in fact show that when $\rho \geq \underline{\rho}$, for each monotonic sequence of enforcement parameters there exists a corresponding sequence of monotonic steady state wages.

In economies with endogenous interest rates (that is, in economies without storage), both the labor and the capital market must clear. In that case, steady state equilibria may fail to exist even in the Cobb-Douglas case. To see this, consider first the extreme case where $\eta = 0$. Then, no borrowing can be supported in equilibrium and managers must self-finance all production. This implies that aggregate savings must exceed the demand for capital whenever $r > -\delta$ since all agents save positive amounts in the second period of their life to finance consumption during retirement.

Equilibria may still fail to exist when η is positive but less than one since savings are not monotonic in r in this environment. Indeed, when the interest rate is low enough, the borrowing constraint we impose on young agents ($a_1 \geq 0$) binds, leaving retirement consumption as the only motive for savings. Then, given our log-specification of preferences, savings are proportional to second period income. As r falls and capital becomes cheaper, labor income and profits rise, as do, therefore, savings. Section 4.3.1 deals with this issue by relaxing the borrowing constraint faced by young workers.

The point of the remainder of this paper is to gauge the quantitative importance of enforcement, and therefore access to finance, on an economy's output and capital intensity under various scenarios.

3 Data

In this section we establish the data benchmarks against which the results of the quantitative section will be compared. Our main source of data is the Penn World Table, Mark 6.2, which is described in detail in Heston, Summers, and Aten (2002). Our sample consists of those countries that report output per worker in 1994 and report investment data for at least 20 years leading up to 1994. This gives us an initial sample of 137 countries.

As a measure of output per worker we use the series RGDPWOK and label it $(Y/L)_D$. Since the capital per worker data is still unavailable at this time for Mark 6.2, we constructed capital stock series for each country using the perpetual inventory method. Our investment series is the product of GDP per capita in 1994 international prices (RGDPL) and the corresponding investment share series (KI). For the initial capital stock we consider two alternative methods. Method 1 sets the initial capital stock equal to the initial investment divided by the growth rate of investment (its geometric average over the sample) plus the depreciation rate (which we set to 10% a year.) This is tantamount to assuming that such an economy is on a balanced growth path in the initial period. Method 2 sets the capital-output ratio in the initial period equal to that in the last period (1994). This is a good approximation if the economy is close to a balanced growth path throughout the period considered. The two methods yield capital stocks (relative to the U.S) in 1994 that are very close to each other, so we choose the first one. Finally, we divide the capital stock in 1994, by GDP per capita (as given by RGDPL) to obtain capital-output ratios. We label the resulting series $(K/Y)_D$.

The first panel in figure 1 plots output per worker against capital-output ratios. Countries in the top decile are **35** times richer than countries in the bottom decile. On the other hand, their average capital-output ratio is only **2.6** times higher than among countries in the bottom decile. These numbers have prompted many development economists to conclude that the dispersion in capital-output ratios across countries cannot account for observed differences in output per worker.[11]

A standard development accounting exercise would in fact lead one to conclude that total factor productivity varies greatly across countries. The standard approach would measure productivity as:

$$\text{Measured Productivity} = \frac{(Y/L)_D^{(1-\alpha)}}{(K/Y)_D^{\alpha}},$$

where α is the capital share. Assuming $\alpha = 0.3$, average measured productivity among countries in the top decile in terms of GDP per worker is roughly **8.5** times higher than among countries in the bottom decile. This statistic does not have the usual interpretation in our model economy, since, in general, aggregate technological opportunities are not well described by a Cobb-Douglas specification. Nonetheless, we report it in all quantitative experiments so that our simulated numbers can be compared to the outcome of standard development accounting exercises.

Measuring the quantity of financial intermediation in a given economy is a more complicated issue. There is a large literature that studies the statistical relationship between proxies for financial intermediation and growth or development.[12] The ideal proxy for our purposes would measure the ability of a manager to obtain financing for a project. In our model economy, *intermediated capital* is the part of capital that does not come from managers' savings, the variable we termed d in the previous section. For sole proprietorships, a natural empirical counterpart for this variable are the funds proprietors borrow from financial institutions. On the other hand, managers who operate an establishment on behalf of a corporation presumably finance a negligible fraction of the resources they manage.[13] In this sense, all but a negligible

[11] This remains true even if one accounts for human capital differences. See for example Prescott (1998) and Hall and Jones (1999).

[12] See Levine (2005) for a review.

[13] In that case, the corporation plays the role of the financial intermediary in our model. In principle, managers can accept a rising compensation profile which amounts to advancing resources to their employer. Quantifying the exact importance of these resources is challenging for obvious reasons, but they are not likely to amount to a large fraction of the capital stock of corporations.

fraction of the capital owned by corporations is intermediated and financed via debt and equity issues.

Correspondingly, and given the set of financial statistics that are available for a reasonably large cross-section of countries, we choose to measure the stock of intermediated capital as the sum of (i) the credit firms obtain from banks and other financial institutions (ii) the credit they obtain from issuing bonds and, (iii) the funds they obtain from issuing stock. The database compiled by Beck, Demirgüç-Kunt, and Levine[14] contains a market value measure of (ii) and (iii), which we use for lack of historical value numbers.[15] The database also provides a measure of total private credit by financial institutions; we subtract household debt from this measure to arrive at a proxy for (i).[16]

The ratio of intermediated capital (so defined) to output in the United States in 1994 was roughly 2.2, which suggests that most capital is intermediated in the U.S. This is consistent with the fact that corporations are the leading form of ownership in the U.S. Using unrelated data and a different methodology, McGrattan and Prescott (2000) arrive at a ratio of 1.8 for the year 2000. Their number is smaller in part because they ignore corporate debt holdings and liabilities.

Figure 1 shows output, capital-output ratios, and measured productivity plotted against intermediated capital to output ratios relative to U.S. values for the 25 countries for which we were able to compute all our variables in 1994. The country with the lowest intermediated capital to output ratio relative to the U.S in this sample is Hungary, at 0.095, while the one with the highest is Japan at 1.05. As is well known, the quantity of finance, the level of development, capital-output ratios, and total factor productivity measures are all positively correlated.

[14]This database is available at http://www.econ.brown.edu/fac/Ross_Levine/Publications.htm.

[15]Whether historical valuations of outstanding stocks are more appropriate for our purposes than market valuations is not obvious, however. Stock values reflect in part the value of retained earnings. These are funds which corporations effectively borrow from their shareholders.

[16]The data for household debt was obtained from national sources and is available upon request.

4 Quantitative Experiments

In our model economy, finance affects development in two ways. First, economies with better financial markets direct more capital to the production sector. Second, they direct capital to more efficient uses (more productive plants). In this section we conduct a series of experiments to measure the quantitative importance of these two channels, and to find out whether financial differences generated by enforcement frictions can account for observed differences in income per worker across countries. We are particularly interested in asking whether the model can generate a lot of dispersion in output while keeping the dispersion of capital-output ratios small.

We begin by studying a benchmark economy with storage where managers and the resources they employ are interpreted strictly as the production sector. Next, we consider an economy where managers are the only source of demand for capital and interest rates adjust to clear capital markets.

4.1 Calibration

Our basic strategy is to set parameters to match salient features of the U.S. economy. We follow standard practice whenever possible, but existing work provides little guidance for calibrating the distribution μ of managerial talent. We will calibrate it via its implications for the employment-size distribution of establishments.

Assuming that an individual's work life lasts 40 years, a period in the model economy corresponds to 20 years. We set the yearly interest rate (the yearly net return to storage) to 4% ($r = 1.04^{20} - 1$) and set $\delta = 0.88$, which implies a yearly depreciation rate of 10%.

To calibrate the parameters governing the production function we begin by assuming an elasticity of substitution of $\sigma = \frac{1}{1-\rho} = 1$. In that case, $F(l, k, z) = zk^{\alpha\nu} l^{(1-\alpha)\nu} + (1 - \delta)k$ for all l, k, $z \geq 0$. When there are no contractual imperfections ($\eta = 1$), the share of value added (Y) in the production sector that accrues to capital is $\frac{K(r+\delta)}{Y} = \alpha\nu$, where K denotes the sum across plants of all capital, while $1 - \nu$ is the share of value added that accrues to managerial services.

Let A denote aggregate savings, i.e. the sum of savings by all agents in the first two periods of their life. In all our simulations with storage, $A > K$ and the proper conceptual counterpart for GDP is $Y + (A - K)(r + \delta)$ – the sum of the value added by the business sector and the gross returns to storage – while capital income is $A(r + \delta)$.[17] Note that in these calculations we assume that the resources invested in the storage technology are treated as investment in National Income and Product Accounts, which is consistent with our interpretation of $A - K$ as (primarily) residential investment.

We follow Atkeson, Khan, and Ohanian (1996) and Atkeson and Kehoe (2001) (among others) and set $\nu = 0.85$ for the degree of strict concavity of the production function. Appendix B.1 provides some additional justification for this value using Internal Revenue Service data for sole proprietorships in the U.S. However, these calculations rely on a number of strong assumptions, and we will verify in section 4.3.3 that our results are robust to broad variations in this parameter. As for the capital share, we make $\alpha\nu = 0.30$, which yields a capital income to GDP ratio of roughly 35% in the benchmark economy.[18]

Next, we set the discount rate, β, so that the ratio of storage investment $(A - K)$ to total investment A in our benchmark economy is roughly 40% which is consistent with US evidence on the stock of housing capital for the post-war period (available from the Bureau of Economic Analysis) and our interpretation of $A - K$ as primarily residential investment.

This leaves us with two objects to calibrate: μ, the distribution of managerial talent, and η, the degree to which contracts can be enforced in the benchmark economy. All else equal, managers with more talent hire more workers, and access to intermediated capital improves when η rises. Correspondingly, we calibrated μ and η together to match the ratio of intermediated capital to GDP in 1994 (calculated using the data and method described in the previous section) and two moments of the labor size distribution of manufacturing establishments in the United States.

Specifically, we assume that μ is a discretized version of a log-normal distribution[19] and

[17]Recall that r is the *net* return to storage.

[18]Due to the presence of contractual imperfections, the capital share is somewhat below 30% in the production sector, but this is offset by the returns to storage.

[19]This implies a distribution of establishment sizes that is approximately log-normal as well, since labor is

calibrate the distribution's two parameters to match 1) the percentage of manufacturing establishments with 9 employees or fewer, which is around 50% according to County Business Patterns Survey data between 1988 and 1998, and 2) the average size of manufacturing establishments in the United States, which is around 50 employees according to the same source. This is meant to capture two salient features of the organization of production in the United States: the majority of establishments are small, but small establishments account for a low fraction of employment.

We initially use manufacturing data to facilitate international comparisons. Data on the organization of production outside of the manufacturing sector are seldom available for developing nations, and are unreliable when they exist, given the preponderance of undeclared, informal activities in services in those countries. We will also present results when the managerial shares and the distribution of talent are calibrated to match the same targets for all sectors as opposed to simply the manufacturing sector.

Appendix B.2 describes the joint calibration of parameters in more detail. The exact values we used for all exogenous parameters in our experiments are in table 1.

4.2 Results

Our experiments consist of comparing steady state equilibria in economies that differ in their degree of contractual imperfections, indexed by η, and thus in their intermediated capital to output ratio. In these comparisons, we focus on output, capital-output ratios, and measured productivity, but we also highlight some subsidiary implications that lend credibility to the framework we propose.

Figure 2 plots steady state statistics against the intermediated capital to output ratio relative to their benchmark values. Alternatively, one can plot these values against η. We chose to plot variables against the intermediated capital to output ratio because it rises monotonically with η and, unlike η, it has a natural empirical counterpart in the available cross-country data. Furthermore, our primary concern in this paper is the relationship between financial and

linear in $z^{\frac{1}{1-\nu}}$ when $\eta = 1$.

economic development.

Figure 2 presents results for the whole economy and for the production sector alone. Not surprisingly, the model is *qualitatively* consistent with the empirical correlation between finance, output, and measured productivity.

The model also correctly predicts a positive correlation between average establishment size and the intermediated capital to output ratio.[20] In fact, the model's quantitative predictions for the organization of production are reasonable given available data. For instance, we calculate Argentina's intermediated capital to GDP ratio to be 8 times lower than in the United States in 1994. The model predicts that in such an economy the average size of establishments should be roughly 4 times lower than in the United States, or around 12.5 employees. According to Argentina's economic census, the average size of manufacturing establishments was around 11 employees in 1993. Mexico's intermediated capital to GDP ratio was about 30% that of the U.S. in 1994.[21] According to its economic census, manufacturing establishments counted 12 employees on average in 1993, while our model predicts roughly 22.[22] Finance disrupts the organization of production in our model by magnitudes that appear reasonable given available data.

Quantitatively, the model also predicts a large impact of financial differences on output. As figure 2 shows, economies where all production is self-financed have 1/3 the output of the benchmark economy. In the sample of countries for which we have data, the two countries with the lowest finance to GDP ratio in 1994 are Hungary and Argentina. Our model implies that these countries should have a relative output per worker of around 50% and 60% respectively. In reality, it was 36% for Hungary and 49% for Argentina. In our model therefore, financial differences alone can account for much of the output differences between the U.S. and middle-income nations.

[20] Average establishment size is a step function of the finance-to-output ratio because the distribution of managerial talent is discrete. Jumps occur where agents of a given managerial talent become indifferent between becoming workers and managers in the second period of their life.

[21] Since 1994 marks the peak of a lending boom, Mexico's intermediation ratio in 1994 is an outlier. Mexico's ratio is near Argentina's for most years other than 1993-1994.

[22] See Tybout (2000) for more quantitative evidence on the organization of production in developing countries. Other emerging nations for which we have both census and financial data typically exclude from their counts establishments below a certain size threshold, which prevents one from calculating an average size.

Due to data limitations, our sample does not contain many of the lowest income nations in the world. But it is clear that our model's maximum income differences of 3 cannot possibly account for development disasters. Nor do we believe disruptions to the intermediation process to be the main problem for these nations.

Because savings rates are similar across economies, capital-output ratios do not change much either as the quality of enforcement drops. In our model, this implies that capital income shares vary little across economies, which is consistent with the evidence presented by Gollin (2002). Since capital intensity varies little across economies, measured TFP, calculated as in section 3, falls markedly as enforcement weakens.

What distinguishes high enforcement economies from low enforcement economies in our model is their ability to channel capital to the production sector. In fact, production sector capital-output ratios vary almost one-for-one with the finance to output ratios. These large differences in production sector capital are broadly consistent with the results of DeLong and Summers (1991) who find that equipment investment rates vary substantially more across countries than overall investment rates. As a consequence of these capital intensity differences, TFP measured using production sector data only varies less than overall TFP. It does fall a bit because resources are misallocated within the production sector, since, in economies where enforcement is poor, marginal products vary greatly across plants, and less productive technologies need to be activated for labor markets to clear. In economies with finance to output ratios similar to nations such as Argentina or Hungary, production sector TFP falls about half as much as overall TFP.

Our model of finance, therefore, improves upon standard development models by generating output dispersion in a context where capital intensity differences are mostly absent. We view this as a step in the right direction in a literature where this is exactly the challenge – differences in output that are primarily associated with differences in measured productivity. We now turn to studying how these results depend on the calibration choices and assumptions we have made so far.

4.3 Sensitivity analysis

4.3.1 Endogenous interest rates

Consider a version of our economy where there is no storage and, therefore, all capital must be used by the production sector. In that case, interest rates must adjust for capital markets to clear. It should be intuitively obvious that the cost of capital will tend to fall when enforcement weakens. Indeed, holding prices constant, a drop in enforcement reduces the demand for capital hence creates an excess supply of capital.

As we discussed in section 2 however, equilibria may not exist in this environment because savings are not monotonic in interest rates due to the borrowing constraint young workers face. To circumvent this issue and to guarantee that equilibria can be found for all possible values of η, we chose to relax the borrowing constraint for young agents who choose to remain workers their entire life. This amounts to assuming that workers can commit to repaying their debt in the second period of their life.

This approach has the advantage of only affecting results in economies where existence is an issue in the first place, that is in economies where η is not high enough. Possible alternatives would entail introducing features such as agents with exogenous endowments, but this would make comparing results across scenarios more problematic since it would introduce additional productive resources.

To make results comparable across experiments, our calibration choices must also be adjusted. Since managers are the sole source of value added and income in the environment without storage, the proper counterpart to GDP becomes Y while the ratio of capital income to GDP is $\frac{K(r+\delta)}{Y}$. Moreover, since managers now account for all contributions to GDP (including returns to residential investment), we need to match the total ratio of intermediated capital (including residential loans) to GDP. Finally, the distribution of managerial talent must be adjusted to continue matching the same size distribution moments as before. Table 1 shows the specific parameters we used to meet the corresponding targets.

The outcome of the resulting experiment is displayed in figure 3. The model with endogenous interest rates produces a pattern for output and capital-intensity that resembles the behavior of the production sector in the economy with storage. The impact of financial differences on output is greater than in the economy with storage, but capital-intensity and the capital-income ratio vary much more than they do in the data or in our benchmark experiment.

Note that the new distribution of managerial talent results in an average size curve with fewer kinks because it puts more mass on agents with high managerial talent. Augmenting the number of grid points near the upper-bound of Z produces a curve more similar to our previous experiments, but we chose to leave the discretization method unchanged for ease of comparison across experiments.

In this environment, the adjustment to financial frictions takes the form of lower interest rates rather than a smaller fraction of productive resources being directed to the production sector. This model does not capture the possibility that in economies where capital is misallocated in the production sector, hence less productive, a larger fraction of investment is directed to other uses.

4.3.2 Distribution of managerial talent

Figure 4 shows the impact of matching moments for the entire distribution of establishments in the U.S. as opposed to moments for the manufacturing sector alone. According to County Business Patterns Survey data between 1988 and 1998, the fraction of establishments with 9 employees or fewer is roughly three quarters across all sectors in the United States, while the average size of establishments is roughly 15 employees. Naturally, our managerial rent target needs to be similarly adjusted. Using Internal Revenue Service data for all sectors between 1989 and 1998 and the same calculations as in appendix B.1 yields a managerial share of approximately 23% instead of 15%. We adjust α correspondingly to continue matching our target for the ratio of capital income to GDP (see table 1.) Figure 4 shows that the broad nature of results changes little under this new calibration approach.

Since results appear robust to even large changes in the distribution of managerial talent, one may wonder whether heterogeneity in talent matters at all. To highlight the critical importance of this aspect of our model, consider a version of our economy in which agents all have the same managerial talent.[23] In this environment it is no longer the case that economies with lower finance to output ratios employ more inefficient technologies, as all managers are equally productive.

Figure 5 shows the quantitative effects of finance in this economy. Production sector output falls by at most a quarter, and this drop is offset by returns to storage. Overall, output falls very little (if at all) as enforcement worsens.

In part, this is because all plants are now equally productive. However, the main reason for the near-constancy of output is more subtle. Because agents are homogenous, they must be indifferent between working and managing in the second period of their life (unless all agents choose to become managers in the second period, which does not happen in any of our experiments given our calibration choices.) This means that all managers have a flat lifetime income profile. In economies with heterogenous agents on the other hand, only marginal managers have flat income profiles, while other managers earn rents and have steeper income profiles. Because steeper income profiles reduce the willingness of managers to transfer resources into the second period, the savings rate of managers is much higher in economies with homogenous agents than in our benchmark economy (at least twice as high in all simulations given our calibration choices.) As a result, a much greater fraction of capital is self-financed in the homogenous agent case, which means that finance has a smaller impact on capital intensity in the production sector.

All told, without managerial heterogeneity, our model generates little to no output dispersion.

[23]Naturally, such an economy delivers very counterfactual predictions for the organization of production. First, all establishments now have the same size. Second, and less obviously, establishments must now be small, even when there are no contractual imperfections. To see this, note that if all agents become managers in the second period of their life, the size of all establishments in equilibrium is one employee. Bigger establishments arise in equilibrium if, and only if, some agents become workers in the second period of their life, while others become managers. In that case, agents must be indifferent between working and managing. Simple manipulations of first-order conditions for profit maximization then imply that the unique equilibrium size in that case is given by the ratio of the labor share to the managerial share, which is roughly 4 employees given our calibration choices. These considerations have critical implications for the impact of finance in this version of the model. Since establishments must be small, many agents must become managers in the second period of their life. In turn, as we explain below, high savings rates among managers in economies with low enforcement have a determinant effect on overall output.

4.3.3 Technological parameters

This section considers the effect of varying technological parameters. We carried out two types of experiments. In the first type, all parameters but the one being considered are left at their benchmark values. In the second, we adjust the distribution of talent, the benchmark enforcement level, and the discount rate to continue matching the same calibration targets as before. For conciseness, we will only discuss results for the first type of experiment. Results for both exercises are almost identical in all cases.

Since the managerial share is not a standard parameter, it is important to verify that our results do not hinge on the value of ν we used in section 4.1. Figure 6 shows the effect of dropping the managerial share from 15% to 10%, and of raising it to 20%, holding the capital share constant.[24] It can readily be seen that this has little effect on results.

Two calibration choices that, on the other hand, have a big impact on our results are the capital share and the elasticity of substitution between capital and labor. Gollin (2002) finds that cross-country capital shares vary between 0.2 and 0.35. Figure 7 shows the results for capital shares of $0.4 = 0.47 \times 0.85$ and $0.2 = 0.24 \times 0.85$ respectively (in the benchmark case we used $0.3 = 0.35 \times 0.85$), while managerial shares are kept constant, and labor shares are adjusted to make up for the difference. The high capital share generates nearly twice as much dispersion in output as in the benchmark parameterization.

Lowering the degree of substitution between capital and labor also leads to more output dispersion. Figure 8 shows the effect of lowering σ to 0.75 and 0.5. As the elasticity of substitution falls, the dispersion in output rises to become comparable to what we observe in the data. Yet, the capital-output ratio dispersion does not differ much from that in the Cobb-Douglas case. Consequently, most of the extra variation in output is reflected in measured productivity, as in the data. With elasticities of substitution between $\sigma = 0.75$ and $\sigma = 0.5$, the model can easily account for the observed dispersion of output and measured productivity across countries.

That lowering the capital share and the degree to which capital and labor can be substituted

[24] In other words, we adjust α so as to leave $\alpha\nu$ unchanged. Keeping α at 0.35 yields very similar results.

for each other leads to more variation in output is well-known (see Caselli (2003) for an extensive discussion of this topic.) If one makes capital more important in production, a given level of variation in capital intensity implies more variation in output for obvious reasons.

Our results thus confirm the importance of correctly measuring the parameters that describe technological opportunities. For now, those parameters remain the cause of much debate.[25] It is not our objective to settle these empirical questions, but our results stress their importance.

In addition, finance changes the relationship between capital intensity and output for independent reasons. Modeling the allocation channel explicitly increases the predicted ratio of output variation to capital intensity variation. This implies that the technological parameter changes one needs in order to generate all the observed dispersion of output across countries are smaller. Put another way, a given change in the elasticity of substitution raises output variation by a larger amount when financial disruptions are explicitly modeled.

4.3.4 Complete markets in the benchmark economy

Our benchmark calibration of $\eta = 0.57$ to match the ratio of financial intermediation to GDP in the U.S. implies that upon default, managers would avoid roughly 40% of the payment they owe the intermediary and lose about 60% (or 12 years worth) of their second period income. Figure 9 shows that our results are not sensitive to this particular aspect of our calibration by plotting the same steady state statistics as before under the assumption that markets are complete ($\eta = 1$) in the benchmark economy so that no manager is constrained. This is because our calibration strategy already implies punishment costs such that the benchmark economy's steady state is near the complete market outcome.

5 Conclusion

Our quantitative experiments attempt to fill what we think is a gap in the literature. Empirical economists have emphasized the strong statistical relationship between financial development

[25] See e.g. Restuccia and Urrutia (2001), Pessoa, Pessoa, and Rob (2003), Antràs (2004), and Duffy and Pappageorgiou (2000).

and economic development. A number of theories have been developed that establish a qualitative connection between finance and development. We quantify the effects of differences in the quantity of financial intermediation on output and productivity in the context of a dynamic general equilibrium model, and find these effects to be significant.

In our model, better financial markets raise output by increasing the capital used in production and by helping direct capital to its best uses. Our calibrated exercises suggest that both channels are quantitatively important. We also show that the quantitative impact of financial frictions on output depends on the importance of capital in production. Lowering the elasticity of substitution between capital and labor or raising the capital share greatly increase the consequences of financial differences.

Our model could be extended to include other potential functions of the financial system. In particular, we ignore the role the financial systems plays in alleviating informational imperfections.[26] Whether modeling these will alter our basic quantitative findings is unclear. Our analysis already encompasses the two extreme cases of financial development: complete markets and pure self-financing.

Moreover, in order to quantify the importance of informational frictions, they must first be calibrated. We believe observed differences in the organization of production to be the natural disciplining benchmark when quantifying disruptions in the allocation of productive resources. A different friction or set of frictions calibrated to produce similar differences in the organization of production across economies as in our model seems unlikely to produce very different output effects.

However, determining the frictions that cause differences in financial development is clearly important. Our experiments suggest that these efforts will enhance our understanding of why some nations are so much richer than others.

[26] For an exercise with asymmetric information, see Erosa and Hidalgo (2004).

A Proofs

A.1 Proof of lemma 1

Rewrite the managers' problem as:

$$\begin{aligned} V(a,z;\eta,w,r) &= \max_{s\leq a, d\geq 0} \Pi(s+d,z;w,r) \\ s.t. &\quad \Phi(s,d,z;\eta,w,r)\geq 0, \end{aligned}$$

where $\Phi(s,d,z;\eta,w,r) \equiv \eta\left[\Pi(s+d,z;w,r)+s(1+r)\right]-(1-\eta)d(1+r)$. As in the text, let $s(a,z;\eta,w,r)$ and $d(a,z;\eta,w,r)$ be the solutions to the above problem and define $a^*(z;\eta,w,r) = inf\{a : s(a,z;\eta,w,r)+d(a,z;\eta,w,r) = k^*(z;w,r)\}$. If $a^*(z;\eta,w,r)=0$, the lemma holds trivially. So assume that $a^*(z;\eta,w,r)>0$ and fix $a<a^*(z;\eta,w,r)$. Start with item (ii). From the envelope theorem, $V_2 = \Pi_2+\lambda\Phi_3>0$, where $\lambda\geq 0$ is the Lagrange multiplier associated with the incentive compatibility constraint. Next, necessary and sufficient[27] conditions for a solution to the constrained maximization problem above are:

$$\Pi_1(s+d,z;w,r)+\lambda\Phi_2(s,d,z;\eta,w,r) = 0 \tag{A.1}$$

$$\lambda\Phi(s,d,z;\eta,w,r) = 0 \tag{A.2}$$

$$\Pi_1(s+d,z;w,r)+\lambda\Phi_1(s,d,z;\eta,w,r)-\nu = 0 \tag{A.3}$$

$$\nu(a-s) = 0 \tag{A.4}$$

where $\nu\geq 0$ is the multiplier associated with the constraint $s\leq a$. (A.1) and (A.3) imply $\lambda(\Phi_1-\Phi_2)=\nu$. Since $a<a^*(z;\eta,w,r)$, the manager is borrowing constrained, i.e. $\Pi_1(s+d,z;w,r)>0$ at the a solution. But A.1 then implies $\lambda>0$ and $\Phi_2<0$. Differentiation of Φ shows that $\Phi_1>0$ when $\eta>0$. Since $\Phi_2<0$, it now follows that $\nu>0$ for constrained agents, hence $s=a$ by (A.4). This establishes item (iii) of the lemma.

To show part (iv), note that for constrained agents $\Phi(a,d(a,z;\eta,w,r),z;\eta,w,r)=0$ in a neighborhood of a. Differentiating with respect to a yields $\Phi_1+\Phi_2\frac{\partial d}{\partial a}=0$, hence $\frac{\partial d}{\partial a}=-\frac{\Phi_1}{\Phi_2}>0$ because $\Phi_1>0$ and $\Phi_2<0$ for constrained agents. For strict concavity, differentiating with respect to a once more gives

$$\frac{\partial^2 d}{\partial a^2} = -\frac{\Phi_2\left(\Phi_{11}+\Phi_{12}\frac{\partial d}{\partial a}\right)-\Phi_1\left(\Phi_{21}+\Phi_{22}\frac{\partial d}{\partial a}\right)}{\Phi_2^2}$$

But differentiation of Φ shows that $\Phi_{12}=\Phi_{11}=\Phi_{22}=\eta\Pi_{11}<0$ for constrained agents. So $\frac{\partial^2 d}{\partial a^2}<0$, as claimed.

For item (i), use the envelope theorem to obtain $V_1(a,z;\eta,w,r)=\nu>0$,. To show strict

[27] These conditions are sufficient because Φ is concave. The functional forms we have assumed for the production function and the incentive compatibility constraint imply that d is interior for constrained agents.

concavity, note that $V_{11}(a, z; \eta, w, r) = \frac{\partial \nu}{\partial a}$. But, from (A.3),

$$\frac{\partial \nu}{\partial a} = \Pi_{11}\left(1 + \frac{\partial d}{\partial a}\right) + \frac{\partial \lambda}{\partial a}\Phi_1 + \lambda\left(\Phi_{11} + \Phi_{12}\frac{\partial d}{\partial a}\right).$$

All we need to insure is that $\frac{\partial \lambda}{\partial a} < 0$. By (A.1),

$$\frac{\partial \lambda}{\partial a} = -\frac{\Pi_{11}\left(1 + \frac{\partial d}{\partial a}\right)\Phi_2 - \left(\Phi_{21} + \Phi_{22}\frac{\partial d}{\partial a}\right)\Pi_1}{\Phi_2^2} < 0.$$

Finally, to establish item (v), differentiating the incentive compatibility constraint w.r.t z yields $\Phi_3 + \Phi_2 \frac{\partial d}{\partial z} = 0$. Thus, $\frac{\partial d}{\partial z} = -\frac{\Phi_3}{\Phi_2} > 0$, as $\Phi_3 > 0$, and $\Phi_2 < 0$ for constrained agents.

A.2 Proof of proposition 2

First note that $ED(\cdot; \eta, r)$ is upper-hemicontinuous, non-empty and convex-valued for all $\eta \in [0, 1]$. Indeed, by the theorem of the maximum, the set of optimal policies for agents of each managerial ability z is non-empty and varies upper-hemicontinuously with w. Because μ has finite support, and integrals in the definition of ED are finite sums, it follows that ED is non-empty and upper-continuous. Since different agent of ability $\underline{z}(\eta, w, r)$ can be assigned to different occupations, ED is also convex-valued. Now, for w high enough $ED(w; \eta, r) \subset I\!R_-$ for all $\eta \in [0, 1]$ since μ has bounded support. A standard application of Kakutani's fixed point theorem now implies that a positive steady state exists provided $ED(w; \eta, r) \cap I\!R_+ \neq \emptyset$ for some $w > 0$ when $\rho \geq 0$. That this is true when $\rho = 1$ (the linear case) is obvious. So it is sufficient to show that when $\rho \in [0, 1)$, $\lim_{w \to 0} d(0, z; \eta, w, r) = +\infty$ for all $\eta > 0$ and $z > 0$. Indeed, this implies that $l(z; \eta, w, r)$ grows without bound while $\underline{z}(\eta, w, r)$ goes to zero as w becomes small. To see this, notice that for all (z, η, w, r), $d(0, z; \eta, w, r)$ is the unique solution to:

$$d = \frac{\eta}{(1 - \eta)(1 + r)}\Pi(d, z; w, r).$$

When $\rho \in [0, 1)$, simple algebra shows that $\frac{\Pi(d,z;w,r)}{d}$ diverges to $+\infty$ as w becomes small for all $z \in Z$ and $d > 0$. (When $\rho < 0$, this is no longer true.) So, whenever $\eta > 0$ and $\rho \geq 0$, $\lim_{w \to 0} d(0, z; \eta, w, r) = +\infty$ and we are done.

B Calibration details

B.1 Managerial share

The degree $1 - \nu$ of strict concavity of the production function is monotonically related to the ratio of managerial income to gross output in our model. One could set it so that this ratio matches the empirical share of proprietors' income in GDP (which is around 10%) in the United States, but this would not be consistent with our theory. We think of F as an establishment technology, and our model does not distinguish between forms of ownership. In that sense, aggregate returns to managerial services should include the income of managers who operate an establishment on behalf of a corporation. Since corporate data do not allow one to measure the share of income that accrues to managerial services, we assume that the repartition of income across factors is independent of the ownership type.

Under that assumption, and under the assumption that in sole proprietorships the owner plays the role of the manager in our model, a good proxy for the managerial share is the ratio of net income to the sum of payments to physical capital, labor, and the owner in sole proprietorships, after correcting for the fact that part of net income rewards the owner's capital input.

For sole proprietorships, gross payments to capital, labor and the sole proprietor are business receipts minus all payments to intermediate inputs plus taxes paid. We calculate payments to intermediate inputs as Cost of Sales and Operations net of Cost of Labor + Supplies + Travel + Utilities + Advertising. Taxes paid need to be allocated to the three factors. Sole proprietors may deduct state and local income taxes, sales taxes, employment taxes, one half of their self-employment tax, personal property, and real estate taxes. We assume that these taxes accrue to the various sectors in proportion to overall factor shares. Then the ratio of payments to the sole proprietor to payments to other inputs can be approximated by the ratio of net income to business receipts minus payments to intermediate inputs minus taxes paid. Manufacturing sole proprietorship data available from the IRS then yields:

Year	(1) Business Receipts	(2) Payments to intermediate inputs	(3) Taxes paid	(4) Net Income	(4) / [(1)-(2)-(3)]
1989	25,400,029	10,499,808†	332,950	3,228,762	0.222
1990	21,839,350	7,957,673†	343,016	2,467,377	0.182
1991	23,354,542	8,920,086†	360,103	2,595,448	0.184
1992	27,243,502	10,934,325†	579,542	3,508,402	0.223
1993	27,157,994	11,228,291	559,901	3,216,585	0.209
1994	32,928,845	13,966,846	649,892	3,927,951	0.215
1995	32,101,683	13,662,657	617,056	4,143,519	0.232
1996	32,057,221	13,247,790	592,845	4,377,188	0.240
1997	32,057,221	13,425,659	593,115	3,941,135	0.218
1998	27,327,211	12,988,688	430,230	3,608,920	0.259

Note: All figures in thousands of dollars. Data are from Internal Revenue Service, Sole Proprietorship Tax Statistics - Manufacturing Sole Proprietorships - General (http://www.irs.gov/taxstats/article/0,,id=96754,00.html.)

† The IRS did not provide the cost of supplies between and 1989 and 1992. We impute it using its average ratio to receipts during the other six years.

Therefore, among manufacturing sole proprietorships in the United States, net income represents (on average) 22% of payments to labor, capital and the owner. Assuming that 30% of net income rewards the owner's capital input (30% is the capital income share in the production sector given our calibration under complete markets) implies then that $1 - \nu = 1 - 0.22 \times 0.70 \approx 0.15$.

Repeating similar calculations for all sectors rather than simply the manufacturing sector yields an average net income to value added ratio of roughly one third which, assuming a capital share of 30% yields a managerial income share of roughly 23% across all sectors, the value we use in section 4.3.2.

B.2 Distribution of managerial talent

Given all parameters, steady state equilibria can be computed using standard techniques. For each set of exogenous parameters, we guess a wage, compute optimal policies for all agents by backward induction, and then update our wage guess until the labor market approximately clears. In all our quantitative experiments, the approximate equilibrium wage proved unique (specifically, the aggregate excess demand for labor function we computed proved monotonic). In this section we make more precise the procedure we use to calibrate the parameters of the distribution of managerial talent, and the degree η to which contracts can be enforced.

We assume that $Z = [0, 30]$ where the width of the domain for managerial talent is a normalization that makes steady state wages near 1 in the benchmark economy. Then we assume that $\frac{z}{30}$ is log-normally distributed with location parameter λ_1 and dispersion parameter $\lambda_2 > 0$. Given λ_1 and λ_2, we use a discretized version of the resulting distribution. Specifically, we assign mass to the set $\{z_i = 30\left(\frac{i}{100}\right)^2 : i = 1, 2 \ldots 100\}$. In particular, the density of mass points is higher near the origin, as we found that doing this improved the precision of the algorithm. Letting $F(\lambda_1, \lambda_2, \cdot)$ denote the cumulative distribution function of the log-normal distribution with parameters (λ_1, λ_2) we set

$$\mu(z_i) = F\left(\lambda_1, \lambda_2, \frac{z_{i+1}/30 - z_i/30}{2}\right) - F\left(\lambda_1, \lambda_2, \frac{z_i/30 - z_{i-1}/30}{2}\right) \text{ for all } i,$$

where $z_0 = 0$, $z_{101} = z_{100} + \frac{z_{100}-z_{99}}{2}$, and μ is re-scaled as necessary so that $\mu\{Z\} = 1$.

Holding all other parameters fixed, denote by $\bar{n}(\lambda_1, \lambda_2, \eta)$, and $\phi(\lambda_1, \lambda_2, \eta)$ the average size of establishments and the fraction of employment in establishments with fewer than 50 employees in equilibrium given λ_1, λ_2, and η. Let $\eta(\lambda_1, \lambda_2)$ be the value for η such that the equilibrium ratio of loans to gross yearly output is 2 given λ_1 and λ_2. Our calibration objective is to choose λ_1 and λ_2 so that:

$$\bar{n}(\lambda_1, \lambda_2, \eta(\lambda_1, \lambda_2)) = 50 \tag{B.1}$$

$$\text{and } \phi(\lambda_1, \lambda_2, \eta(\lambda_1, \lambda_2)) = 0.50 \tag{B.2}$$

We meet those goals using a downhill simplex method[28] to minimize:

$$[50 - \bar{n}(\lambda_1, \lambda_2, \eta(\lambda_1, \lambda_2))]^2 + [50 - 100\phi(\lambda_1, \lambda_2, \eta(\lambda_1, \lambda_2))]^2$$

In all our experiments, we were able to meet our two calibration targets almost exactly. Furthermore, even drastic changes in initial conditions (the coordinates of the initial simplex) led to the same results suggesting that despite the complexity of the mapping between parameters and $\bar{n}$ and ϕ, solutions to (B.1-B.2) are unique.

[28]Our C routine uses Amoeba.c from Numerical Recipes.

The following table shows the resulting parameters in the benchmark experiment, the experiment with endogenous interest rates and the experiment where the size distribution of establishments is calibrated to data for all sectors. Variations considered in other sensitivity analysis are described in the text.

Table 1: Parameters

	Benchmark	No storage	All sectors
β	0.45	0.45	0.45
α	0.35	0.47	0.39
ν	0.85	0.85	0.77
δ	0.88	0.88	0.88
η_{US}	0.57	0.64	0.48
σ	1.00	1.00	1.00
λ_1	-2.10	-5.45	-2.71
λ_2	0.66	0.53	0.58

References

Acemoglu, D., and F. Zilibotti (2001): "Productivity Differences," *Quarterly Journal of Economics*, 116, 563–606.

Amaral, P., and E. Quintin (2006): "A Competitive Model of the Informal Sector," *Journal of Monetary Economics*, 53(7), 1541–53.

Antràs, P. (2004): "Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution," *Contributions to Macroeconomics*, 4.

Atkeson, A., and P. Kehoe (2001): "The Transition to a New Economy After the Second Industrial Revolution," Working Paper 606, Federal Reserve Bank of Minneapolis.

Atkeson, A., A. Khan, and L. Ohanian (1996): "Are Data on Industry Evolution and Gross Job Turnover Relevant for Macroeconomics?," *Carnegie-Rochester Conference Series on Public Policy*, 44, 215–50.

Bencivenga, V., and B. Smith (1991): "Financial Intermediation and Economic Growth," *Review of Economic Studies*, 58, 195–209.

Caselli, F. (2003): "The Missing Input: Accounting for Cross-Country Income Differences," in *Handbook of Economic Growth, forthcoming*, ed. by P. Aghion, and S. Durlauf.

Cooley, T., R. Marimon, and V. Quadrini (2004): "The Aggregate Consequences of Limited Enforceability," *Journal of Political Economy*, 112, 817–847.

DeLong, B., and L. Summers (1991): "Equipment Investment and Economic Growth," *Quarterly Journal of Economics*, 106, 445–502.

Duffy, J., and C. Pappageorgiou (2000): "A Cross-Country Empirical Investigation of the Aggregate Production Function Specification," *Journal of Economic Growth*, 5, 87–120.

Erosa, E., and A. Hidalgo (2004): "On Capital Market Imperfections as a Source of Low TFP and Economic Rents," unpublished.

Gentry, A., and R. Hubbard (2004): "Entrepreneurship and Household Saving," *Advances in Economic Analysis & Policy, Berkeley Electronic Press*, 4.

Goldsmith, R. (1969): *Financial Structure and Development*. Yale University Press, New Haven, CT.

Gollin, D. (2002): "Getting Income Shares Right," *Journal of Political Economy*, 110, 458–74.

Greenwood, J., and B. Jovanovic (1990): "Financial Development, Growth, and the Distribution of Income," *Journal of Political Economy*, 98, 1076–107.

Hall, R., and C. Jones (1999): "Why Do Some Countries Produce So Much More Output per Worker Than Others?," *Quarterly Journal of Economics*, 114, 83–116.

Herrendorf, B., and A. Teixeira (2004): "Monopoly Rights Can Reduce Income Big Time," unpublished.

Heston, A., R. Summers, and B. Aten (2002): "Penn World Table Version 6.1," *Center for International Comparisons at the University of Pennsylvania (CICUP).*

Hsieh, C.-T., and P. Klenow (2004): "Relative Prices and Relative Prosperity," unpublished.

Jeong, H., and R. Townsend (2004): "Discovering the Sources of TFP Growth: Occupational Choice and Financial Deepening," unpublished.

Khan, A. (2001): "Financial Development and Economic Growth," *Macroeconomic Dynamics*, 5, 413–433.

King, R., and R. Levine (1993): "Finance and Growth: Schumpeter Might Be Right," *Quarterly Journal of Economics*, 108, 717–737.

Levine, R. (2005): "Finance and Growth: Theory and Evidence," in *Handbook of Economic Growth, forthcoming*, ed. by P. Aghion, and S. Durlauf. Elsevier.

Levine, R., N. Loayza, and T. Beck (2000): "Financial Intermediary Development and Growth: Causes and Causality," *Journal of Monetary Economics*, 46, 31–77.

Lucas, R. (1978): "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9, 508–23.

McGrattan, E., and E. Prescott (2000): "Is the Stock Market Overvalued," *Federal Reserve Bank of Minneapolis Quarterly Review*, 25(4), 20–40.

McKinnon, R. (1973): *Money and Capital in Economic Development*. The Brookings Institution, Washington, DC.

Pessoa, S., S. Pessoa, and R. Rob (2003): "Elasticity of Substitution Between Capital and Labor: A Panel Data Approach," unpublished.

Prescott, E. (1998): "Needed: A Theory of Total Factor Productivity," *International Economic Review*, 39, 529–49.

Quadrini, V. (1999): "The importance of entrepreneurship for wealth concentration and mobility," *Review of Income and Wealth*, 45, 1–19.

Restuccia, D. (2004): "Barriers to Capital Accumulation and Aggregate Total Factor Productivity," *International Economic Review*, 45, 225–38.

Restuccia, D., and R. Rogerson (2003): "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," unpublished.

Restuccia, D., and C. Urrutia (2001): "Relative Prices and Investment Rates," *Journal of Monetary Economics*, 47, 93–121.

Shaw, E. (1973): *Financial Deepening in Economic Development*. Oxford University Press, New York, NY.

Shleifer, A., and D. Wolfenzon (2002): "Investor Protection and Equity Markets," *Journal of Financial Economics*, 66, 3–27.

Townsend, R., and K. Ueda (2005): "Financial Deepening, Inequality, and Growth: A Model-Based Quantitative Evaluation," forthcoming, Review of Economic Studies.

Tybout, J. (2000): "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?," *Journal of Economic Literature*, 28, 11–44.

Figure 1: **Data**

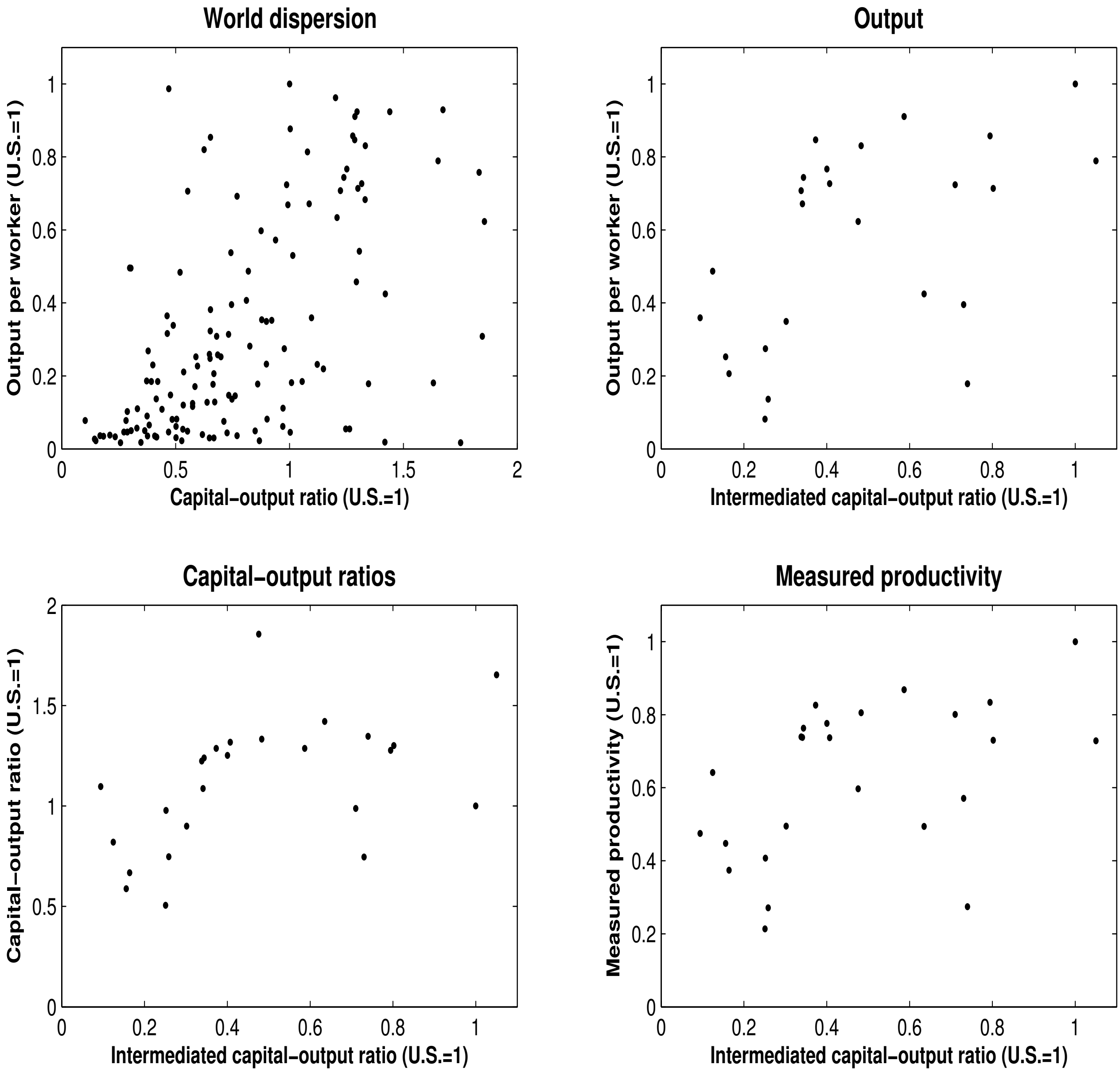

Figure 2: **Cobb-Douglas**

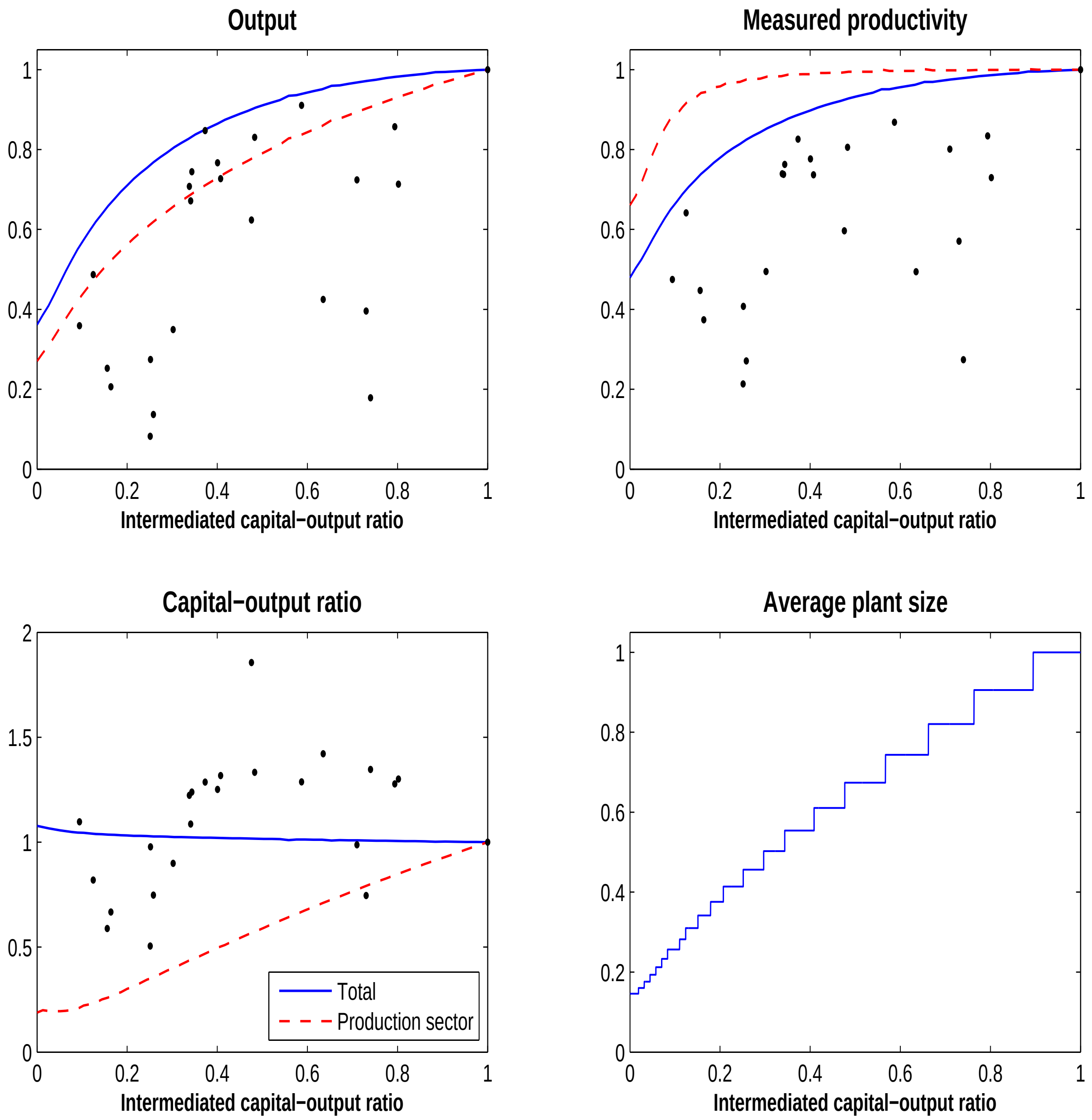

Figure 3: **Capital market clearing**

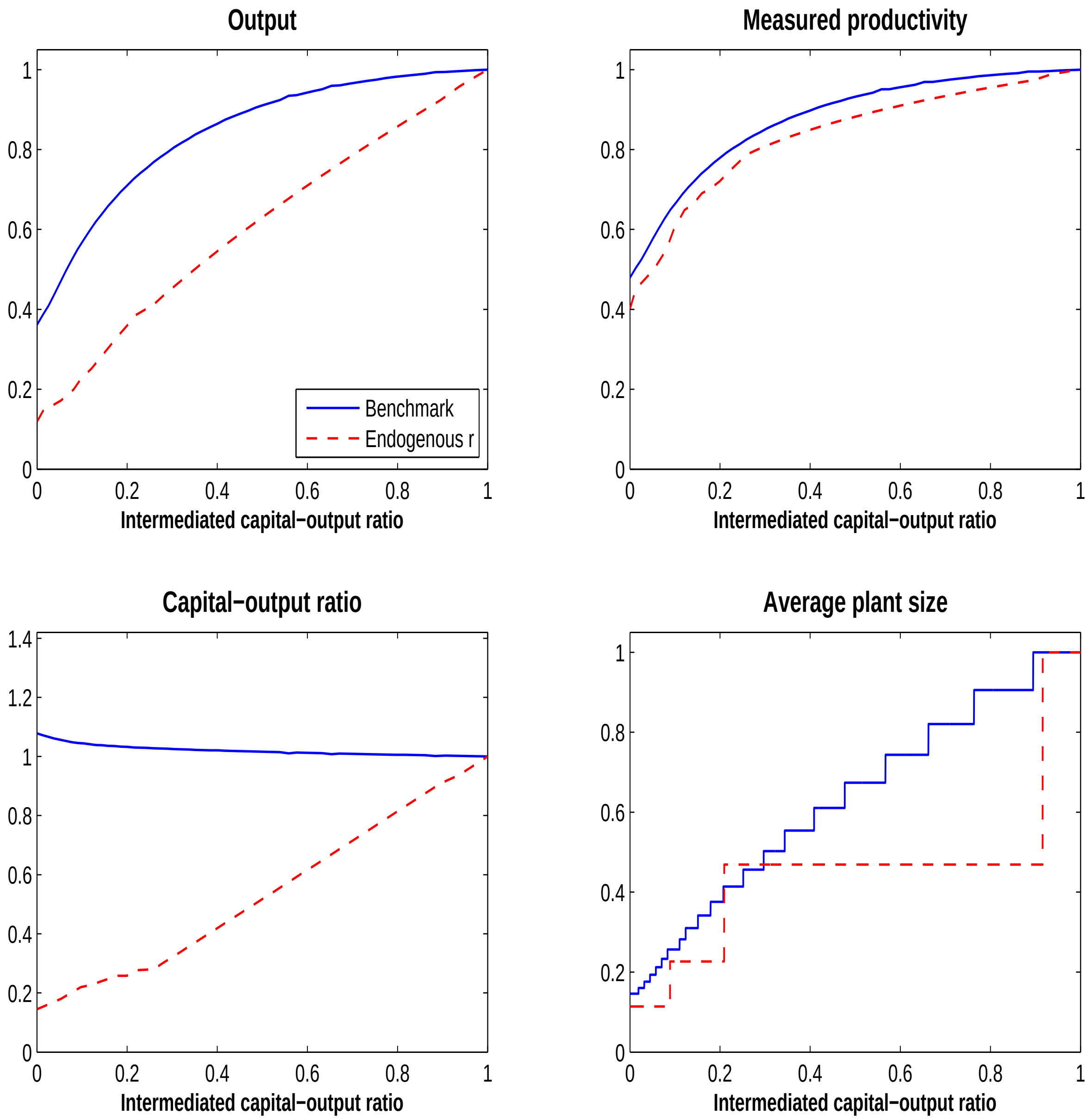

Figure 4: **All sectors**

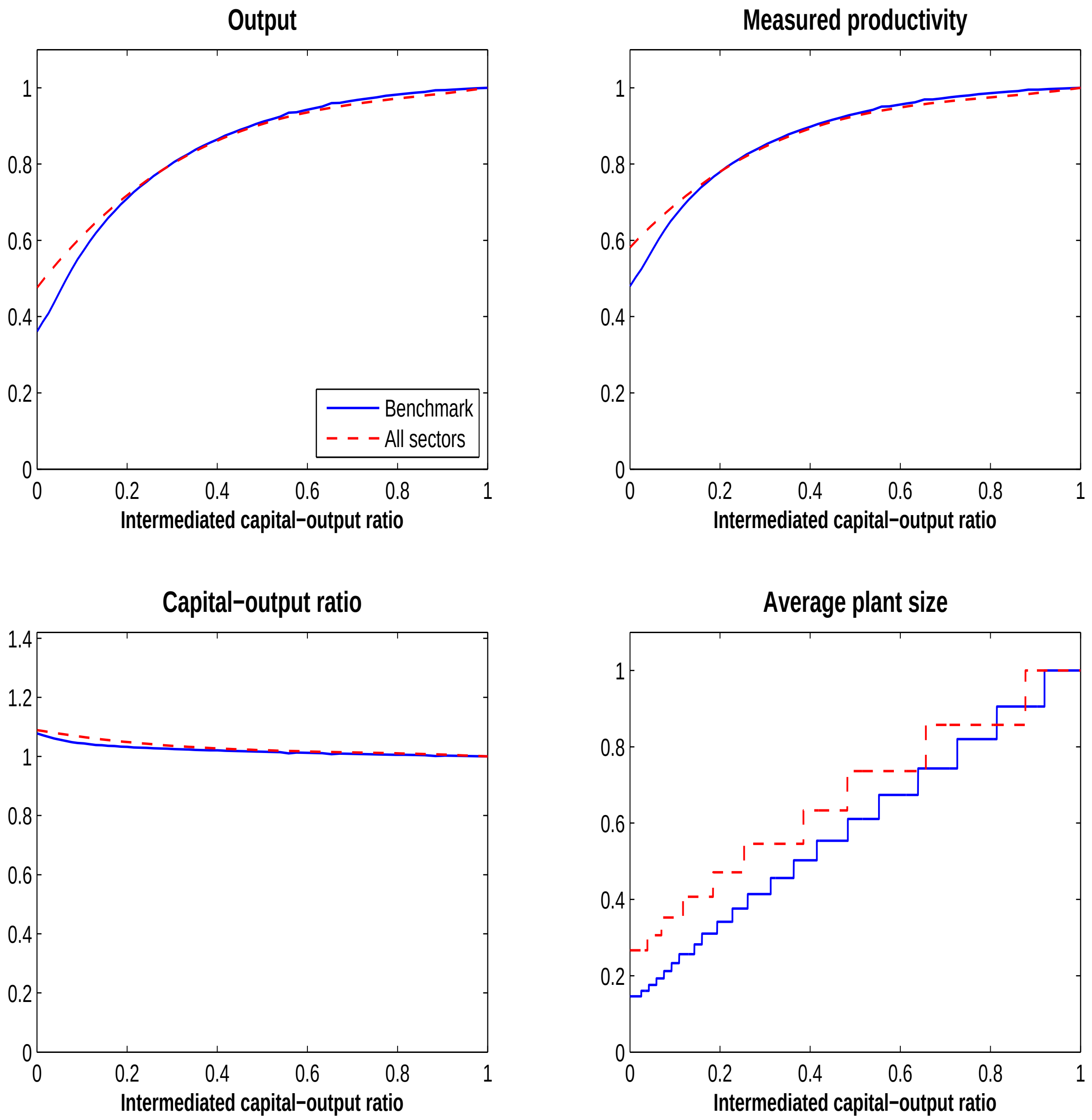

Figure 5: **Homogenous agents**

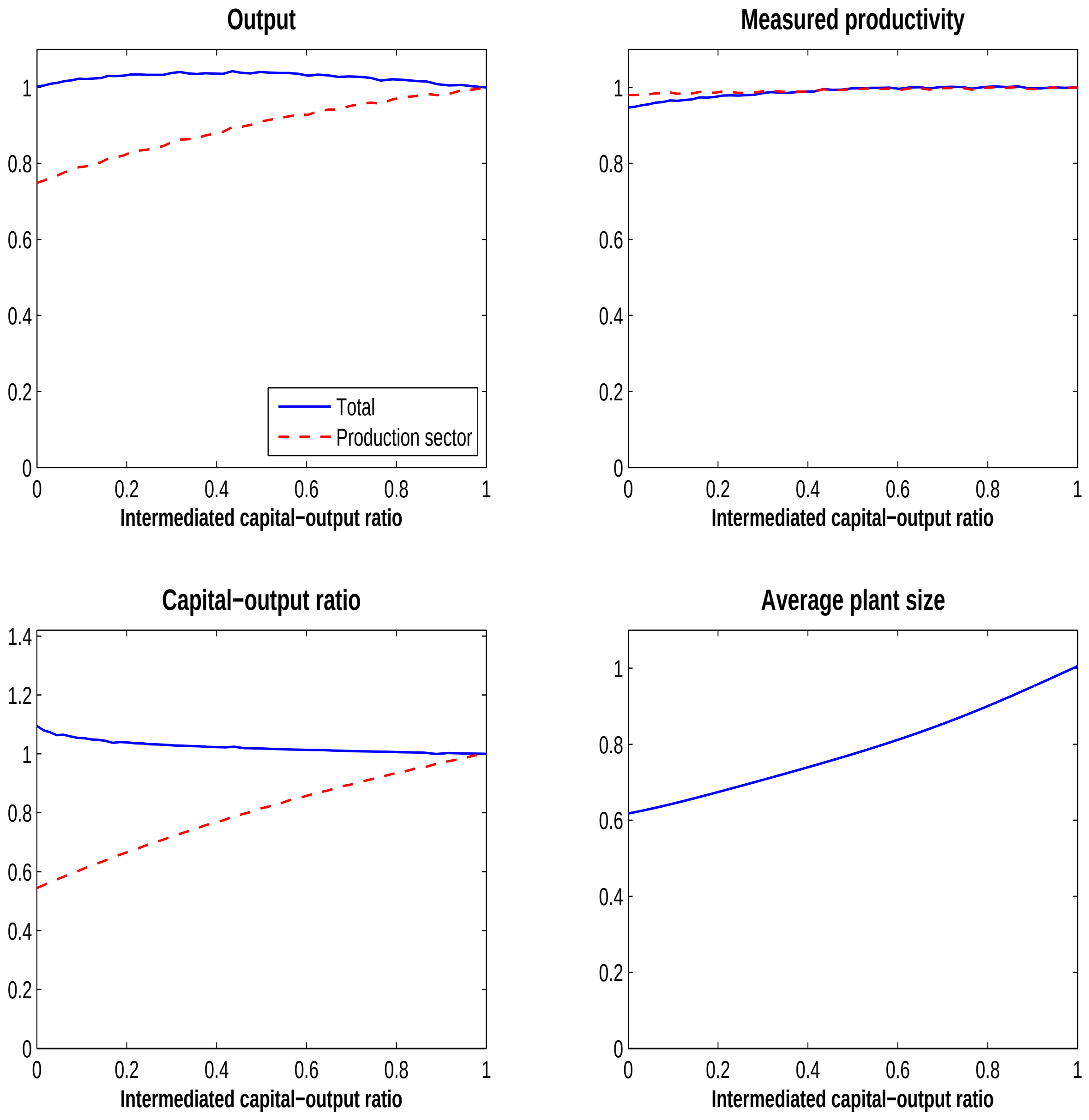

Figure 6: **Different managerial shares**

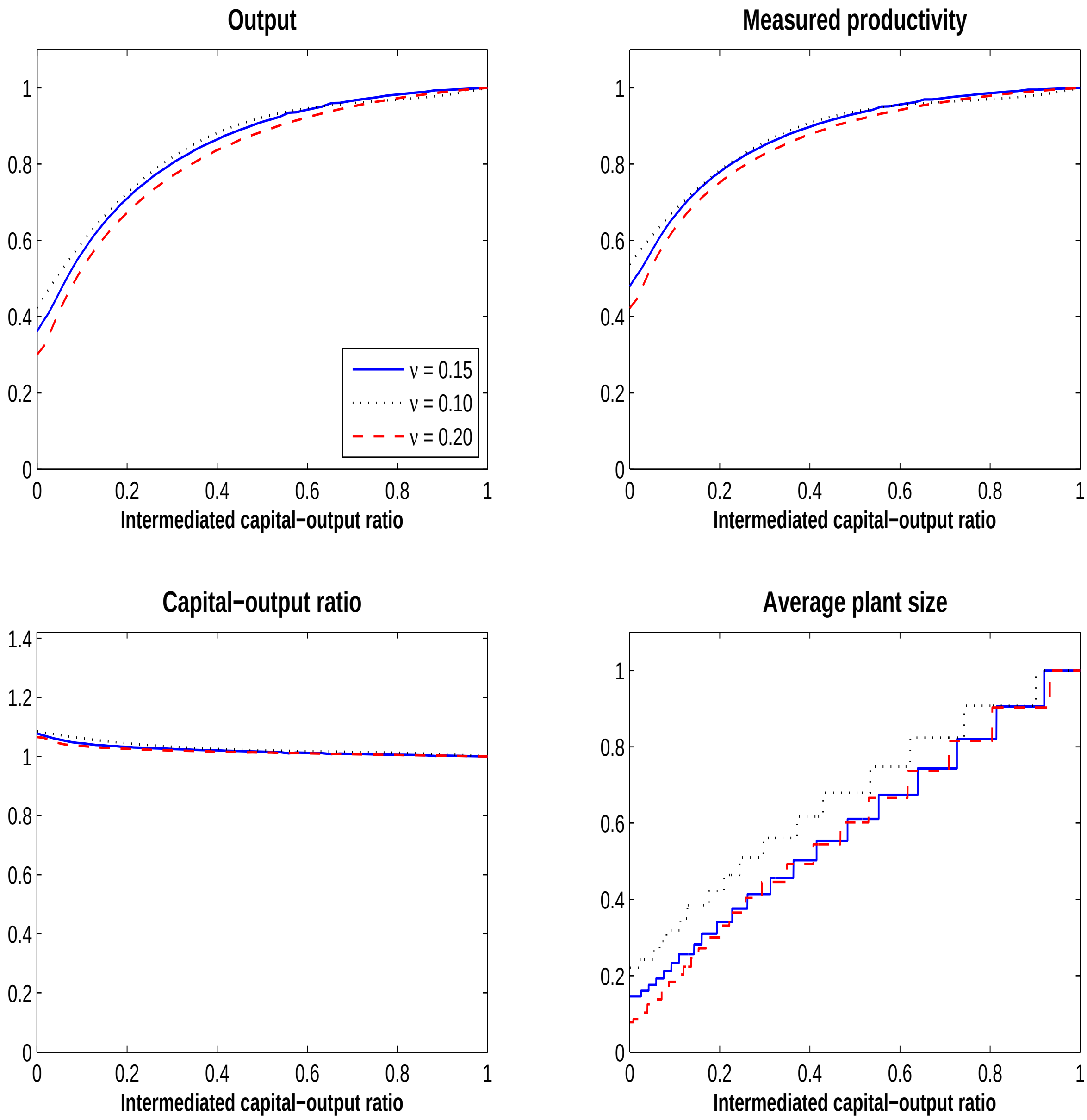

Figure 7: **Different capital shares**

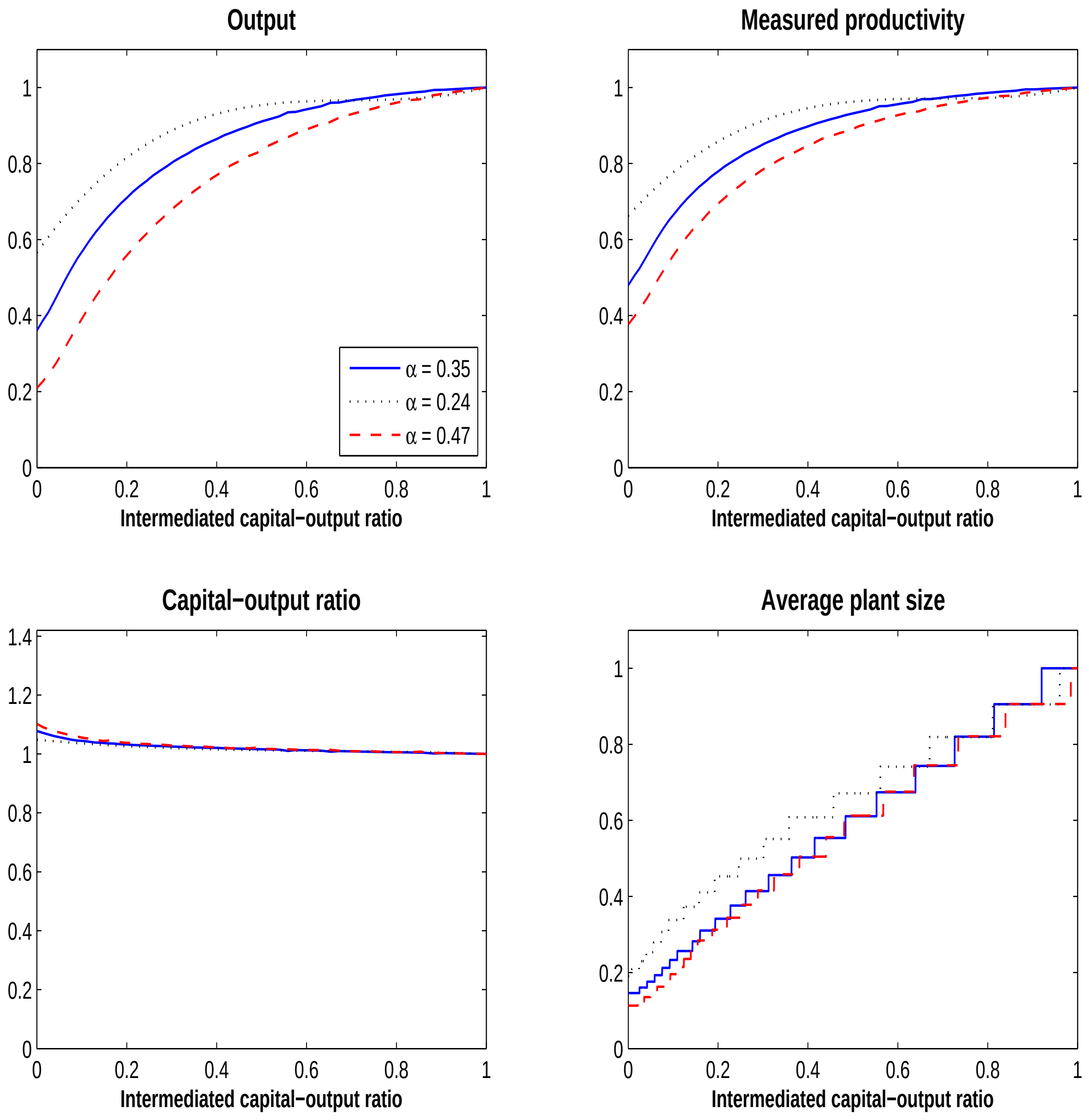

Figure 8: **Different elasticities of substitution**

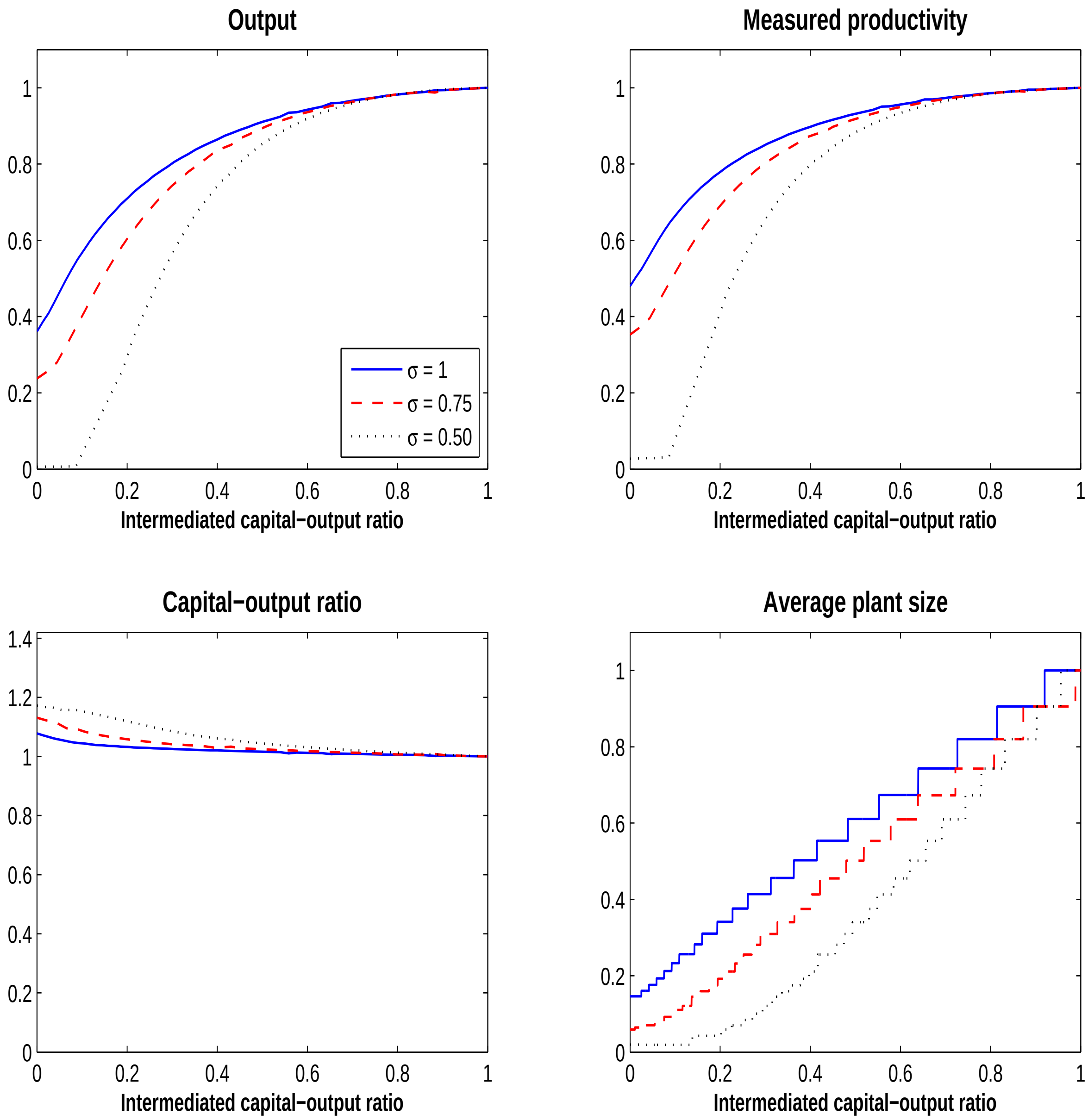

Figure 9: **Complete markets in the benchmark economy**