Financial Inclusion, Entrepreneurship and Employment Creation: Theory and a Quantitative Assessment*

Timothy Besley LSE and CIFAR

Konrad Burchardi IIES, Stockholm

Maitreesh Ghatak LSE

September 2017

Abstract

This paper develops a general equilibrium model with credit market frictions where agents differ in entrepreneurial ability and wealth to study the benefits of financial inclusion. As well as modeling the impact of credit market frictions, we allow for an increase in financial market access. We calibrate the model using the US size distribution of firms to create a benchmark. We show that it is access to finance which is quantitatively more important than factors which affect credit market frictions such as use of collateral or competition in credit markets. The main mechanism is through selection of the highest quality entrepreneurs, and the main beneficiaries of credit market access are wage laborers due to an expansion in labor demand and creation of large firms. Indeed, the size distribution of firms comes to resemble that in a more developed economy as credit market access expands.

JEL Classification: E44, G28, O16

Keywords: Financial Inclusion, Entrepreneurship. Employment Creation

*We are grateful to the ESRC-DFID growth research program for financial support (Grant reference ES/L012103 /1). We have received valuable research assistance from Kanishak Goyal, Pallavi Jindal, Kosha Modi, Tanmay Sahni, and Saurav Sinha.

1 Introduction

Increasing financial inclusion is now regarded as one of the principal development challenges (see, for example, World Bank, 2014). Although estimates vary, it appears that around half the world's population does not have access to formal banking services. Not surprisingly, financial exclusion is concentrated among the poorest people in the poorest countries. There is a wide variety of ways in which lack of access to financial services results in economic costs to those concerned. Lacking the capacity to save in reliable ways can damage the ability to build up assets, smooth against shocks as well as make provisions for old-age. And entrepreneurs who lack access to finance may not be able to start or scale up the operation of a business.

This paper models the gains from extending the reach of financial markets with a focus on expanding access to capital for entrepreneurs. We develop a general equilibrium model with four key features: (i) only a sub-group of the population are able to access financial markets, the remainder being in financial autarky (ii) all individuals in the economy, who may differ in their entrepreneurial productivity and their initial wealth, make an occupational choice, between being an entrepreneur and a wage labourer (iii) those who access financial markets, negotiate optimal contracts with lenders respecting the possibility of moral hazard (iii) wages are determined in general equilibrium.

The model highlights an important channel by which increasing financial inclusion affects the economy, namely the employment channel. As access to financial markets increases, labour demand is increased which pushes up equilibrium wages. Since the vast majority of workers are wage labourers, this extends the benefits of financial inclusion across the economy. As wages rise, there is stronger selection of entrepreneurs from the pool of those who are more productive which, in turn, leads the size distribution of firms to shift towards larger employers as capital is deepened in the economy. Hence the substantial gains from financial inclusion are through its impact on employment creation.

A key innovation in this paper is in the detailed modelling of credit market frictions in a general equilibrium setting. To date, much of the literature has used a reduced form model where credit access is limited by repayment technologies. While instructive, it creates a somewhat mechanical friction in capital markets as opposed to deriving it from optimal contracts. This makes it difficult to assess whether it is access to credit or the implications of second-best contracting frictions that matter quantitatively when assessing the importance of credit market frictions. One benefit of our approach is that we are able to separate these things out. One of our main findings is that embedding such frictions in a market equilibrium settings with an endogenously determined outside option actually leads to smaller losses in productivity and welfare, compared to limited market access. Being able to show this is a benefit from having a more detailed model of agency problems which derives optimal contracts.

The model offers a specific window on entrepreneurship and the development process. One of the striking features of low income economies is the large population of self-employed small-scale entrepreneurs. This can be viewed as being concomitant with a low wage economy and poor access to finance by large swathes of the population. As financial inclusion expands, large firms constitute a higher fraction of firms as the more productive firms are the ones that stay in business with the higher wages that now ensue. As the financial market access extends further, we get only a small fraction of the population running their own firms with the vast majority relying on selling their labour power, but this is good for wage labour as wages are higher. Indeed, it is developments in the labor market which explain many of the quantitative effects that the model finds.

To get a sense of the quantitative magnitude of these effects, we calibrate the model. As well as exploring the theoretical mechanism in detail and quantifying the aggregate gains from financial inclusion, we can also look at distributional effects across the population with two dimensions of heterogeneity: wealth and entrepreneurial talent. As financial inclusion increases, inequality is influenced more by who has entrepreneurial talent and less by who owns wealth. This approach also allows us to explore how two features of the contracting environment affect the outcomes, namely improving property rights which enhances access to capital and increasing competition between lenders.

The paper shows that there are indeed large aggregate benefits from extending financial market access. Moving from autarky to full inclusion in our calibrated economy increases wages from 40% of the US wage to 90%. However, these are driven almost exclusively through an employment-cum occupational choice channel which emerges from an aggregate general equilibrium perspective. When financial inclusion is first introduced into an "unbanked economy", the most dramatic effect that we can observe is on the proportion of self-employed entrepreneurs in the economy. The big beneficiaries of financial inclusion are wage labourers due to the employment creation that occurs and we show how the size distribution of firms increases with small firms being squeezed as marginal entrepreneurs exit. We also show that financial inclusion breaks the link between wealth and occupational choice.

Increasing competition does mean that entrepreneurs who access capital markets get a larger share of surplus. Otherwise, the benefits of inclusion will tend to accrue to lenders. We show that this is mainly a distributional issue between firms and lenders while wage labourers always gain with raising wages. Property rights also matter and increase efficiency. However, the efficiency effects of property rights expansion and increased competition are quantitively small relative to the effects of expanding the domain of financial markets.

This paper explores micro-economic factors which determine differences in the *level* of income per capita. The development accounting literature, such as Caselli (2005), has shown that it is differences in total factor productivity across economies that are key. Our paper is in the spirit of Hsieh and Klenow [2010] who tied this explicitly to factor misallocation. This links to older and long-standing debates in the development economics literature on how contracting frictions and imperfect markets matter for under-development. For example, authors such as Bardhan (1984) and Stiglitiz (1988) have highlighted a range of such frictions but without providing an approach to assess their implications quantitatively.

The idea that development of the financial sector has important implications for the economy has a long history with pioneering contributions by Gerschenkron (1962) and Goldsmith (1969). Both put the development banking system at the heart of understanding differences in the trajectories taken by economies. A large body of work has established a strong correlation between measures of financial market development and economic performance at the aggregate level (see, for example, Levine, 2005, Cihak et al, 2013). In parallel, there has also been a theoretical literature on the importance of financial frictions in affecting growth and development including Banerjee and Newman (1993) and Galor and Zeira (1993).[1] This literature does not typically focus on heterogeneous managerial ability (Lloyd-Ellis and Bernhardt, 2000 being an exception) and as a result, better functioning credit markets increase rather than decrease the fraction of entrepreneurs in the economy. In our model, improved financial market access

[1]Reviews of the literature can be found in Banerjee and Duflo (2005) and Matsuyama (2007).

enables more able individuals to become entrepreneurs and hire more workers.

There is also an extensive theoretical and empirical literature on how financial arrangements affect households and businesses, particularly how market frictions due to transactions costs and informational constraints may lead to borrowing constraints, and possibly, to poverty traps (see, for example, Banerjee and Duflo 2010, Karlan and Zinman, 2009 and Townsend and Ueda, 2008). This literature looks much at the ground up and focuses on heterogeneity and distributional implications of credit market activity.

A number of papers relate financial frictions to aggregate economic performance in ways that combine theory, data and calibration methods. In Jeong and Townsend (2007), there is a modern and subsistence economy with agents differing in wealth and talent. There is a fixed cost of setting up a firm and some agents, as in this paper, lack access to credit markets. They calibrate the model to Thai data showing that credit access is an important factor in explaining TFP dynamics. Buera et al (2011) also study the aggregate implications of credit market access emphasizing that there may be differences between manufacturing and services. They also introduce a non-convexity due to an entry cost. They model the financial friction as imperfect enforcement which limits the use of capital that an entrepreneur can use. After calibrating the model to U.S. data, they find that the variation in financial frictions which they explore can bring down output per worker to less than half of the perfect-credit benchmark. Moll (2014) builds on these approaches and explores the implications of productivity shocks which lead to inefficient capital allocation which can persist in the long-run. This provides a link to research which has looked at the macroeconomic effects of micro-economic distortions such as Bartelsman, Haltiwanger, and Scarpetta (2013), Hsieh and Klenow (2009) and Restuccia and Rogerson (2008). In general, these models therefore do not generate any equilibrium defaults even though they induce capital misallocation.

Our paper builds on these contributions by providing a more complete model of credit market distortions which has the possibility of default in equilibrium due to the presence of shocks, a realistic feature of credit markets. Moreover, the default rate is determined endogenously for each type of borrower in an optimally designed credit contract with outside options determined in general equilibrium. We show that the default rate is a sufficient statistic for credit misallocation for each type of borrower. We then explore the impact of extending the reach of credit markets exploring their impact on the size distribution of firms and equilibrium wages.

The recent concern about financial inclusion builds on these observations and tries to find metrics to study access to different kinds of financial markets. This has highlighted how having large populations of unbanked populations is a key issue in many countries around the world. The Global Financial Inclusion 2014 ("Global Findex") database based on a survey of 150,000 individuals in 148 countries shows sharp differences across countries showing less use of financial products in poor countries and generally among low income individuals. For example, in developing countries, the top quintile of earners is more than twice as likely to have a bank account than the bottom quintile and the cost of having an account or distance from the nearest branch are frequently cited as the reason (Demirgüç-Kunt and Klapper, 2013). A number of papers have explored the consequences of rolling out banking services. Burgess and Pande (2005) exploit a natural experiment due to bank-branching rules in India and find a significant impact on agricultural wages. Dupas et al (2017) looked at experimental variation in access to banking services in three countries: Uganda, Malawi and Chile. They suggest that there is a puzzlingly low take-up rate of banking services further underlining the challenge of expanding the outreach of financial services. Our approach provides a way of looking at the potential gains from expanded financial access if it can lead to greater borrowing.

The remainder of the paper is organized as follows. In the next section, we lay out the theoretical framework that we use. Section three moves from the model to the data and shows how it can be calibrated. Section four develops the results, first on the structure of credit contracts in general equilibrium and second on the effects of extending financial inclusion. Section five concludes.

2 Theory

The model developed here constitutes a significant development of the framework put forward in Besley et al (2012) to allow for endogenous occupational choice and wages. A group of agents who are heterogeneous in two dimensions: productivity and wealth can choose one of two possible occupations: becoming an entrepreneur or working as a wage labourer. If they choose to be entrepreneurs, then they have to decide how much capital and labour to employ. Capital can come from their own resources, i.e. their wealth, but can be augmented by borrowing if they have access to financial markets. However, lending is risky and some entrepreneurs end up defaulting on their loans. Labor is then hired by the successful entrepreneurs.

When entrepreneurs can access credit markets, these are subject to frictions. Credit market frictions are created by a lack of wealth which limits collateral and hence creates a moral hazard problem with respect to the effort put into project success. Wealth can be limited either because the borrowers are intrinsically poor or because of imperfect property rights which limit the use of wealth as collateral. Lenders can design contracts which optimize the terms of access to credit to maximize their profits, given that wealth is limited and effort is unobserved. If borrowers have some wealth that can be used as collateral, this will diminish the problem of moral hazard.

The first part of the theory considers optimal credit contracts to reflect heterogeneity in borrowers. In this part of the paper, the outside option of the borrower will be fixed exogenously. Hence, we think of this as a partial equilibrium setting. We show precisely how frictions in the credit market lead to misallocation of capital since lenders have to charge a risk premium to compensate for the probability of default. A high equilibrium default probability leads to less capital being allocated to an entrepreneur all else equal. This section, also illustrates the importance of competition in the credit market. This is modelled very simply as the fraction of the surplus in a lending relationship which goes to the borrower versus the lender. The more competitive is the credit market, the better the borrower does.

Before proceeding to general equilibrium, we introduce a financial inclusion parameter which, following Jeong and Townsend (2007), denotes the fraction of individuals with access to financial markets. As in Townsend, (1978), we think of this as reflecting a prohibitively high transaction cost which some agents face, for example due to their geographical location or level of knowledge. The next step is to consider who becomes an entrepreneur as a function of talent and wealth. In general, this will be the most wealthy and productive individuals. Finally, we consider the aggregate implications of the model for the level of output and equilibrium wages. To that end we will make specific functional form assumptions. In particular, we suppose that the production process is a standard decreasing-returns Lucas "span of control" model. This will allow us to characterize the factor allocation. We will jointly determined the equilibrium wage and the fraction of agents who become entrepreneurs. This endogenously determines the outside option of agents who borrow to be entrepreneurs.

2.1 Entrepreneurs, Lenders and Credit Contracts

The economy is populated by a continuum of agents all of who are endowed with a unit of labour power. Each agent is characterized by an ability level if she becomes an entrepreneur, denoted by θ and a level of wealth, denoted by a. Heterogeneous productivity can be interpreted either as "ability" or being endowed with a particular production technology.

Workers Workers are risk neutral, and all workers can earn p_l in expectation in a competitive labour market. In equilibrium, there are two classes of workers: managers and wage labourers. All agents are assumed to be equally productive in both labour market roles. The wage of a wage labourer is p_l and of a manager is p_e. As we detail below, managers also need to be compensated for risk since some may be contracted to work for firms which turn out to be unsuccessful. In equilibrium, managers and wage labourers earn the same expected wage.

Entrepreneurs Any agent can set up a firm and work as an entrepreneur. All firms have access to a production technology which allows them to earn a profit by using labour and capital. However, their idiosyncratic productivity level θ affects their productivity as entrepreneurs. The output price is p_y.

Let $f(k, l; \theta)$ denote the production function where k is the value of capital employed and l is wage labour employed; we spell out its properties in Assumption 1 below. Entrepreneurs will be residual claimants on the firm's profit stream and therefore form a capitalist class in this model.

Managerial Labor Firm success is stochastic and depends on managerial input. Specifically, the probability of being able to produce successfully is denoted by $g(e)$ where $e \in [0, 1]$ denotes the level of "managerial labour": The more managerial labour is hired, the higher is the success probability of the firm. In our model managers only affect the success probability of firms, not their output in case of success. By choosing e the entrepreneur trades of higher profits in case of success with a higher success probability. An alternative interpretation of e is therefore to understand it as a measure of entrepreneurial risk taking. This formulation allows for the existence of a managerial labour market as well as owner-managers.

Henceforth, let $\mathbf{p} = (p_y, p_l, p_e)$ denote the price vector. Without loss of generality, and for notational compactness, we allow *all* of the functions that depend on *any* price to be functions of the *entire* price vector even if only some prices are relevant for some specific decisions. We denote the cost of managerial labour by $d(e; \theta, \mathbf{p})$ which is assumed to be increasing in e, p_e and θ.[2] Since θ will affect firm size positively, this formulation assumes that larger firms need to

[2] A priori, we expect this only to depend on p_e and the most natural case would be where

$$d(e; \theta, \mathbf{p}) = p_e D(e; \theta)$$

which is the case when we calibrate the model below.

hire more managers. Further properties of $g(e)$ and $d(e;\theta,\mathbf{p})$ are introduced in Assumptions 1 and 2 below.

We make the following regularity assumption throughout:

Assumption 1 *The following conditions hold for the functions $g(e)$, $d(e;\theta,\mathbf{p})$ and $f(k,l;\theta)$:*

(i) $g(e)$ is strictly increasing, twice-continuously differentiable, strictly concave for all $e \in [0,1]$, $p(0) = 0$ and $p(1) \in [0,1]$.

(ii) $f(k,l;\theta)$ is twice-continuously differentiable and strictly increasing in $k \in \mathbb{R}^+$ and $l \in \mathbb{R}^+$, strictly concave in l and is increasing in θ with $f_{\theta k} > 0$ and $f_{\theta l} > 0$. Further $f(k,l;\theta) \geq 0$ for all $(k,l) \in \mathbb{R}^+ \times \mathbb{R}^+$.

(iii) $d(e;\theta,\mathbf{p})$ is strictly increasing, twice-continuously differentiable and convex in e; it is increasing in θ and p_e.

(iv) $\epsilon(e;\theta,\mathbf{p}) \equiv \frac{d_{ee}(e;\theta,\mathbf{p})}{g_e(e)} - \frac{d_e(e;\theta,\mathbf{p})g_{ee}(e)}{\{g_e(e)\}^2}$ is continuous and increasing for $e \in [0,1]$.

Wealth and Collateral Each entrepreneur has a level of wealth a measured in units of labour endowment. In addition to their own wealth, entrepreneurs can approach lenders to borrow money. Hence total capital available to an entrepreneur is $k = x + a$ where x is the amount that she borrows.

Collateral is from individual wealth. As in Besley et al (2012), we suppose that only a fraction τa of wealth can be used as collateral where $\tau < 1$ if property rights are imperfectly established.

Lenders Credit contracts are described by a vector (x,r,c) comprising (i) an amount borrowed, x, (ii) an amount to be repaid if the firm is successful, r, (iii) an amount of financial collateral c. For notational simplicity we will use $\mathbf{t} = (x,r,c)$ to denote a credit contract.

Lenders can all access funds at the same opportunity cost $\gamma > 1$, which is the gross interest rate (principal plus the net interest rate). A lender's expected profit when agreeing to lend to a producer with collateral τa is therefore:

$$\Pi(\mathbf{t};e) = g(e)\,r + (1 - g(e))\,\tau a - \gamma x. \tag{1}$$

This reflects the fact that, with probability $g(e)$, the lender is repaid and with probability $(1 - g(e))$ there is default in which case the lender seizes the producer's collateral.

There is a finite set of lenders with whom entrepreneurs can contract. To model competition between lenders, we suppose that there is a Bertrand-style price setting game. Imagine that there are two lenders with identical access to the capital market, γ and the same enforcement technologies. In principle this should lead to borrowers capturing all of the surplus as lenders compete for borrowers until ex ante payoffs are zero. However, there are good reasons to doubt that this is a reasonable model and there are likely to be costs of switching between lenders. Rather than being specific about the friction, we capture imperfectly competitive credit markets by supposing that an alternative lender provides an outside option worth a share ϕ of the total surplus created by their lending contract. If $\phi = 1$, then all of the surplus over and above the entrepreneur's outside option accrues to the entrepreneur rather than the lender. This is the competitive benchmark. On the other hand, if ϕ is small, then the lender has a lot of market power.

Timing The timing of production for a type (a, θ) is as follows.

1. Workers choose whether to become an entrepreneur or worker.

2a. If she chooses to become a worker, she supplies one unit of labour to the labour market.

2b. If she is an entrepreneur, then each lender offers her a contract (x, r, c). After deciding whether to accept this contract, she chooses managerial labour, e.

3a. With probability $g(e)$, she is a successful entrepreneur and has a viable project. Then she chooses how much labour to hire, l. Output is realized, wages are paid to managerial and wage labourers, and the loan repayment, r, is made.

3b. With probability $1 - g(e)$, an entrepreneur produces nothing and forfeits collateral, c.

We now work backwards through these decisions to determine the optimal contract. Here, we suppose that prices $\mathbf{p}$ are fixed. We then explore the general equilibrium where these are determined.

Labor Hiring With probability $g(e)$, the firm produces in which case it decides how many wage labourers to hire to maximize profits, i.e.

$$l^*(k; \theta, \mathbf{p}) = \arg\max_{l} \{p_y f(k, l; \theta) - p_l l\} \tag{2}$$

and define $\pi(k; \theta, \mathbf{p}) \equiv p_y f(k, l^*(k; \theta, \mathbf{p}); \theta) - p_l l^*(k; \theta, \mathbf{p})$ as the conditional profit function given an allocation of capital k. Throughout we make the following assumption, that ensures well-defined interior solutions.

Assumption 2 *The following conditions hold for $g(e)$ and $\pi(k; \theta, \mathbf{p})$:*

(i) $\pi(k; \theta, \mathbf{p})$ *is strictly concave for all* $k \in \mathbb{R}^+$.

(ii) $g(e)\pi(k; \theta, \mathbf{p})$ *is strictly concave for all* $(e, k) \in [0, 1] \times \mathbb{R}^+$.

(iii) $\lim_{e \to 0} g_e(e)\pi(k; \theta, \mathbf{p}) - [d_e(e; \theta, \mathbf{p}) + g(e)\epsilon(e; \theta, \mathbf{p})] > 0$ *for all* $k > 0$;
$\lim_{k \to 0} g(e)\pi_k(k; \theta, \mathbf{p}) > \gamma$ *for all* $e > 0$.

Choice of Managerial Labor We allow lenders to offer credit to entrepreneurs which are tailored to an entrepreneur's characteristics, (a, θ). The timing of the managerial hiring decision captures the possibility of moral hazard in the credit market which is a source of credit market frictions, i.e. since managerial labour is costly and unobserved to the lender, there is a risk that a firm will shirk.

The expected payoff of an entrepreneur who borrows under contract $\mathbf{t}$ is given by:

$$V(e; \mathbf{t}, a, \theta, \mathbf{p}) = g(e)[\pi(x + a; \theta, \mathbf{p}) - r + c] - c - d(e; \theta, \mathbf{p}). \tag{3}$$

Observe that this is decreasing in the amount of collateral, all else equal.

The first order condition for managerial labour is :

$$g_e(e)[\pi(x + a; \theta, \mathbf{p}) - r + c] = d_e(e; \theta, \mathbf{p}). \tag{4}$$

Managerial "effort" – the amount of managerial labour hired – is increasing in collateral c and the amount borrowed, x. However, it is decreasing in r all else equal, i.e. asking for a higher loan repayment blunts incentives and increases the default rate. Equation (4) is an incentive compatibility constraint on credit contracts.

Managers face a risk when they work for a firm since it may turn out not to be successful. Since labour is equally productive in either role, persuading agents to work as managers therefore requires that they be compensated for risk. This implies that the managerial wage rate is $p_e = p_l / g(e)$. This will vary by firm if e varies and hence riskier firms will have to pay managers more.

Acceptable Credit Contracts As well as being incentive compatible, entrepreneurs must enter into contracts voluntarily at stage 2. Hence all credit contracts offered to an entrepreneur of type (a, θ) must generate a payoff which exceeds what is available elsewhere which we denote by u. The participation constraint is therefore:

$$V(e, \mathbf{t}; a, \theta, \mathbf{p}) \geq u. \tag{5}$$

In equilibrium, u is determined endogenously and depends on θ, a and $\mathbf{p}$. It can be thought of as a price which endogenously clears the credit market given outside opportunities available to an entrepreneur. In other words, it determines the expected returns from entrepreneurship striking a balance between the demand and supply for different occupations in the economy, which in turn depends on economic fundamentals, such as the distribution of talent and wealth and prices. Below we will determine p_e and p_l endogenously but all individuals take prices as given when making their decisions.

2.2 Credit Contracts in Partial Equilibrium

In this section, we explore access to credit holding fixed who decides to become an entrepreneur and the price vector $\mathbf{p}$. We begin with three key observations on the properties of such contracts.

First, as long as first-best effort cannot be implemented, the lender will choose $c = \tau a$, i.e. collateral is at it's highest possible value. As long as $c < \tau a$ surplus can always be extracted more efficiently by reducing r and increasing c, thereby relaxing the incentive compatibility constraint and leaving the borrower's participation constraint unchanged.

Second, when the participation constraint (5) is binding, combining (3), (4) and (5) yields that the borrower's optimal effort level is $\xi(v; \theta, \mathbf{p})$ defined by

$$\frac{d_e(\xi(v;\theta,\mathbf{p});\theta,\mathbf{p})g(\xi(v;\theta,\mathbf{p}))}{g_e(\xi(v;\theta,\mathbf{p}))} - d(\xi(v;\theta,\mathbf{p});\theta,p) = v,$$

where $v = u + \tau a$. Under Assumption 1, second-best effort $\xi(v; \theta, \mathbf{p})$ is increasing in v, and therefore in the outside option of the borrower and his collateralizable wealth. It also depends on prices as we are allowing the cost of managerial effort to depend on the wage (an element of $\mathbf{p}$). When the participation constraint is non-binding, we can combine (3) and (4) and maximize lender profit over e and x. In this case optimal effort and capital, denoted $e_0(\theta, \mathbf{p})$ and $k_0(\theta, \mathbf{p})$, are independent of v.

The value of v also determines whether the outside option is binding and/or whether the first-best level of surplus is attained, as summarized in our next result:

Proposition 1 *There exists* $[\underline{v}(\theta,\mathbf{p}), \bar{v}(\theta,\mathbf{p})]$ *such that optimal lending contracts implement effort* e *as follows:*

$$\hat{e}(v;\theta,\mathbf{p}) = \begin{cases} e_0(\theta,\mathbf{p}) & \text{for } v \leq \underline{v}(\theta,\mathbf{p}) \\ \xi(v;\theta,\mathbf{p}) & \text{for } \underline{v}(\theta,\mathbf{p}) < v < \overline{v}(\theta,\mathbf{p}) \\ e^*(\theta,\mathbf{p}) & \text{for } v \geq \overline{v}(\theta,\mathbf{p}) \end{cases}$$

where $e_0(\theta,\mathbf{p})$ *is a constant,* $e^*(\theta,\mathbf{p})$ *is a constant equal to first best effort,* $\lim_{v\to\underline{v}(a,\theta,\mathbf{p})}\xi(v;\theta,\mathbf{p}) = e_0(\theta,\mathbf{p})$ *and* $\lim_{v\to\overline{v}(a,\theta,\mathbf{p})}\xi(v;\theta,\mathbf{p}) = e^*(\theta,\mathbf{p})$.

When v is high then effort is first best, defined by

$$g_e(e^*(\theta,\mathbf{p}))[\pi(x(\overline{v}(\theta,\mathbf{p})+a;\theta,\mathbf{p});\theta,\mathbf{p})] = d_e(e^*(\theta,\mathbf{p});\theta,\mathbf{p}),$$

i.e. sets the marginal benefit equal to the marginal cost when the agent is a full residual claimant. At the other extreme, for low v, the effort level is set so that the outside option does not bind and the agent obtains an "efficiency" utility. This is characterized by

$$g_e(e_0(\theta,\mathbf{p}))\,\pi(x(\underline{v}(\theta,\mathbf{p})+a;\theta,\mathbf{p});\theta,\mathbf{p}) = \epsilon(e_0(\theta,\mathbf{p});\theta,\mathbf{p}) + d_e(e_0(\theta,\mathbf{p});\theta,\mathbf{p}).$$

In this case, it is "as if" the cost of effort is increased by the term $\epsilon(e_0(\theta,\mathbf{p});\theta,\mathbf{p})$ which represents the marginal "agency cost" due to moral hazard. At intermediate levels of v the effort distortion is decreasing in v.

Our third observation concerns the optimal allocation of credit which is determined by maximizing (1) with respect to x, subject to the constraints. We can show the following result:

Proposition 2 *Firm capital,* $\hat{k}(v;\theta,\mathbf{p})$*, and therefore* $\hat{x}(v;\theta,\mathbf{p}) = \hat{k}(v;\theta,\mathbf{p}) - a$*, is defined by*[3]

$$\pi_k\left(\hat{k}(v;\theta,\mathbf{p});\theta,\mathbf{p}\right) = \frac{\gamma}{g(\hat{e}(v;\theta,\mathbf{p}))}. \tag{6}$$

This is the core equation for capital allocation. It says that capital will be allocated on a risk-adjusted basis to reflect the equilibrium default probability. So the marginal return to capital is not equalized across firms to the extent that there are different probabilities of default. Capital is only misallocated to the extent that capital market frictions lead to distortions in e.

The repayment r is determined, conditional on $\hat{e}(v;\theta,\mathbf{p})$ and $\hat{k}(v;\theta,\mathbf{p})$, from (4) as:

$$r(v;a,\theta,\mathbf{p}) = \pi\left(\hat{k}(v;\theta,\mathbf{p});\theta,\mathbf{p}\right) + \tau a - \frac{d_e(\hat{e}(v;\theta,\mathbf{p});\theta,\mathbf{p})}{g_e(\hat{e}(v;\theta,\mathbf{p}))}.$$

Total surplus in a lending relationship is:

$$S(v;\theta,\mathbf{p}) = g(\hat{e}(v;\theta,\mathbf{p}))[\pi(\hat{x}(v;\theta,\mathbf{p})+a;\theta,\mathbf{p})] - d(\hat{e}(v;\theta,\mathbf{p});\theta,\mathbf{p}) - \gamma\cdot\hat{x}(v;\theta,\mathbf{p}).$$

The result in Proposition 1 therefore provides a convenient way of summarizing optimal contracts since v is a sufficient statistic for the efficiency of the lending arrangement, determining the level of effort and hence capital.

The following result gives a characterization of the ranges in which v can fall in terms of the surplus function, where S_v denotes the partial derivative of the surplus function with respect to v.

[3] Where $\hat{x}(v;\theta,\mathbf{p}) < 0$ the entrepreneur has sufficient wealth to self-finance at first best and will not borrow.

Corollary 1 *The surplus function, $S(v;\theta,\mathbf{p})$, is increasing in v whenever $S_v(v;\theta,\mathbf{p}) \in [0,1]$. For $v \geq \overline{v}(\theta,\mathbf{p})$ we have $S_v(\overline{v}(\theta,\mathbf{p});\theta,\mathbf{p}) = 0$. For $v < \underline{v}(\theta,\mathbf{p})$ the participation constraint of the entrepreneur does not bind, and at $\underline{v}(\theta,\mathbf{p})$ we have $S_v(\underline{v}(\theta,\mathbf{p});\theta,\mathbf{p}) = 1$.*

Credit contracts will implement first best effort as long as entrepreneurs can provide sufficient collateral, i.e. has high a, or have a high outside option. In the intermediate v range, greater collateral allows for more efficient lending since it relaxes (4). The lender then offers a higher x, which amplifies the effect of collateral on the incentive compatibility constraint. Similarly, a higher outside option increases lending efficiency. The lender has to transfer a greater share of surplus to the entrepreneur, and this is optimally implemented by reducing r and increasing x, which in turn increase effort. However, for $v \leq \underline{v}(\theta,\mathbf{p})$ the lender will always implement $e_0(\theta,\mathbf{p})$. In this range – due to the concavity of $g(e)$ – a reduction in r increases surplus by more than it transfers surplus to the entrepreneur. Therefore it is in the interest of the lender to offer a contract which leaves the entrepreneur with an expected income greater than the outside option. It is optimal to transfer surplus by decreasing r. In this region, the lender reacts to an increased c by increasing r by the same amount, and leaving both e and x unchanged. Surplus stays unchanged, but is transferred from the borrower to the lender.

The Lender's Participation Constraint Whether a lender wishes to lend to an entrepreneur of type (a,θ) depends upon whether they can make a profit by doing so. Hence for an entrepreneur of type (a,θ) to be offered any credit requires that

$$\hat{\Pi}(v;a,\theta,\mathbf{p}) \geq 0.$$

Determining the Entrepreneur's Outside Option The final part of the partial equilibrium analysis is to determine the entrepreneur's outside option endogenously. This will be the maximum of three things: (i) what she can obtain by borrowing from another lender, (ii) self-financing the project with the (limited) wealth owned and (iii) working for a wage. We now explore this in detail.

Let $\hat{u}(\phi;a,\theta,\mathbf{p})$ be defined by:

$$\phi \cdot S(\hat{u}(\phi;a,\theta,\mathbf{p}) + \tau a;\theta,\mathbf{p}) = \hat{u}(\phi;a,\theta,\mathbf{p}).$$

This implicitly defines the equilibrium payoff of an entrepreneur if the only outside option is to receives a share ϕ of the surplus in a lending relationship. Note that this is not the payoff from borrowing since the efficiency utility in Proposition 1 bounds the borrower's payoff from below when ϕ and/or τa are low.[4]

Now consider the payoff where the agent chooses to self-finance, i.e. use only his own wealth. This is given by

$$V^{self}(a,\theta,\mathbf{p}) = \max_{(e,k)} \{g(e)\,\pi(k;\theta,\mathbf{p}) - d(e;\theta,\mathbf{p}) - \gamma k : k \leq a\}. \tag{7}$$

Let $\{e^{self}(a,\theta,\mathbf{p}), k^{self}(a,\theta,\mathbf{p})\}$ denote the solutions to the maximization problem (7). Lastly the entrepreneur could choose to become a wage labourer. The entrepreneurs outside option will therefore be given by

$$\underline{u}(a,\theta,\mathbf{p}) = \max\{V^{self}(a,\theta,\mathbf{p}), \hat{u}(a,\theta,\mathbf{p}), p_l\}.$$

[4]Note that even with $\phi = 0$, the lender does not necessarily receive u since, as we observed Proposition 1, the entrepreneur's participation constraint might not be binding.

Comparative Statics We now have the following result for entrepreneur payoffs:

Proposition 3 *For $v > \underline{v}(a, \theta, \mathbf{p})$, the entrepreneur's expected profit increases with more competition (ϕ) and greater wealth (a). In the absence of further assumptions, the effect of productivity (θ) on the outcome is indeterminate.*

Thus entrepreneurs benefit from increased competition since they get a larger share of the surplus in the credit market. They also do better when they have more collateral to post. Equally, more productive entrepreneurs are better off.

2.3 General Equilibrium

So far, we have taken the price vector $\mathbf{p}$ and the occupational structure as given. Our general equilibrium analysis determines these endogenously.

Financial Market Access A fraction $z(a, \theta) \in [0, 1]$ of agents of type (a, θ) has access to financial markets. Denote with $\chi \in \{0, 1\}$ whether any given individual has access to credit markets. Let $h(a, \theta)$ denote the joint density associated with the distribution of (a, θ). Total financial inclusion in the economy is defined by

$$\bar{\chi} \equiv \int \int z(a, \theta) \, h(a, \theta) \, da d\theta,$$

i.e. as the proportion of agents who have market access. If they have access then they can access credit markets as described in the previous section.

Occupational Choice Let $\sigma \in \{0, 1\}$ denote whether an individual becomes an entrepreneur where $\sigma = 1$ is entrepreneurship. They will choose this when their expected payoff from being an entrepreneur exceeds that from being a wage labourer. Formally,

$$\sigma(a, \theta, \chi, \mathbf{p}) = \begin{cases} 1 & \text{if } \chi = 1 \text{ and } \hat{\Pi}(\underline{u}(a, \theta, \mathbf{p}) + \tau a; a, \theta, \mathbf{p}) \geq 0 \\ 1 & \text{if } V^{self}(a, \theta, \mathbf{p}) \geq p_l \\ 0 & \text{otherwise.} \end{cases}$$

The borrower will always choose to become an entrepreneur if the autarchy payoff is bigger than the wage. If she has access to credit markets, she will also become an entrepreneur if the lender can offer a profitable credit contract (satisfying the borrower's outside option and incentive constraint). Clearly this depends on the individuals type (a, θ). Moreover, since the payoff from entrepreneurship is increasing in a and θ, if a type (a, θ) becomes an entrepreneur then so do all individuals with higher wealth and productivity. Hence, there will be critical values of wealth and productivity that define the entrepreneurial class. How dense this is depends on the joint distribution of wealth and productivity.

Equilibrium Wages To determine equilibrium wages, we need to solve for aggregate labour supply and demand in the economy. This means aggregating over the distribution of wealth and productivity. Aggregate labour supply is determined by the fraction of individuals who choose not to become entrepreneurs, i.e.

$$L^S(\mathbf{p}) = \int \int z(a, \theta) \left[1 - \sigma(\theta, a, 1, \mathbf{p})\right] + (1 - z(a, \theta)) \left[1 - \sigma(\theta, a, 0, \mathbf{p})\right] \, h(a, \theta) \, da d\theta. \quad (8)$$

Denote the managerial labour demand, conditional on becoming entrepreneur, by

$$\hat{e}(a,\theta,\chi,\mathbf{p}) = \chi(\hat{e}(\underline{u}\,(a,\theta,\mathbf{p}) + \tau a; a,\theta,\mathbf{p}) + (1-\chi)e^{self}\,(a,\theta,\mathbf{p})\,,$$

and firm capital, conditional on becoming entrepreneur, by

$$\hat{k}(a,\theta,\chi,\mathbf{p}) = \chi(\hat{k}(\underline{u}\,(a,\theta,\mathbf{p}) + \tau a; a,\theta,\mathbf{p}) + (1-\chi)k^{self}\,(a,\theta,\mathbf{p})\,.$$

To solve for aggregate labour demand we need to take into account the fraction of firms that are operational given the equilibrium default probability which we denote by

$$\hat{g}\,(a,\theta,\chi,\mathbf{p}) = g(\hat{e}(a,\theta,\chi,\mathbf{p})).$$

Note that this also depends on $\mathbf{p}$ through its affects on profits and the cost of managerial labour. Labor demand also depends on the amount of labour hired by each firm, conditional on producing. We will denote this by

$$\hat{l}\,(a,\theta,\chi,\mathbf{p}) = l^*(\hat{k}(a,\theta,\chi,\mathbf{p});\theta,\mathbf{p})$$

using (2). Aggregate labour demand is then given by

$$\begin{aligned} L^D\,(\mathbf{p}) \;&=\; \int\!\!\int z(a,\theta)\left[\sigma\,(a,\theta,1,\mathbf{p})\cdot\left(\hat{l}\,(a,\theta,1,p)\cdot\hat{g}\,(a,\theta,1,\mathbf{p})+\hat{e}(a,\theta,1,\mathbf{p})\right)\right]h\,(a,\theta)\,dad\theta \\ &+\; \int\!\!\int(1-z(a,\theta))\left[\sigma\,(a,\theta,0,\mathbf{p})\cdot\left(\hat{l}\,(a,\theta,0,p)\cdot\hat{g}\,(a,\theta,0,\mathbf{p})+\hat{e}(a,\theta,0,\mathbf{p})\right)\right]h\,(a,\theta)\,dad\theta \end{aligned} \tag{9}$$

This is the sum over the labour demand functions of individuals, characterized by (a,θ,χ), who choose to become entrepreneurs at prevailing prices $\mathbf{p}$.

The equilibrium wage now equates supply and demand, i.e. solves

$$L^S\,(\hat{\mathbf{p}}) = L^D\,(\hat{\mathbf{p}})$$

where $\hat{\mathbf{p}}$ is the equilibrium price vector. This depends implicitly on all dimensions of choice: occupational choice, credit contracts which determines use of capital and labour demand. It also depends on the extent of financial access since this will affect who becomes an entrepreneur and the amount of labour demand among those who do, depending on whether they can access financial markets.

2.4 Two Benchmarks

Before proceeding to study the calibration of the model, it is worth considering two special cases that will serve as useful benchmarks in what follows: autarky and the first best.

Autarky We define autarky purely in terms of credit markets, i.e. to describe a situation where there is only trade in labour and goods markets, but not in capital. Formally, this is a case where $z\,(a,\theta) = 0$ for all (a,θ). In this cases, the only way in which individuals can access credit is via their own wealth. The choice of managerial labour and capital are given by (7). In autarky there can be wide dispersion in the marginal product of capital across entrepreneurs: an entrepreneurs' firm's capital is constrained by his personal wealth. Associated with autarky will be a price vector $\mathbf{p}^{aut}$ which clears the labour market given the occupational choice decisions.

By misallocating capital, autarky also results in lower labour demand. This in turn depresses wages. This means that wages will tend to be lower so autarky can actually encourage people to become entrepreneurs compared to a situation where capital markets are functioning well.

The First-Best We now consider what would happen with perfect capital markets. This has two dimensions. First, there is complete access to financial markets, $z(a,\theta) = 1$ for all (a,θ), and there is no moral hazard problem. In effect, the latter implies that a lender can specify a level of managerial input as part of the lending contract.

This would result in effort and capital solving

$$V^*(\theta,\mathbf{p}) = \max_{e,k}\{g(e)\pi(k;\theta,\mathbf{p}) - d(e;\theta,\mathbf{p}) - \gamma k\}$$

and capital allocation follows

$$\pi(k^*(\theta,\mathbf{p});\theta,\mathbf{p}) = \gamma/g(e^*(\theta,\mathbf{p}))$$

Note that the first best does have a level of default associated with it. However, these decisions and payoffs are independent of a, i.e. the entrepreneur's level of wealth is irrelevant.

Occupational choice is given by

$$\sigma^*(\theta,\mathbf{p}) = \begin{cases} 1 & \text{if } V^*(\theta,\mathbf{p}) > p_l \\ 0 & \text{otherwise} \end{cases}$$

which is also independent of a. Associated with first-best will be a price vector $\mathbf{p}^*$ which clears the labour market given the occupational choice, capital allocation and labour demand decisions. The wage rate will be endogenous and set to clear the labour market.

3 From Theory to Data

The model allows us to think about two main things. First, we can think about the effect of credit market frictions on optimal credit contracts. We can explore the effect of two specific frictions as represented by ϕ and τ. Second, we can look at impact of changing market access as represented by χ.

Changing market frictions affects labour demand for a given wage in (9) through three channels. First, it increases access to capital and this increases labour demand since capital and labour are complements. Second, it reduces the default probability by increasing effort. Third, it lowers the threshold productivity and wealth levels at which agents choose to become entrepreneurs. Increasing χ has a direct effect on labor demand since some entrepreneurs now get access to more capital.

General equilibrium effects are largely driven by shifts in labour demand and occupation choice which affect the wage which, in turn, feeds back on to the participation constraint of entrepreneurs and hence to the terms of credit contracts. Wages also affect the amount of managerial labour applied by changing profitability and the amount of capital used.

The model is able to give a clear sense of the different "moving parts" that affect credit market frictions in a general equilibrium model with endogenous occupational choice. Our next step is to put the model to work to explore different aspects of what the model predicts quantitatively. For this, we will need to give a specific parametrization and simulate the model's predictions which will give insights in three main areas.

Next, we describe how we apply the model by introducing the specific functional forms. We then discuss how various key parameters are calibrated.

3.1 Parametrization

The production function, $f(k, l; \theta)$ is Cobb-Douglass with diminishing returns:

$$f(k, l; \theta) = \theta^{1-\eta-\alpha} \left(l^{1-\beta} k^{\beta}\right)^{\eta}, \tag{10}$$

where θ is the firm specific productivity parameter and $\alpha, \beta, \eta \in (0, 1)$ are parameters governing the shape of the production function. Thus the model is essentially a classic Lucas-style "span of control" model η representing the extent of diminishing returns and pure profits can be thought of as payment to an untraded factor such as technology or ability.

Using this, a firm's labour demand, conditional on k, is given by:

$$l^{*}(k; \theta, \mathbf{p}) = \left[\eta(1-\beta) \frac{p_y}{p_l} \theta^{1-\eta-\alpha} k^{\eta\beta}\right]^{\frac{1}{1-\eta(1-\beta)}} \tag{11}$$

and the conditional profit function is

$$\pi(k; \theta, \mathbf{p}) = (1 - \eta(1-\beta)) \left[\left(\frac{\eta(1-\beta)}{p_l}\right)^{\eta(1-\beta)} p_y \theta^{1-\eta-\alpha} k^{\eta\beta}\right]^{\frac{1}{1-\eta(1-\beta)}}. \tag{12}$$

The marginal product of capital is therefore given by:

$$\pi_k(k; \theta, \mathbf{p}) = \eta\beta \left[\left(\frac{\eta(1-\beta)}{p_l}\right)^{\eta(1-\beta)} p_y \theta^{1-\eta-\alpha} k^{\eta-1}\right]^{\frac{1}{1-\eta(1-\beta)}} \tag{13}$$

In addition to the productivity level θ the producer's credit market access is dependent on $v\,(= u + \tau a)$, which also affects collateralizable wealth as we saw (6) above. In particular, an entrepreneur faces a cost of capital equal to $\gamma / g(\hat{e}(v; a, \theta, \mathbf{p}))$ where v is determined in a credit market equilibrium and will therefore depend on $\mathbf{p}$.

For the managerial labour technology, we also use a constant-elasticity functional form where:

$$\begin{aligned} g(e) &= \lambda e^{\alpha} \\ &\text{and} \\ d(e; \theta, \mathbf{p}) &= p_e \theta^{\delta} e, \text{with } \delta \geq 0. \end{aligned}$$

We can think of $\theta^{\delta} e$ as the amount of managerial labour required to set up the project given that is going to be successful with probability e. Each agent who works as a laborer is indifferent between standard labour (providing input l) and managerial labour; they are paid at rate p_l, or - alternatively - at a risk-adjusted wage rate $p_e\,(= p_l / g(e))$ in case of success. The parameter δ governs the dependence of the effort cost of θ, in effect the link to firm size. If $\delta = 0$, then the cost of securing a given level of default does not depend on firm size whereas $\delta > 0$ means that achieving the same default in a large firm requires more managerial input. The parameter α in the technology above governs the elasticity of the success probability with respect to managerial effort.[5] Together with the assumption in (10) this functional form implies

[5]The parameter α will be chosen such that first best default probabilities $g(e^{*}(\theta, \mathbf{p}))$ match their empirical counter-part, including at the highest level of θ. Both for lower levels of θ and in second-best default probabilities, i.e. success probabilities will be lower. Therefore no additional assumption is required to guarantee that $g(e) \in [0, 1]$.

that output has constant returns to scale in managerial labour (e), capital (k), labour (l) and managerial talent (θ). Finally, the parameter λ captures the general productivity of managerial labour at achieving project success. In the next section, we will show how to use data on the firm size distribution and heterogeneous default probabilities by firm size to calibrate $(\delta, \alpha, \lambda)$.

3.2 Calibration

Without loss of generality, we can impose values for the price vector $\mathbf{p} = (p_y, p_l, p_e)$. We will take the output price to be the same across countries and choose the unit of measurement such that $p_y = 1$. Since we will think of the price of capital goods (but not necessarily the rental rate) to be equal across countries, we measure capital, k in value terms. Further we will assume $p_e/g(e) = p_l$. Any wage or income level in the distorted model will then be measured relative to the US wage.

Model Parameters We calibrate a subset of the model parameters using evidence from existing studies. First, we assume that β, which in first best measures the share of output paid to capital relative to labour[6], is 1/3 in line with standard calibrations used in the macro-economic literature. Secondly, we take the marginal cost of capital γ to be 1.1, which roughly corresponds to long run real interest rates in the US since the 1980's (Yi and Zhang, 2016) with an allowance for capital depreciation. Thirdly, we set η to 3/4, following the assumption of Bloom (2009) in a related context.

The remaining parameters are chosen by calibrating the model to US data, assuming that this is an example of perfectly functioning credit markets. While this assumption is somewhat extreme, it may still serve as a reasonable approximation of the difference between US credit markets and developing countries' credit markets which is our main focus of attention. What makes this assumption convenient is that all of the model's predictions are independent of the asset distribution. This, in turn, allows us to calibrate the unknown parameters without knowledge of the asset distribution. We can then specify any asset distribution when we simulate second best outcomes.

Managerial Labour and the Distribution of Productivity Once we suppose that the U.S. is first best, we can calibrate the distribution of θ jointly with α and δ. We first show how α and δ determine the pattern of corporate default rates across firm sizes, conditional on the distribution of θ. The distribution of θ can be backed out from data on the distribution of firm size, conditional on α and δ. Jointly, these allow to back out the parameters affecting default risk and the cost of effort (α, δ) and the distribution of θ from the US firm size distribution and the pattern of corporate default rates across firm sizes. We normalize, without loss of generality, the US wage to be one. Further we assume $\lambda = 1.05$ for the calibration while in the simulations we will set $\lambda = 1.0$. This assumes that managerial labour productivity is 5% higher in the US than in the simulated economy.

From the first order conditions we solve for the first best level of managerial labour and capital (e^*, k^*) in closed form:

[6] Note that this only holds when defining the labour income share as payments to l, not e.

$$e^{*}(\theta,\mathbf{p}) = \left[\theta^{1-\eta-\alpha}\left(\frac{\eta(1-\beta)}{p_l}\right)^{\eta(1-\beta)}\left(\frac{\eta\beta}{\gamma}\right)^{\eta\beta}\left(\frac{\lambda\alpha(1-\eta(1-\beta))}{p_l\theta^{\delta}}\right)^{1-\eta}\right]^{\frac{1}{(1-\alpha)(1-\eta)-\alpha\eta\beta}} \tag{14}$$

$$k^{*}(\theta,\mathbf{p}) = \left[\theta^{1-\eta-\alpha}\left(\frac{\eta(1-\beta)}{p_l}\right)^{\eta(1-\beta)}\left(\frac{\eta\beta}{\gamma}\right)^{(1-\alpha)(1-\eta(1-\beta))}\left(\frac{\lambda\alpha(1-\eta(1-\beta))}{p_l\theta^{\delta}}\right)^{\alpha(1-\eta(1-\beta))}\right]^{\frac{1}{(1-\alpha)(1-\eta)-\alpha\eta\beta}} \tag{15}$$

Note that for $(\delta,\alpha) = (\frac{1-\alpha-\eta}{1-\eta},\alpha)$ the first best effort level, and therefore the default probability, is independent of the scale of the firm θ. Other levels of (δ,α) imply that first best default probabilities increase or decrease with first best firm size. Given a distribution of productivities, α and δ can then be chosen such that the implied pattern of default probabilities across firm sizes matches the empirical pattern. In particular, we calibrate α and δ such that the smallest firm operating in equilibrium has a default probability of 0.10 and the largest firm has a default probability of 0.01. Note that in our model any default implies full "charge-off". Hence we suppose that default rate is best approximated by the charge-off rates of corporate loans which are approximately 0.8 percentage points over the last 30 years, see Board of Governors of the Federal Reserve (2016).[7]

Next we show how the marginal distribution of θ can be calibrated from data on the distribution of firms sizes, conditional on α and δ. Plugging (15) into (11) we can write equilibrium labour demand, l^*, as a function of θ up to a constant of proportionality. Inverting this relationship, we have that:

$$\theta = (l^*)^{\psi}\cdot\Psi \tag{16}$$

where Ψ and ψ are known constants. Equation (16) shows that the distribution of θ conditional on entrepreneurship can be backed out from data on the distribution of the firm level labour force, l^*. Empirically the distribution of firm sizes measured in terms of the size of the labour force l^* is well approximated by a Pareto distribution, with shape parameters $\sigma_l = 1.059$ (Axtell, 2001). Given the functional form in (16), θ also follows also a Pareto distribution with known shape parameter.[8] We take both the firm size and θ distributions to follow upper-truncated Pareto distribution, where the point of truncation is defined by the largest firm observed in the Axtell (2001) dataset. Note that this does not pin down the scale parameter of the θ distribution, $\underline{\theta}$, since l^* is only observed for firms with $\sigma(a,\theta,\mathbf{p}) = 1$. We choose $\underline{\theta}$ to clear the labour market, i.e. solve $L^S(\mathbf{p}) = L^D(\mathbf{p})$ at $p_l = 1$, i.e. we assume that US labour markets are in equilibrium and find the distribution of θ such that the equilibrium wage predicted by the model matches the observed US wage.

In what we describe above, the calibration of (δ,α) is conditional on the distribution θ, and vice versa. We find the values of δ, α and the distribution of θ to simultaneously to match the specified pattern of default probabilities, the observed firm size distribution, and imply that labour markets clear at wage equal to 1.

[7] Delinquency rates are higher, mechanically.

[8] A Pareto distribution with scale parameter $\underline{l}$ and shape parameter σ_l has a c.d.f. $P(L \leq l) = 1 - \left(\frac{\underline{l}}{l}\right)^{\sigma_l}$. We find the c.d.f. of θ as $P(t \leq \theta) = 1 - \left(\frac{(\underline{\theta}/\Psi)^{\frac{1}{\psi}}}{(\theta/\Psi)^{\frac{1}{\psi}}}\right)^{\sigma_l} = 1 - \left(\frac{\underline{\theta}}{\theta}\right)^{\frac{\sigma_l}{\psi}}$. This is again a Pareto distribution, with shape parameter σ_l/ψ and lowest value as $\underline{\theta} = \underline{l}^{\psi}\Psi$.

The Distribution of Wealth We can specify the marginal asset distribution to follow any observed or hypothetical wealth distribution. For our baseline simulations we choose the marginal distribution of assets to approximate the wealth distribution in India. We obtained data on the Indian wealth distribution from the Global Wealth Report 2015 (Credit Suisse, 2015). This provides information on the Gini coefficient of the Indian wealth distribution, mean wealth, median wealth and the fraction of the population in four wealth classes: 0-10k, 10k-100k, 100k-1m and over 10m USD. The median wealth in India is 1.75% of median wealth in the US, and the mean wealth is 1.24% of mean wealth in the US. We assume the Indian wealth distribution to be of the Pareto family, which has been shown to be a reasonable approximation in a number of countries. This reduces the calibration to choosing a shape and scale parameter of that distribution. Moreover, given the Pareto assumption, the shape parameter has a known monotonic relation to the Gini coefficient. We use this relation together with the aforementioned data on the empirical Gini coefficient to back out the shape parameter. Specifically, the scale parameter is chosen to minimize the sum of squared differences between the empirical probability mass and the probability mass of the calibrated Pareto distribution in each of the four wealth categories, where the summation is across wealth categories.

Lastly, we need to specify the joint distribution $h(a, \theta)$ of assets and productivities. This is difficult to back out non-parametrically from data. In a world with first best credit contracts, knowledge of individual wealth levels, occupational status, and the size of the labour force of firms held by entrepreneurs, would be sufficient to back out the joint distribution of a and θ for the subset of individuals with a θ high enough to become entrepreneurs. However, for all individuals with a value of θ that does not lead to them becoming entrepreneurs, θ is fundamentally unobserved. In our simulations we therefore work with several hypothetical joint distributions.

To this end, we can specify a pattern of dependency between a and θ using the statistical concept of copulas.[9] According to Sklar's theorem (Sklar, 1959), the multivariate density function $h(a, \theta)$ can be rewritten as $h(a, \theta) = h_a(a) \cdot c(H_a(a), H_\theta(\theta)) \cdot h_\theta(\theta)$, where $H_a(\cdot)$ and $H_\theta(\cdot)$ are the cumulative density functions of the marginal distribution of a and θ, respectively, $h_a(\cdot)$ and $h_\theta(\cdot)$ are the corresponding probability density functions, and $c : [0, 1]^2 \to R^+$ is the density function of the copula. We assume that the dependency between a and θ is characterized by a Normal copula. This implies that the only free parameters that have to be specified is the covariance which we choose such that the induced correlation between a and θ matches one of a range of "target values" of the correlation: $\rho \in \{0.0, 0.05, 0.1, 0.2, 0.3\}$. As we increase ρ we are postulating a stronger and stronger link between productivity and wealth. Using this approach, we can simulate our model given each value of ρ to trace out the implications of different degrees of correlation between a and θ for credit market outcomes.

3.3 Computation

In order to compute the model, we approximate the continuous distribution of a and θ by a distribution with 1000 and 10000 discrete values, respectively, both in the calibration and the subsequent simulations. These discrete values approximately represent equally spaced centiles of the continuous distribution.

When calibrating the model we solve jointly for the distribution of productivities, α, δ using

[9] See Nelson (1999) and Trivedi and Zimmer (2007) for accessible introductions.

an iterative process as follows. We start from an initial trial value of the parameters affecting default risk and the cost of effort, (δ, α), and then find the distribution of productivity, θ, to match the empirical firm size distribution and ensure that the labour markets clear at a wage (p_l) of one as described above. Conditional on this distribution of θ we then update the value of (δ, α) to generate default probabilities of 0.1 and 0.01 for the smallest and largest firms which are active in the equilibrium. We then iterate this process until the values of both α and δ converge in the sense that their values change each by less than 0.1 percentage points relative to the previous iterations.

The core problem of the simulations is to find the equilibrium wage at each level of ρ, τ and ϕ. We implement this computationally using the bisection method. A wage is accepted as a solution once labour demand relative to labour supply deviates by less than 0.001 from 1. Given any ρ, τ and ϕ and wage, the simulations involve computing the credit contracts for each of the 1000×10000 tuples for (a, θ). In order to speed up the computation, we make use of the result that if a potential entrepreneur decides to become a worker at (a, θ), all individuals with the same productivity and lower wealth will also choose to become workers.

4 Results

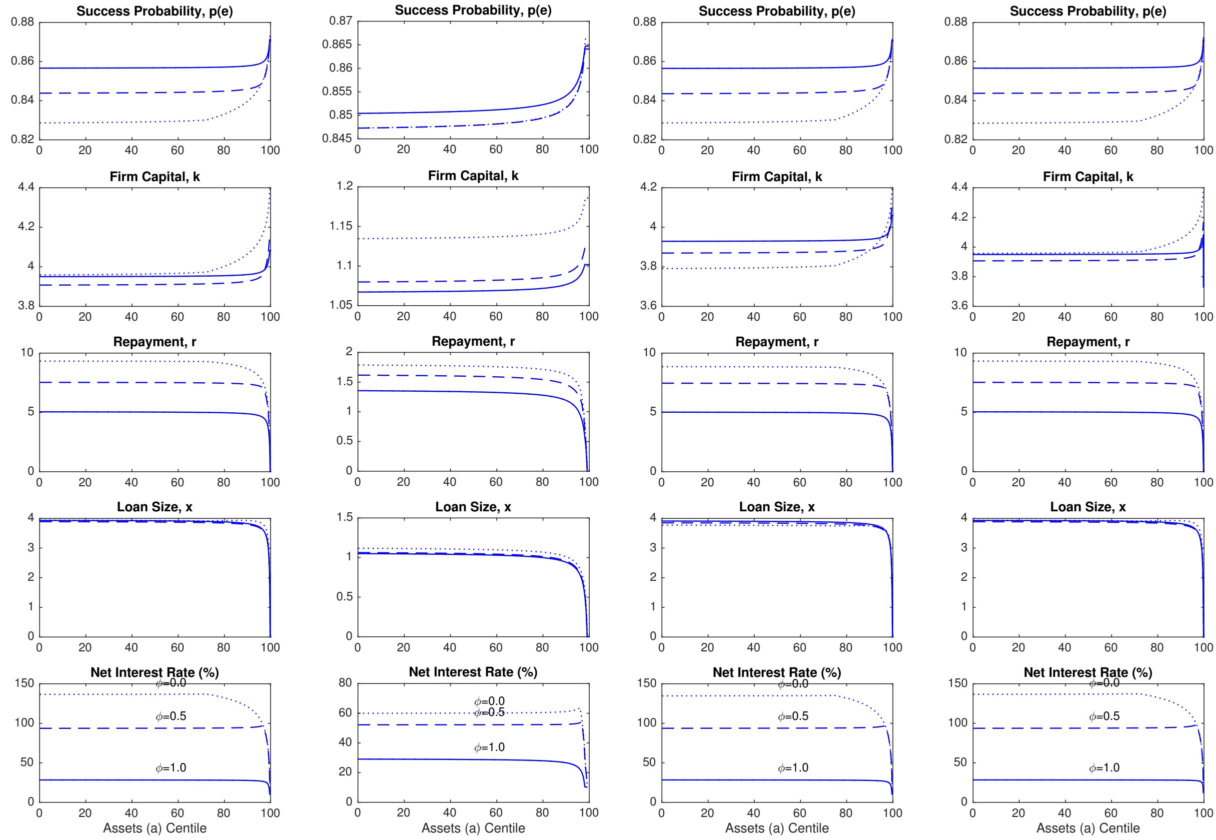
For the results that follow, we will consider an economy where the productivity distribution is based on the US and the wealth distribution on India as detailed in the previous section. The benchmark that we study has no correlation between wealth and productivity ($\rho = 0$). For the core results presented here, we set $\lambda = 1$ so that our US benchmark is 5% more productive than the economy that we are studying translating managerial effort into repayment success. We then set $\tau = 1$ so that property rights to wealth are perfect, i.e. all wealth can be used as collateral. All of these assumptions will be maintained in what follows unless we explicitly state otherwise. Capital can be acquired by lenders at borrowing rate of 10% so that $\gamma = 1.1$. In the first best, the marginal product of capital will be equal to this.

4.1 Credit Contracts

Baseline We begin by looking at credit contracts and capital allocation and how these vary with an entrepreneur's position in the wealth distribution. In all cases, we take a highly productive entrepreneur (at the 99th percent of the productivity distribution). Since around 5% of the population are entrepreneurs when there is full access to credit markets, this constitutes the top 20% in the distribution of entrepreneurial productivity and corresponds to a firm size of around 9 employees. This may still seem quite small. However, the firm-size distribution implied by the calibrated values is highly skewed. Such individuals are always active as entrepreneurs in our calibration even if they have little wealth and hence we do not need to worry about their occupational choice as we vary parameter values. In all cases, Figure 1 illustrates the outcome for different values of the parameter ϕ. Recall that $\phi = 0$ is the lowest level of competition and $\phi = 1$ is the highest. We also give the contracts for a middle level of competition: $\phi = 1/2$, half the surplus goes to the entrepreneur and half to the borrower.

When interpreting the figures that follow, it should be borne in mind that there are two effects of changing the level of competition. The first is a direct effect whereby the entrepreneur's share of the surplus varies. This affects the total amount of surplus to the extent that incentives

FIGURE 1: EQUILIBRIUM CREDIT CONTRACTS



(A) BASELINE (B) θ = 96TH PERCENTILE (C) ρ = 0.3 (D) τ = 0.50

Notes: The figures describe characteristics of the equilibrium credit contracts for individuals across all centiles of the asset distribution, holding constant θ. Subfigure (A) presents the baseline scenario where we impose perfect collateralisability of wealth ($\tau = 1$), assume that the distribution of asset holdings and productivities are independent ($\rho = 0$) and presents results for the 99th centile of the productivity distribution. Subfigures (B), (C) and (D) preserve the baseline scenario with one exception each: In subfigure (B) we presents results for the 96th centile of the productivity distribution, in subfigure (C) we assume that the distribution of asset holdings and productivities are correlated ($\rho = 0.3$) and in subfigure (D) we impose imperfect collateralisability of wealth ($\tau = 0.5$). Firm outcomes are shown for three distinct levels of competitiveness of credit markets: full competition ($\phi = 1.0$), monopolistic competition ($\phi = 0.0$), and an intermediate level ($\phi = 0.5$). Credit contracts are characterised by the gross interest payment r and the loan size x. Both are measured in absolute terms and in units of the annual income of a US wage labourer (taken to be 43k US$). The net interest rate is calculated as $(r/x - 1) \times 100$. Together with the entrepreneurs own capital x determines the firm capital k. The level of r and x also determine through the effort level the projects success probability $g(e)$.

for managerial effort vary. The second is a general equilibrium effect of competition. Changing competition affects aggregate labour demand, L_D and hence wages thus also changing entrepreneurial profits.

In the top panel of Figure 1, the default probability which depends on managerial labour hired by the entrepreneur is illustrated for different wealth levels and competition. Two things are immediate. First, the default rates implied by the model are around 15% across the wealth distribution. This reflects the fact that, in general equilibrium, even low wealth entrepreneurs face good outside options (even when competition is low). This is because, for marginal entrepreneurs, this is the option of being a wage labourer and for higher wealth individuals this is the possibility of self-finance. The default rate is flat across most of the wealth distribution but then decreases for very high wealth individuals who are closer to first-best self-financing. Competition does have real effects since default is lower when there is more competition. This is because the payoffs to entrepreneurs from being successful are higher when there is greater competition.

The second panel gives the use of capital as a function of the position in the wealth distribution. We know from equation (6) that this is the flip-side of the repayment probability as illustrated in the top panel. A higher repayment rate naturally means more capital as the marginal product of capital will be lower.

Firm capital (k) is lower for lower levels of assets. However, for most wealth levels, the effect of competition on firm capital is non-monotonic. Moving competition for $\phi = 0$ to $\phi = 1/2$, decreases firm capital for individuals at most percentiles of the wealth distribution. However, capital usage typically increases for the move from $\phi = 1/2$ to $\phi = 1$. At the highest centiles of the wealth distribution increased credit market competition leads monotonically to a decrease in capital usage. For these entrepreneurs the increased capital market competition does not lead to substantially improved credit access, and the positive effect of credit market competition on wages depresses capital usage. Thus the model predicts a heterogeneous, non-linear and often non-monotonic effect of competition on capital allocation which could only be seen by disaggregating by wealth level. That said, the magnitude of these effects is relatively modest, i.e. around a 5% decrease in the amount borrowed for high wealth individuals when competition moves from $\phi = 0$ to $\phi = 0.5$.

The third panel gives the amount repaid by the entrepreneur for the loan that she takes out and the fourth panel gives the loan size. The latter shows that the amount borrowed is almost unaffected by competition. Moreover, the amount borrowed does not depend much on wealth except at very high wealth levels where self-financing substitutes for credit. This pattern reflects the fact the marginal product of capital does not move much with assets in our calibrations.

The repayment level by contrast does vary quite a bit with competition and is highest when competition is low. This reflects the division of surplus between the lenders and entrepreneurs. In highly uncompetitive environments a good amount of the profits that entrepreneurs make are captured by investors. The only limit on this process when $\phi = 0$ is the outside option available and/or the possibility that an entrepreneur receives his efficiency utility. The ability to capture entrepreneurial surplus is diminished for high wealth investors since they have a very good autarky outside option.

In popular discussion, the interest rate is frequently used as a barometer of credit conditions. In general our model shows that this is a poor sufficient statistic for matters which should be gauged from capital allocation and surplus sharing. The reason why the interest rate, $(r/x - 1) \times 100$, is a very poor indicator of capital allocation is that both numerator and denominator are functions of the default rate and the underlying source of heterogeneity (a, θ).

However, it is still interesting to see what the model predicts and how well it relates to efficiency and distribution in the credit market. And we know from many studies of developing country credit markets that interest rates charged by monopolistic borrowers can be very high. With very low competition ($\phi = 0$), the model predicts an interest rate of between 40% and 150% for almost all wealth classes. Only for very high wealth individuals is the rate less than this and it falls quite rapidly for the top of the distribution even with low competition which is due to the fact that outside options for such borrowers are very good (if they need credit at all). The interest rate profile for middle levels of competition is also comparatively flat but again turns down for very high wealth levels, though it should be noted that the horizontal scale is in terms of asset centiles, not absolute values. For the highest level of competition, the interest rate is consistently quite low. So competition does seem to have a significant bearing on the interest rate offered to borrowers.

Lower Productivity Benchmark We now consider what happens when we look at a more marginal group of entrepreneurs by focusing on the 96th percentile in the productivity distribution. Such entrepreneurs typically employ around two workers so are a quarter of the size of the firms in the baseline case. We will look at how changing this focus affects credit contracts and credit allocation. These results are depicted in Subfigure (B) of Figure 1.

Note first the relationship between competition and default probabilities is much less for these marginal entrepreneurs. However, we still see that at high levels of wealth (above the 80th percentile of the distribution), the repayment probability rises quite steeply. Capital allocation is now much lower (due to productivity being lower) but, in common with the baseline, it is very flat during low levels of the wealth distribution. At the highest levels of the asset distribution both the success probability and the firm capital are independent of assets. Here the first best allocation is achieved. Although repayment and loan size are lower, the same broad relationship with wealth and competition is observed as in the higher productivity case. Interest rates are lower for these less productive borrowers, which is driven by the outside option of both wage labour and self-financing being more attractive in relative terms for these borrowers.

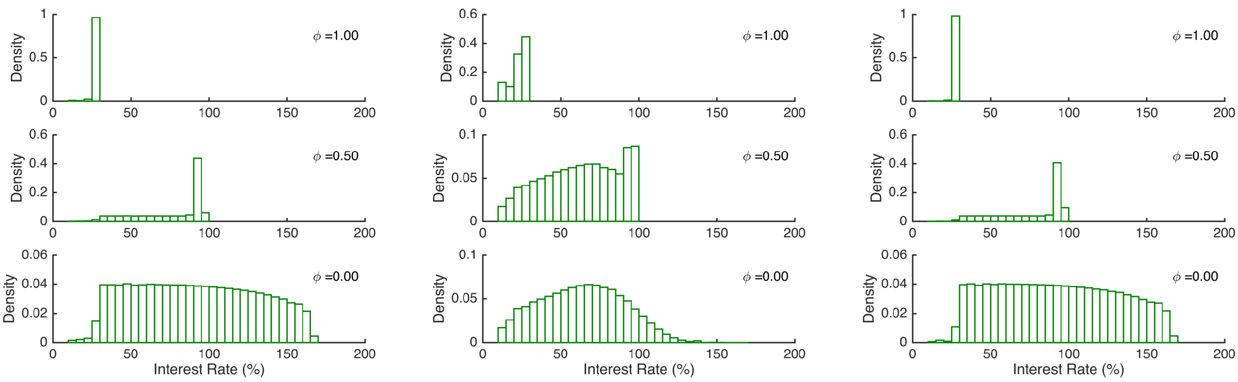
Overall, we find a common pattern that, with optimal second best credit contracts, wealth does not have a strong quantitative effect on the allocation of resources over a wide range of wealth levels. This is because, in a general equilibrium setting, the outside option of wage labour and/or the possibility of receiving an efficiency utility level, does most of the work. This is a general lesson from our model and would only be found by taking a general equilibrium perspective which solves explicitly for the outside options that entrepreneurs face.

Higher Correlation Between Productivity and Wealth In subfigure (C) of Figure 1, we look at what happens when we allow for a stronger correlation between wealth and productivity, which increases the density of high wealth/high productivity individuals. The optimal contracts for each (a, θ) are essentially preserved compared to the baseline. The only exception is a change in the allocation of capital. We now find a monotonic positive effect of credit market competition on firm capital for a wide range of low asset individuals. This is driven by a general equilibrium effect of high productivity individuals having likely also high wealth. This decreases the dependence of high productivity individuals on credit access, and increases wages. Improvements in credit market competition have less of an effect on wages. This in turn means that, for most wealth levels, the positive effect of increased competition on credit access dominates the weak general equilibrium effect on wages throughout all levels of competition.

Imperfect Property Rights We now consider what happens when $\tau = 0.5$ so that only 50% of wealth can be used as collateral. In effect, collateralizable wealth in the economy is halved. This constitutes the de Soto effect in this model. We already have a hint from the first panel that this will not matter much for the lower wealth part of the distribution as we observed that over the range 0 to 80%, most variables – including the repayment rate and capital do not vary with wealth. Given this, we would not expect a strong equilibrium effect in the economy. Subfigure (D) of Figure 1 confirms that this is not the case and there is virtually no effect of changing property rights affecting the use of collateral in this setting. This, perhaps surprising, finding can be put down to the general equilibrium setting where the (endogenous) wage rate plays a key role in determining the outside option of borrowers. Hence, even when competition is low, borrowers have relatively little capacity to exploit their market power. Moreover, the possibility of an efficiency utility also limits the impact of the outside option on contracts at lower wealth levels. At higher wealth levels the possibility of self-financing provides a relevant outside option.

The Distribution of Interest Rates and Default Probabilities Figure 2 also looks at interest rates but this time, the distribution of such rates across types of borrowers with different levels of competition. As we would expect from Figure 1, when competition is very high then there is no variation in interest rates at all. A feature of the high competition case is the emergence of a modal interest rate; almost every borrower is being offered the same interest rate. The spread of interest rates on offer starts to increase as competition is reduced. However, with $\phi = 0.5$, there is still quite a bit of bunching.

FIGURE 2: DISTRIBUTION OF INTEREST RATES



(A) BASELINE (B) $\rho = 0.3$ (C) $\tau = 0.5$

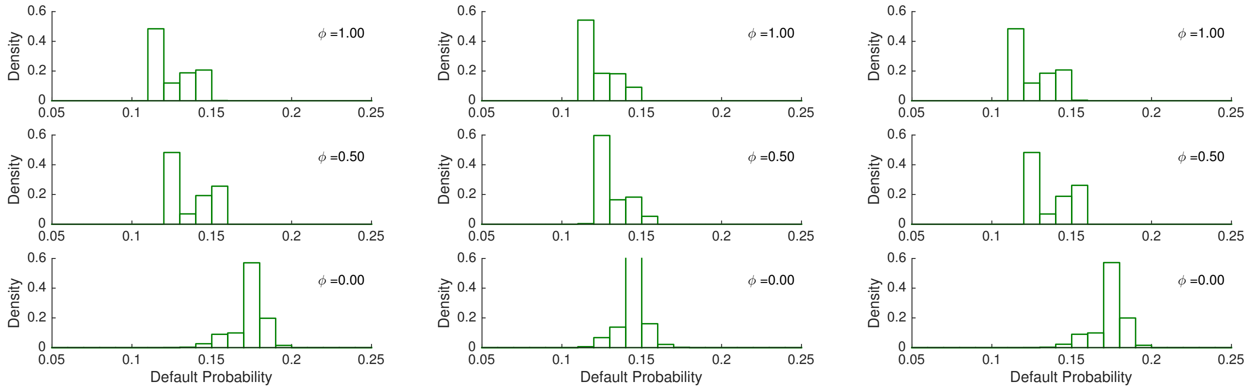
Notes: All graphs depict the distribution of interest rates ($(r/x - 1) \times 100$) paid by individuals who borrow in equilibrium. Subfigure (A) presents the baseline scenario where we impose perfect collateralisability of wealth ($\tau = 1$), assume that the distribution of asset holdings and productivities are independent ($\rho = 0$). Subfigures (B) and (C) preserve the baseline scenario with one exception each: in subfigure (B) we assume that the distribution of asset holdings and productivities are correlated ($\rho = 0.3$) and in subfigure (C) we impose imperfect collateralisability of wealth ($\tau = 0.5$). For each scenario we show the distribution of interest rates for three distinct levels of competitiveness of credit markets: full competition ($\phi = 1.0$, top figure), monopolistic competition ($\phi = 0.0$, bottom figure), and an intermediate level ($\phi = 0.5$, middle figure).

When $\phi = 0$, i.e. no competition, then there is the widest range of interest rates which are essential chosen to extract all of the surplus from a lending relationship between an entrepreneur and lender. Variation in the interest rate are caused by a subset of entrepreneurs having attractive outside options, either as wage labourers or by self-financing.

Figure 2 also allows these distributions to vary across three cases: higher productivity-wealth correlation and worse property rights. The latter, as above, leaves things almost unchanged. However, the effect of increase ρ, which tends to increase the wage, is more visible with a greater spread in interest rates. This partly reflects that a larger number of entrepreneurs has attractive outside options, given our assumption on the correlation of θ and a.

In Figure 3, we look at the distribution of the credit market distortion across borrower types. A sufficient statistic for this is the default probability $1 - g(e)$. We present figures showing this for different levels of competition. When competition is highest (in the top panel), then there is a modal outcome with only a few borrowers having lower default probabilities (those with higher wealth). As competition is reduced across the second and third panels, this mode shifts to the right (higher default) but the same broad pattern occurs. The distribution visibly widens

FIGURE 3: DISTRIBUTION OF DEFAULT PROBABILITY



(A) BASELINE (B) $\rho = 0.3$ (C) $\tau = 0.5$

Notes: All graphs depict the distribution of default probabilities $(1 - g(e))$ of individuals who borrow in equilibrium. Subfigure (A) presents the baseline scenario where we impose perfect collateralisability of wealth ($\tau = 1$), assume that the distribution of asset holdings and productivities are independent ($\rho = 0$). Subfigures (B) and (C) preserve the baseline scenario with one exception each: in subfigure (B) we assume that the distribution of asset holdings and productivities are correlated ($\rho = 0.3$) and in subfigure (C) we impose imperfect collateralisability of wealth ($\tau = 0.5$). For each scenario we show the distribution of default probabilities for three distinct levels of competitiveness of credit markets: full competition ($\phi = 1.0$, top figure), monopolistic competition ($\phi = 0.0$, bottom figure), and an intermediate level ($\phi = 0.5$, middle figure).

with a greater reduction in default for higher wealth entrepreneurs. When $\phi = 0$ (the lowest panel), i.e. no competition, then there is a wider distribution of default probabilities. This spreading out occurs in the right tail as more entrepreneurs who face worse outside options hire less managerial labour. These distributions vary somewhat with a change in ρ but are largely insensitive (as with all of our results) with variation in τ.

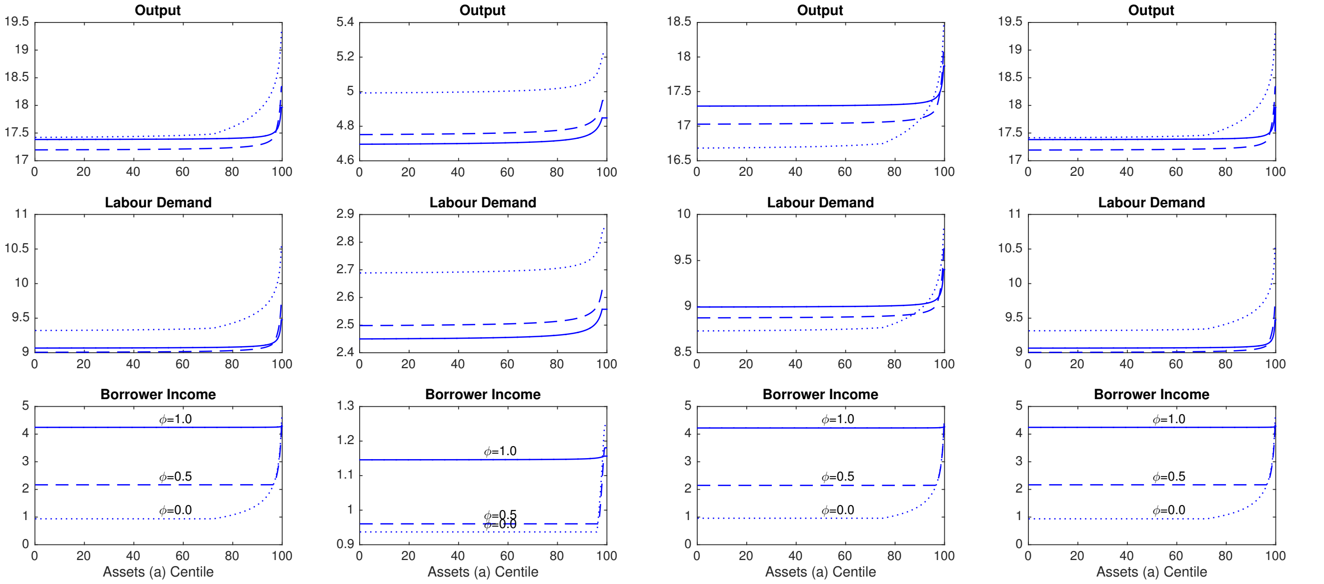
Output In Figure 4, we look at outcomes among entrepreneurs for the same core parameter values that we looked at in the previous section, i.e. θ in the 99th percentile. We look at output, labour demand and the income of entrepreneurs. Once again, we look at this in four main cases – the same as those studied when look at credit contracts and capital allocation.

In the top panel, we look at output (for successful entrepreneurs) for various wealth groups. This shows that output is largely unaffected by wealth up to the 80th percentile of the wealth distribution regardless of the level of competition. However, for the lowest level of competition, output increases from about the 80th percentile onwards. With high competition this effect is only apparent for wealth levels in the top 5 percentiles of the wealth distribution. One striking feature of this panel is that output is highest when competition is lowest. This may seem paradoxical. However, it is worth bearing in mind that the wage increases when competition increases which squeezes the income of entrepreneurs. So entrepreneurs actually produce more when competition is lower. Another feature of this top panel is that once again, the effect of changing competition is highly non-linear with an effect only observable between $\phi = 1/2$ and $\phi = 0$.

The second panel in Figure 4 looks at labour demand. This is highest when competition is lowest since the wage rate is lowest then. It is clear that taking a general equilibrium model is needed to bring this out. If ϕ could be increased for one entrepreneur rather than for all at once, then the wage would be unchanged and labour demand would fall when competition is decreased. There is no visible effect comparing $\phi = 1/2$ and $\phi = 1$.

In the third panel, we look at the income of entrepreneurs. Naturally this depends strongly on the level of competition as the share of the surplus is affected by this. However, it is fairly

FIGURE 4: EQUILIBRIUM ENTREPRENEURS



(A) BASELINE (B) θ = 96TH PERCENTILE (C) ρ = 0.3 (D) τ = 0.5

Notes: The figures describe characteristics of the equilibrium firm level outcomes for entrepreneurs across all centiles of the asset distribution, holding constant θ. Subfigure (A) presents the baseline scenario where we impose perfect collateralisability of wealth ($\tau = 1$), assume that the distribution of asset holdings and productivities are independent ($\rho = 0$) and presents results for the 99th centile of the productivity distribution. Subfigures (B), (C) and (D) preserve the baseline scenario with one exception each: In subfigure (B) we presents results for the 96th centile of the productivity distribution, in subfigure (C) we assume that the distribution of asset holdings and productivities are correlated ($\rho = 0.3$) and in subfigure (D) we impose imperfect collateralisability of wealth ($\tau = 0.5$). Firm outcomes are shown for three distinct levels of competitiveness of credit markets: full competition ($\phi = 1.0$), monopolistic competition ($\phi = 0.0$), and an intermediate level ($\phi = 0.5$). Expected output is $\hat{e}^{\alpha}\theta^{1-\eta-\alpha}(\hat{l}^{1-\beta}\hat{k}^{\beta})^{\eta}$, measured in units of average annual wage income. Total expected firm level labour demand is $\hat{e}^{\alpha}l + \theta^{\delta}\hat{e}$, measured in units of one persons annual labour supply. Expected borrower income is $\hat{e}^{\alpha}(\Pi(\hat{k};\theta) - \hat{r}) - (1-\hat{e}^{\alpha})\hat{c} - \theta^{\delta}\hat{e}w - \gamma\kappa a$, measured in units of average annual wage income. Both are measured in absolute terms and in units of the annual income of a wage labourer in the US.

flat with respect to the level of wealth except when there is very low competition when it turns upwards at the highest wealth levels.

The remaining panels look at the same variants as in Figure 1. As we might expect, Subfigure (B) shows that having smaller entrepreneurs scales down the size of incomes and labour demand. Competition increases labour demand and income. Increasing the correlation between wealth and productivity implies that competition has a negative effect on labour demand for most wealth levels. This is driven by the same general equilibrium effect that also drives the pattern of firm capital in Figure 2. When competition is low, now labour demand shifts down more noticeably with the level of competition. Finally, in subfigure (D), we continue to see little evidence of a general equilibrium de Soto effect.

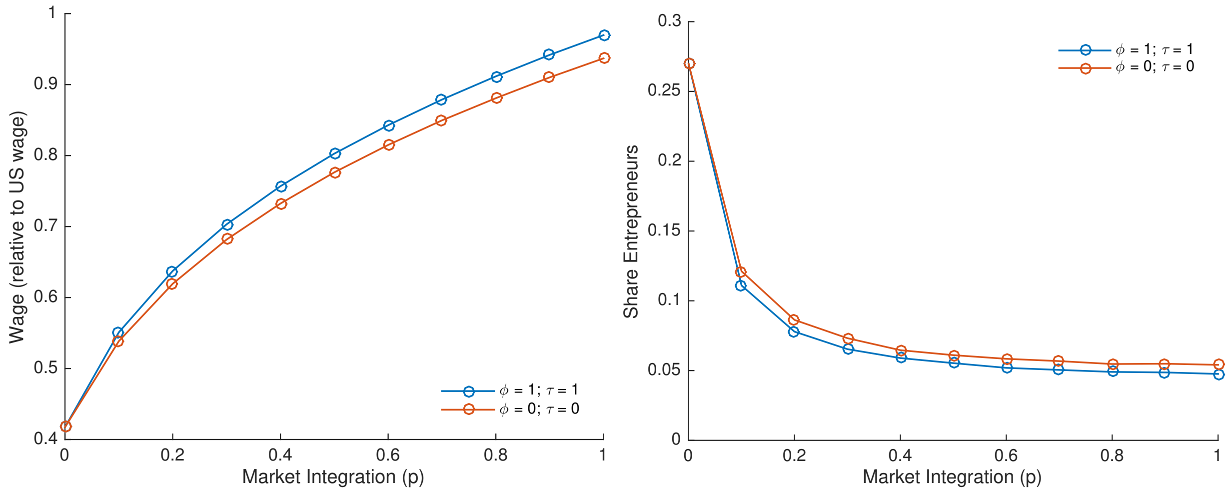
4.2 Expanding Market Access

In this section, we look at aggregate implications of extending financial inclusion. We consider different values of $\chi(a, \theta)$ assuming that this is independent of (a, θ) but we allow credit to be extended to a wider and wider set of individuals.

Wages and Self-Employment Figure 5 gives the core aggregate outcomes. The first is the wage as a fraction of the US wage moving form Autarky through to full credit market access. The second gives the fraction of the population that become entrepreneurs. We illustrate this for two cases. The first, the blue lines in Figure 5 sets $\phi = 1$ (full competition) and $\tau = 1$

(full collateralizability of wealth) which is the best possible scenario for credit markets. The second, the red line in Figure 5, sets $\phi = 0$ (no competition) and $\tau = 0$ (no collateralizability of wealth).

FIGURE 5: AGGREGATE IMPLICATIONS OF MARKET INTEGRATION AND MARKET IMPERFECTIONS



(A) WAGE RATE

(B) ENTREPRENEURSHIP

Notes: This figure presents the equilibrium wage rate (Figure A) and the share of entrepreneurs in the population (Figure B) across levels of market integration, ranking from autarky ($p = 0.0$) to full market integration ($p = 1.0$). In each figure we present the outcome of interest for the case of perfectly functioning credit markets, subject to credit market access existing ($\phi = 1.0; \tau = 1.0$) and the case of imperfectly functioning credit markets, subject to credit market access existing ($\phi = 0.0; \tau = 0.0$). Throughout we assume that the distribution of asset holdings and productivities are independent ($\rho = 0$).

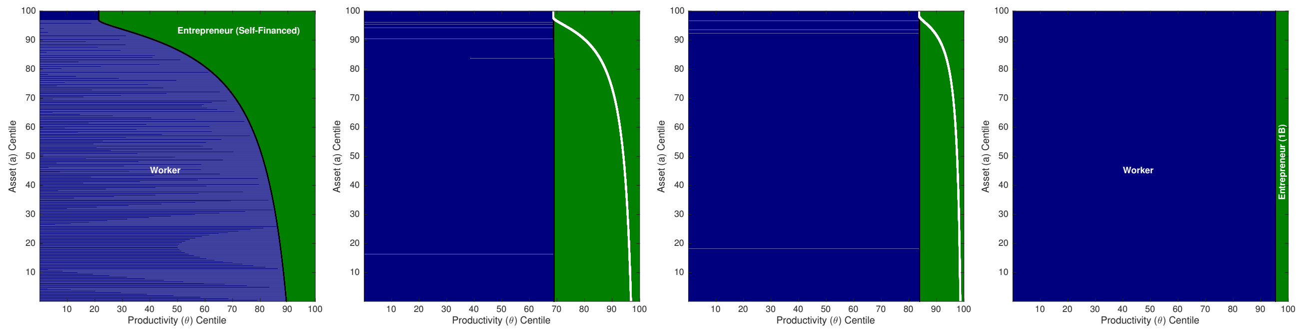
The left hand panel of Figure 5 shows that the wage moves from around 40% of the US wage in autarky to over 90% when there is full credit market access. The upwards sloping curve shows a concave relationship. However, there are still good-sized gains when moving for example, from 40% to 80% access. Comparing the blue and red lines, we find that moving from the least to most efficient credit markets (conditional on a level of access) leads to modest aggregate wage gains although (naturally) the gap between these lines increases as credit market access expands.

The right hand panel in Figure 5 shows the proportion of the population that is an entrepreneur. In autarky this is around 27% of the population. However, it falls rapidly as credit market access expands and the levels off from around 40% access at a little above 5%. Individuals are driven out of self-employment by the increasing wage which makes become a wage labourer more attractive and squeezes the profits of marginal entrepreneurs. This figure illustrates why looking at the rate of self-employment is not a good guide to economic outcomes. Individuals are only self-employed because wages are low and they lack access to borrowing opportunities. Thus, allowing for capital to flow to its most productive uses will ensure that only the most productive entrepreneurs (regardless of their wealth) will become entrepreneurs and employment will be concentrated in such firms. Thus the gain are not because the economy has become more intrinsically productive just because gains from trade in labour and capital are realized.

Occupational Choice In Figure 6, we look at the occupational choice at all points in the (a, θ) distribution. The blue shaded area illustrates the space in which individuals choose to be workers and the green shaded area where they choose to be entrepreneurs. Almost all of these entrepreneurs choose to borrow. However, those with very high wealth choose to self-finance.

On the extreme left and right hand panels, we illustrate autarky and the first best. The left hand panel shows, not surprisingly, that the low wealth and low productivity individuals are all workers. However, of the 27% or so who choose to become entrepreneurs, this goes quite far down the productivity distribution for people with high wealth. In the first best, only around the top 4% of the productivity distribution becoming entrepreneurs. Moreover, initial wealth does not matter now since capital allocation in a firm is not dependent on this, only on the productivity of an enterprise.

FIGURE 6: OCCUPATIONAL CHOICE



(A) AUTARKY | (B) $p = 0.10$ | (C) $p = 0.25$ | (D) FIRST BEST

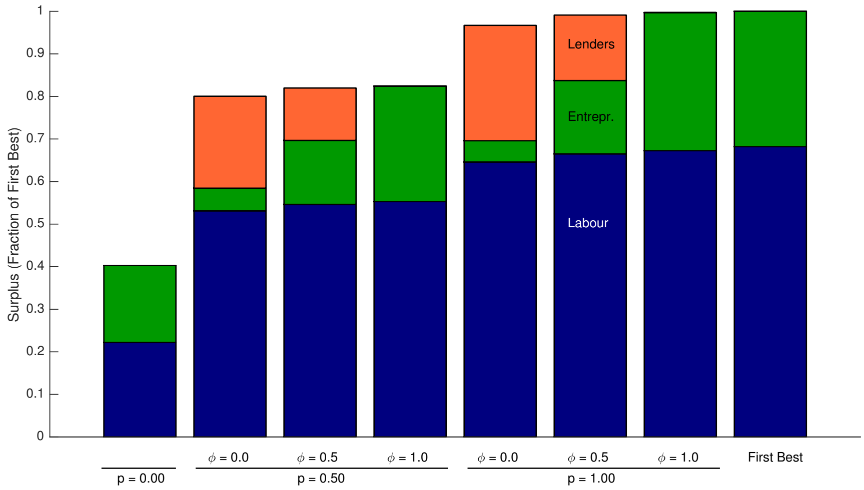
Notes: These figure depicts the occupational choice and lending decision in the a - θ space. Throughout we assume $\rho = 0$. In Figure 6a we present the occupational choice in autarky; in Figure 6b we assume $p = 0.1$ and ($\phi = 1.0; \tau = 1.0; \kappa = 1.0$); in Figure 6c we assume $p = 0.25$ and ($\phi = 1.0; \tau = 1.0; \kappa = 1.0$); and in Figure 6d we present results in first best. Blue areas indicate individuals who become workers when they have no credit market access. Green areas indicate individuals who become entrepreneurs when they have no credit market access. The while line indicates the occupational choice of individuals with credit market access: individuals to the lower-left of it become workers, individuals to the upper-right of it become entrepreneurs.

In the middle two panels, we illustrate the occupational choice for two intermediate values of credit market access where credit contracts are second-best optimal. We choose $\chi = 0.10$, i.e. 10% of the population has access to credit markets and $\chi = 0.25$ where it is 25%. They show how as credit market access expands, there are fewer entrepreneurs. Even with very limited access, there is switch away from low-productivity high wealth individuals choosing to become entrepreneurs. Hence, the selection effect is quite powerful. The line between the blue and green areas is close to vertical. It then shifts to the right as credit market access expands.

Competition and Distribution We now look at the division of income between labourers, lenders and entrepreneurs. We will consider this as we vary credit market access, χ, and the level of competition, ϕ. The size of the columns in Figure 7 illustrate the level of income as a fraction of US income.

On the extreme left is autarky. Here, a little less than half of national income is in the form entrepreneurial profits. The next three bars are for the case where 50% of the population have access to credit markets but for three levels of competition. Note, however, that with $\phi = 0$, then most of the gains from credit markets are appropriated by lenders and entrepreneurial profits are squeezed relative to autarky. The main effect of increasing competition is to redistribute surplus between lenders and entrepreneurs. There is a modest increase in the wage, due to competition with most of the gain (as we would expect from Figure 5) coming from increasing market access. A similar pattern of surplus redistribution is found for the case of full credit market access.

FIGURE 7: DISTRIBUTION OF SURPLUS ACROSS LEVELS OF COMPETITIVENESS



Notes: This figure depicts the size and distribution of total surplus in the economy, in units of first best surplus. It depicts these across levels of market integration, ranking from autarky ($p = 0.0$) to full market integration ($p = 1.0$), and for distinct levels of competitiveness of credit markets, ranging from monopolistic competition ($\phi = 0.0$) to full competition ($\phi = 1.0$). Throughout we impose perfect collateralisability of wealth ($\tau = 1$) and assume that the distribution of asset holdings and productivities are independent ($\rho = 0$).

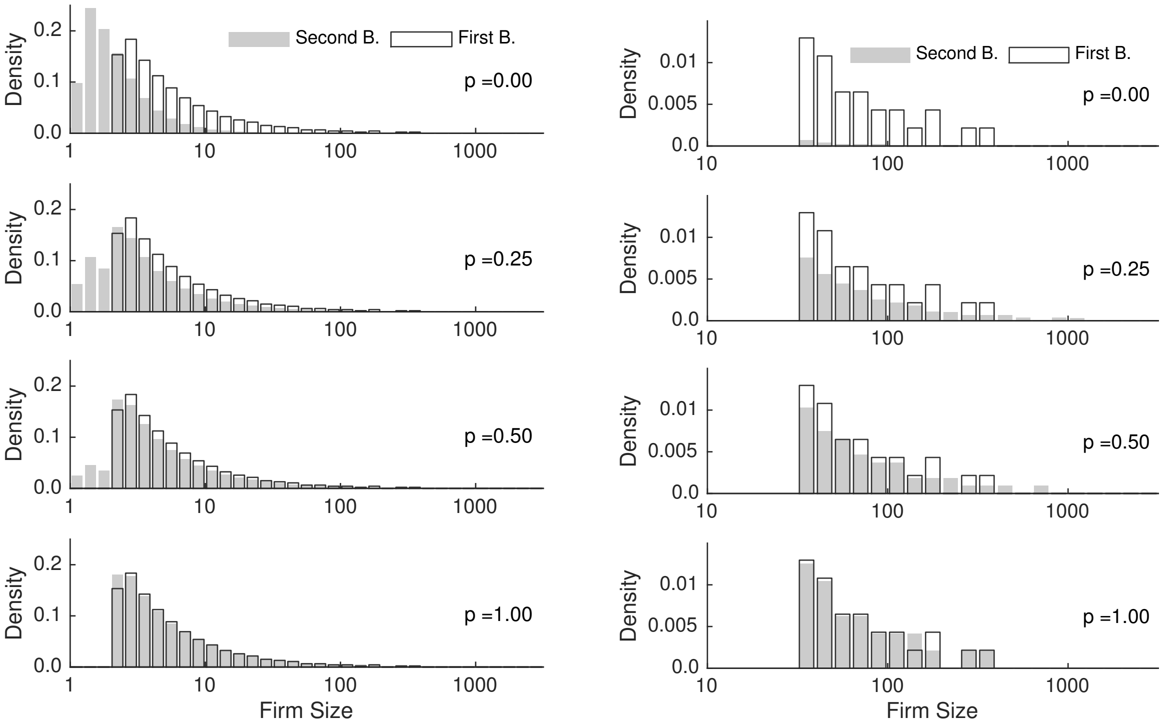
The Distribution of Firm Size In Figure 8, we look at the distribution of firm size. In each case, we look at variation in market access for four values varying from autarky to full access. Throughout Figure 6, the solid bars gives the distribution in the first best and the colors (green and orange) show deviations in second best from the first best. The left hand panel gives the full range of firms in the economy whereas the right hand panel gives the distribution of largest firms – the upper tail of the distribution. In each panel, we give the first best distribution of firms so that we can compare this to the distribution implied by the second-best.

In the top panel left panel we compare autarky to the first best. Here, we find a very clear shift in the distribution towards small firms. This made even more apparent in the top right-hand panel which shows that there are virtually no large firms in the economy. This lack of labour demand is what keeps the wage low. This broad pattern is found in all of the panels. However, as credit market access varies, the deviation from the "first best" distribution of firm size diminishes. By the time of full credit market access, the first best is very similar to distribution generated by second-best credit markets.

The gains from credit market access Figure 9 gives an insight into how the gains from participation in credit markets is distributed across the population. It contains two panels based on the level of competition. Comparing the two panels, it is clear that the gains and losses of credit markets, relative to autarky, are highly heterogeneous.

Those who are wage labourers in autarky are all better off with the possibility of trading in credit markets. However, among those who were entrepreneurs in autarky, a large fraction also lose from the introduction of credit markets due to rising wages. These losses are concentrated amongst entrepreneurs with higher level of wealth, holding θ constant. These entrepreneurs had good access to capital even in autarky, and benefited from low wage levels. This effect is particularly pronounced in the right hand panel of Figure 10, where we compare payoffs with credit markets (with $\phi = 1$) to autarky. Here, high productivity entrepreneurs with low

FIGURE 8: DISTRIBUTION OF FIRM SIZES

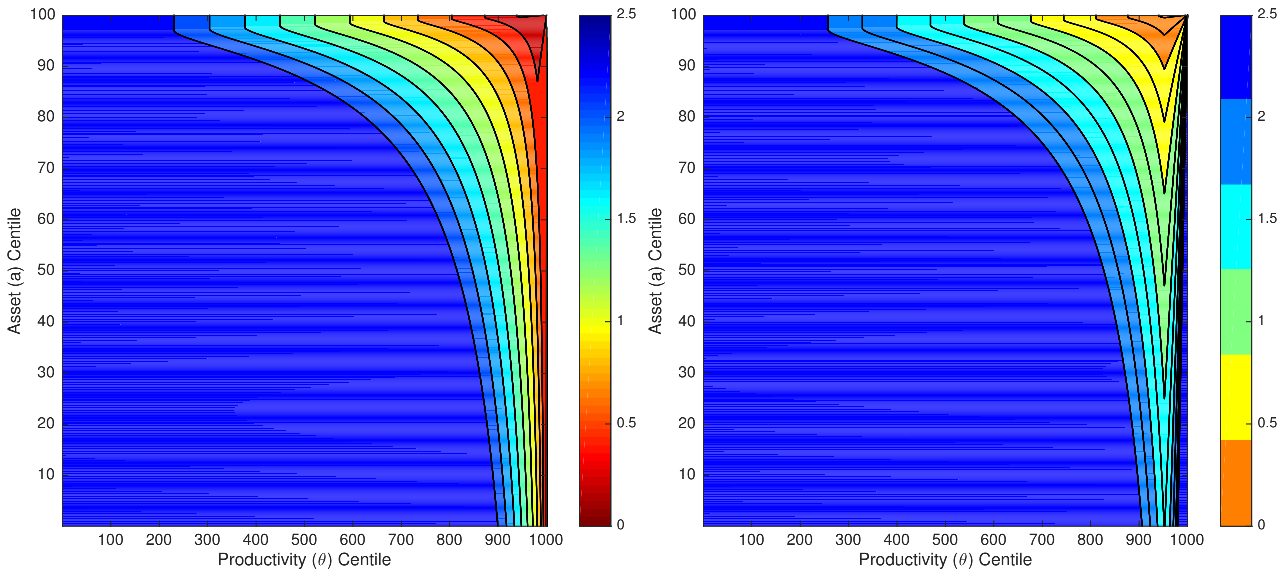


(A) FULL DISTRIBUTION

(B) UPPER TAIL OF DISTRIBUTION

Notes: All graphs depict the distribution of firm sizes in labour units in equilibrium (grey filled bars) We show the distribution of firm sizes for four distinct levels of credit market access: autarky ($p = 0.0$, top figure), low credit market access ($p = 0.25$, second from top figure), half the population has credit market access ($p = 5.0$, second from top figure), and full access ($p = 1.0$, top figure). Throughout we impose perfect collateralisability of wealth ($\tau = 1$), perfect credit market competition ($\phi = 1.0$) assume that the distribution of asset holdings and productivities are independent ($\rho = 0$). Subfigure (A) presents the full distribution of firm sizes with $\log_{10}$ scale on the x-axis; subfigure (B) presents a zoomed-in version of the right tail of the firm size distribution. In all figures the distribution of firm sizes in first best is also shown as benchmark (black outlined bars).

FIGURE 9: RELATIVE INCOME GAINS: CREDIT MARKETS VS. AUTARKY



(A) NO COMPETITION, $\phi = 0$

(B) FULL COMPETITION, $\phi = 1$ (WINZORISED)

Notes: We calculate for each level of assets a and productivity θ the ratio of income when credit markets exists over income in autarky. Both figures present contour maps of these ratios over the (a, θ) space. Figure (A) assumes fully uncompetitive credit markets ($\phi = 0$) and Figure (B) assumes fully competitive credit markets ($\phi = 1$). Note that in Figure (B) the income ratio is winzorised at 2.5 for visual clarity. This affects the top 2% of the θ distribution, and increasingly at lower levels of assets. The highest relative income gain with a ratio of 28.04 is observed for the highest level of θ and lowest level of a. Throughout we assume no correlation between assets and productivity ($\rho = 0$) and perfect collateralisability of wealth ($\tau = 1$).

levels of assets benefit form the increased access to credit, despite the fact that this comes with sizeable wage increases. However, entrepreneurs with similarly high productivity and high levels of assets do still loose out relative to autarchy. For them the increased credit access is less important, since they can to a large extent, just self-finance their investment. In the absence of competition, all of the most productive entrepreneurs lose relative to autarchy, since the increase in wages is not compensated by better credit access.

5 Concluding Comments

This paper has provided a quantitative exploration of the aggregate implications of credit market frictions where agents, who differ in wealth and productivity, make an occupational choice between being a worker and an entrepreneur. The paper has explored optimal credit contracts in a general equilibrium setting. It has then explored the impact of extending access to finance to a wider and wider group of the population.

The main aim of the paper has been to explore how lack of market access and credit market frictions affect the level of income per capita and wages for a fixed technology. Having specified a model, a key finding is that the wage moves from around 40% of the US wage in autarky to over 90% when there is full credit market access. Moreover, it is access to credit markets rather than frictions due to moral hazard which are quantitatively most important.

The paper focuses on an important general equilibrium channel by which credit market access affects development due to better selection of entrepreneurs, based on their talent and an increase in labour demand which increases the wage. Thus, credit market access spillovers benefit the population at large. Hence, looking for the impact of credit market distortions purely by focusing on capital markets and local effects where factor prices are fixed will often miss the big picture that expansion of the market to the unbanked is where the big gains in poverty reduction lie given that most of the poor in developing countries are dependent on wage labour.

The paper also highlights the importance of diminishing the extent of small-scale self-employment as development progresses with such individuals mainly switching to wage labor. This is consistent with anecdotal observation and evidence on the structural transformation which takes place alongside extending credit market access. Hence while entrepreneurship is important to development, it is the capacity of the market system to get capital to those entrepreneurs with high productivity which matters most. When market access is limited, many of them have to rely on their own wealth and resources which implies that they cannot operate at optimum scale.

The model has assumed a specific technology for providing credit where limits on conventional collateral create a friction, conditional on access. In future, we plan to explore further the gains to be made by expanding collateral to non-pecuniary punishments, so-called social collateral. This would allow us to link the paper to discussions about the role of micro-credit in development.

The paper has focused on differences in the level of income. A natural next step in the research program is to look at dynamic implications. Once there are new technologies and shocks to individual firm productivity due to this then the economy has to continuously reallocate capital. Limited market access plays a role in limiting this too if those who currently enjoy market access can get credit and those with new technologies are excluded. This will affect the ability of the economy to benefit from growth opportunities. More generally, it would be interesting to develop a dynamic version of the model with wealth accumulation.

References

[1] Axtell, R. L. [2001] "Zipf Distribution of U.S. Firm Sizes" Science, 293(5536), 1818–1820.

[2] Banerjee, Abhijit, and Andrew F. Newman, [1993] "Occupational Choice and the Process of Development." Journal of Political Economy, 101 (2): 274–98.

[3] Banerjee, Abhijit V. and Esther Duflo [2005]: "Growth Theory through the Lens of Development Economics", Chapter 7 in Philippe Aghion and Steven N. Durlauf (ed.s) *Handbook of Economic Growth*, Volume 1, Part A, 473-552.

[4] Banerjee, Abhijit, and Esther Duflo [2010], "Giving Credit Where it is Due," Journal of Economic Perspectives, 24(3), 61-79.

[5] Bardhan, Pranab, [1984], Land, Labor and Rural Poverty, Columbia University Press, New York and Oxford University Press, New Delhi.

[6] Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, [2013],"Cross-Country Differences in Productivity: The Role of Allocation and Selection." American Economic Review, 103 (1): 305–34.

[7] Besley, Timothy, Konrad Burchardi, and Maitreesh Ghatak [2012], "Incentives and the De Soto Effect,", The Quarterly Journal of Economics, 127(1), 237-282.

[8] Bloom, Nicholas, [2009], "The Impact of Uncertainty Shocks" Econometrica, 77(3), 623-685.

[9] Board of Governors of the Federal Reserve [2016] "Charge-Off and Delinquency Rates on Loans and Leases at Commercial Banks," [Online; accessed 18-November-2016]

[10] Buera, Francisco, Joseph Kaboski, and Yongseok Shin, [2011], "Finance and Development: A Tale of Two Sectors." American Economic Review, 101 (5): 1964–2002.

[11] Burgess, Robin and Rohini Pande, [2005], "Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment," American Economic Review, 95(3), 780-795.

[12] Caselli, Francesco, [2005], "Accounting for Cross-Country Income Differences," Chapter 9 in Philippe Aghion and Steven N. Durlauf (eds), Handbook of Economic Growth, Volume 1, Part A, 679-741.

[13] Cihak, Martin, Asli Demirguc -Kunt, Eric Feyen, and Ross Levine, [2013], "Financial Development in 205 Economies, 1960 to 2010," Working paper 18946, National Bureau of Economic Research.

[14] Credit Suisse [2015], "Global Wealth Report," Credit Suisse Research Institute.

[15] Demirguc-Kunt, Asli and Leora Klapper, [2013], "Measuring Financial Inclusion: Explaining Variation in Use of Financial Services across and within Countries," Brookings Papers on Economic Activity (Spring), 279-327.

[16] Dupas, Pascaline, Dean Karlan, Jonathan Robinson and Diego Ubfal, [2017], "Banking the Unbanked? Evidence from three countries," unpublished typescript.

[17] Galor, Oded, and Joseph Zeira, [1993] “Income Distribution and Macroeconomics.” Review of Economic Studies 60 (1): 35–52.

[18] Gerschenkron, Alexander [1962], Economic Backwardness in Historical Perspective – A Book of Essays, Harvard University Press, Cambridge, MA.

[19] Goldsmith, Raymond W. [1969], Financial Structure and Development. Yale University Press, New Haven, CT.

[20] Jeong, Hyeok, and Robert M. Townsend, [2007] “Sources of TFP Growth: Occupational Choice and Financial Deepening.” Economic Theory 32 (1): 179–221.

[21] Hsieh, Chang-Tai, and Peter J. Klenow, [2009], “Misallocation and Manufacturing TFP in China and India.” Quarterly Journal of Economics 124 (4): 1403–48.

[22] Hsieh, Chang-Tai, and Peter J. Klenow, [2010], “Development Accounting,” American Economic Journal: Macroeconomics, 2(1), 207-23.

[23] Karlan, Dean and Jonathan Zinman, [2009], “Observing Unobservables: Identifying Information Asymmetries With a Consumer Credit Field Experiment,” Econometrica, 77(6), 1993-2008.

[24] Levine, Ross, [2005] “Finance and Growth: Theory and Evidence.” In Handbook of Economic Growth. Vol. 1, Part A, edited by Philippe Aghion and Steven Durlauf, 865–934. Amsterdam: Elsevier B.V.

[25] Lloyd-Ellis, Huw, and Dan Bernhardt.[2000]: “Enterprise, Inequality and Economic Development.” *The Review of Economic Studies*, 67(1), 147–168.

[26] Matsuyama, Kiminori, [2007], “Aggregate Implications of Credit Market Imperfections.” In NBER Macroeconomics Annual 2007, Vol. 22, edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, 1–60.

[27] Moll, Benjamin [2014], “Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?,” American Economic Review 104 (10), 3186-3221.

[28] Nelsen, Roger B. [1999], An Introduction to Copulas, Springer.

[29] Restuccia, Diego, and Richard Rogerson, [2008] “Policy Distortions and Aggregate Productivity with Heterogeneous Plants.” Review of Economic Dynamics 11 (4): 707–20.

[30] Sklar, Abe, [1959], “Fonctions de répartition à n dimensions et leurs marges”, Publications de l’Institut de Statistique de l’Université de Paris, 8, 229–231.

[31] Stiglitz, Joseph, [1988], “Economic organization, information, and development,” Chapter 5 in Hollis B. Chenery and T.N. Srinivasan (eds), Handbook of Development Economics Volume 1, 93-160.

[32] Townsend, Robert, [1978], “Intermediation with Costly Bilateral Exchange,” Review of Economic Studies, 45(3), 417-25.

[33] Townsend, Robert and Kenichi Ueda, [2006], "Financial Deepening, Inequality, and Growth: A Model-Based Quantitative Evaluation," Review of Economic Studies, 73(1), 251-93.

[34] Trivedi, Pravin K. and David M. Zimmer [2007], "Copula Modeling: An Introduction for Practitioners," Foundations and Trends in Econometrics, 1(1), 1-111.

[35] World Bank, [2014], Global Financial Development Report 2014: Financial Inclusion, Washington DC, World Bank.

[36] Yi, K. , and J. Zhang, [2016], "Real Interest Rates Over the Long Run," Economic Policy Paper, 16(10), Federal Reserve Bank of Minneapolis.

A Proofs of Propositions

Proof of Proposition 1: The proof is a generalisation of the proof of Proposition 2 in Besley et al (2012) in two ways: first, $\pi(k;\theta,\mathbf{p})$ and $d(e;\theta,\mathbf{p})$ are allowed to depend on θ and $\mathbf{p}$; second, $d(e;\theta,\mathbf{p})$ is not necessarily linear in e. The first generalisation is trivial, since lender and borrower take θ and $\mathbf{p}$ as given in a given contracting problem. These are therefore additional constant, multiplicative factors, immaterial to the proof. The second generalisation is straightforward and affects the proof in two ways.

First, when the participation constraint is non-binding (Step 3 of the Proof of Proposition 2 in Besley et al (2012)), the optimal contracting problem can be written as

$$\max_{(e,k)} g(e)\left[\pi(k;\theta,\mathbf{p}) - \frac{d_e(e;\theta,\mathbf{p})}{g_e(e)}\right] + \tau a - \gamma k.$$

We have that $q(e)\pi(k;\theta,\mathbf{p})$ is strictly concave by Assumption 2 (ii) and $-q(e)\left[\frac{d_e(e;\theta,\mathbf{p})}{g_e(e)}\right]$ is strictly concave by Assumption 1 (i), (iii), and (iv). Therefore the maximisation is well-behaved, and by standard arguments, a unique global maximum (e_0, k_0) exists. The first-order necessary conditions for an interior optimum are:

$$g_e(e_0)\pi(k_0;\theta,\mathbf{p}) = d_e(e_0;\theta,\mathbf{p}) + g(e_0)\epsilon(e_0;\theta,\mathbf{p}) \tag{17}$$

$$g(e_0)\pi_k(k_0;\theta,\mathbf{p}) = \gamma. \tag{18}$$

We have that $\epsilon(e;\theta,\mathbf{p}) = \frac{d_{ee}(e;\theta,\mathbf{p})}{g_e(e)} - \frac{d_e(e;\theta,\mathbf{p})g_{ee}(e)}{[g_e(e)]^2} > 0$ since $g(e)$ is strictly concave, $g(e)$ and $d(e;\theta,\mathbf{p})$ are strictly increasing and $d(e;\theta,\mathbf{p})$ is weakly convex. By Assumption 2 (iii) the unique global maximum (e_0, k_0) is then an interior solution.

Secondly, if the participation constraint is binding (Step 4 of the Proof of Proposition 2 in Besley et al (2012)), we can use the binding participation constraint, incentive compatibility constraint and limited liability constraint to find optimal effort ξ defined by

$$g(\xi(v;\theta,\mathbf{p}))\left[\frac{d_e(\xi(v;\theta,\mathbf{p});\theta,\mathbf{p})}{g_e(\xi(v;\theta,\mathbf{p}))}\right] - d(\xi(v;\theta,\mathbf{p});\theta,\mathbf{p}) = v. \tag{19}$$

We have $\xi_v(v;\theta,\mathbf{p}) = g_e(\xi)\left[\frac{d_e(\xi;\theta,\mathbf{p})}{g_e(\xi)}\right] + g(\xi)\left[\frac{d_{ee}(\xi;\theta,\mathbf{p})}{g_e(\xi)} - \frac{d_e(\xi;\theta,\mathbf{p})g_{ee}(\xi)}{[g_e(\xi)]^2}\right] > 0$ by Assumption 1 (iv), and the fact that $g(e)$ and $d(e;\theta,\mathbf{p})$ are increasing.

Proof of Proposition 2: When the participation constraint is non-binding, capital is defined by (18). When the participation constraint is binding, effort is pinned down by (19). Using the incentive compatibility constraint and the lender's objective function, the optimal contracting problem becomes

$$\max_{k} g(\xi(v;\theta,\mathbf{p}))\left[\pi(k;\theta,\mathbf{p}) - \frac{d_e(\xi(v;\theta,\mathbf{p});\theta,\mathbf{p})}{g_e(\xi(v;\theta,\mathbf{p}))}\right] + \tau a - \gamma k.$$

The first order condition with respect to k again takes the form:

$$g(\xi(v;\theta,\mathbf{p}))\pi_k(k;\theta,\mathbf{p}) = \gamma.$$

Proof of Corollary 1: The proof is directly analogous to the proof of Lemma 1 in Besley et al (2012).

Proof of Proposition 3: We use the fact that

$$\phi \cdot S(\hat{u}(\phi; a, \theta, \mathbf{p}) + \tau a; \theta, \mathbf{p}) = \hat{u}(\phi; a, \theta, \mathbf{p}).$$

Now differentiate $\hat{u}(\phi; a, \theta, \mathbf{p})$ with respect to the various parameters to yield

$$\begin{aligned}
\frac{\partial \hat{u}(\phi; a, \theta, \mathbf{p})}{\partial \theta} &= \frac{S_\theta(\hat{u}(\phi; a, \theta, \mathbf{p}) + \tau a; \theta, \mathbf{p})}{1 - \phi S_v(\hat{u}(\phi; a, \theta, \mathbf{p}) + \tau a; \theta, \mathbf{p})} \\
\frac{\partial \hat{u}(\phi; a, \theta, \mathbf{p})}{\partial a} &= \frac{S_v(\hat{u}(\phi; a, \theta, \mathbf{p}) + \tau a; \theta, \mathbf{p})\tau}{1 - \phi S_v(\hat{u}(\phi; a, \theta, \mathbf{p}) + \tau a; \theta, \mathbf{p})} > 0 \\
\frac{\partial \hat{u}(\phi; a, \theta, \mathbf{p})}{\partial \phi} &= \frac{S(\hat{u}(\phi; a, \theta, \mathbf{p}) + \tau a; \theta, \mathbf{p})}{1 - \phi S_v(\hat{u}(\phi; a, \theta, \mathbf{p}) + \tau a; \theta, \mathbf{p})} > 0.
\end{aligned}$$

The inequalities follow noting that $1 - \phi S_v(\hat{u}(\phi; a, \theta, \mathbf{p}) + \tau a; \theta, \mathbf{p}) > 0$ for all $v \in (\underline{v}(\theta, \mathbf{p}), \bar{v}(\theta, \mathbf{p}))$ and $0 \leq S_v(\hat{u}(\phi; a, \theta, \mathbf{p}) + \tau a; \theta, \mathbf{p}) \leq 1$. Further we can derive

$$S_\theta = [g_e(\hat{e})\pi(\hat{k}; \theta, \mathbf{p}) - d_e(\hat{e}; \theta, \mathbf{p})]\hat{e}_\theta + g(\hat{e})\pi_\theta(\hat{k}; \theta, \mathbf{p}) - d_\theta(\hat{e}; \theta, \mathbf{p}),$$

where we use the first order condition with respect to k. We have $g_e(\hat{e})\pi(\hat{k}; \theta, \mathbf{p}) - d_e(\hat{e}; \theta, \mathbf{p}) > 0$ as long as the first best effort is not implemented; $g(\hat{e})\pi_\theta(\hat{k}; \theta, \mathbf{p}) > 0$ by Assumption 1 (i) and (ii); $-d_\theta(\hat{e}; \theta, \mathbf{p}) < 0$ by Assumption 1 (iii). From (19) it is straight-forward to show that the sign of $\hat{e}_\theta$ is indeterminate and depends on $d_{e\theta}(e; \theta, \mathbf{p})$, $d_\theta(e; \theta, \mathbf{p})$, and the shape of $g(e)$. Examples can be constructed where $S_\theta > 0$ and $S_\theta < 0$.