Impact of Entry Costs on Aggregate Productivity: Financial Development Matters

Saeed Shaker

Abstract

This paper revisits the question: what is the impact of entry costs on cross country differences in output and total factor productivity (TFP)? I argue that for the countries with low levels of financial development the answer is the conventional one in the literature, that higher entry costs cause misallocation of productive factors and lower TFP. But for the countries with reasonably high levels of financial development the conventional answer does not hold. Motivated by observations on cross-country data, I propose a new theory on the impact of entry costs on TFP. In my mechanism, there are two competing forces that affect TFP when entry cost changes: A wealth-based selection force, and a productivity-based selection force. This results in TFP being a hump-shaped function of entry costs. That is, entry costs are not inherently bad for TFP if their target is to deter low productivity individuals from starting business. I develop an analytically tractable model of firm dynamics with entry barriers and financial frictions and derive the sufficient conditions for the impact of entry cost on TFP in both wealth- and productivity-based selection phases.

JEL: E440, O16, O41,O43, L510

1 Introduction

There has been lots of attention and efforts in macroeconomics literature to explain the cross country differences in output and productivity. Startup entry costs, which are categorized as a form of institutional obstacles for businesses, is one of the important factors that has been taken into account. The main channel introduced in the literature for the interaction between entry costs and TFP is the misallocation channel. That is, higher entry costs would lower the number of entrants where low productivity firms may stay in business and get larger because of lack of competition on both output and input markets which would lower the TFP. This mechanism is backed by some empirical evidence using cross country data. The main starting point is that there is a negative correlation of roughly around -0.5 between entry costs and TFP. Some empirical works have established this inverse relationship between entry cost and TFP using different data sets and econometric methods.

I argue that the mechanism explained above may not hold depending on the current state of the economy. For a more appropriate analysis we need to take into account the channels that relate the entry costs and TFP to each other. Financial development and business density are two main potential links between entry costs and TFP that I considered in this paper.[1] In other words, the current state of the economy in terms of its level of financial development and business density is a main determinant of how the TFP responds to changes in entry costs. There will be more emphasis and focus throughout the paper on the role of financial development, but there is an important role for business density which is more implicit and I will discuss this in later sections of the paper. I split my sample of countries into two groups based on their level of financial development, and I repeat the simple correlation exercise for these subsets of countries.[2] I observe that for the countries with low levels of financial development, the correlation is still around -0.50, but for those with high levels of financial development the correlation becomes positive though very small, and as we narrow the set of countries to include those with even higher levels of financial development, the correlation becomes even larger.[3] This simple exercise tells us that for the countries with

[1]Other potential candidates could be: 1. average firm size, which is highly related to business density, 2. entry rate, 3. prices including wages and interest rates, 4. distributional links such as wealth and income distributions, all of which are considered in the dynamics of my analysis. There are also factors such as sectoral structure of the economy, and adjustment costs of labor and capital, etc. which will not be feasible to consider because of either data or modeling limitations.

[2]In this particular exercise I split the countries based on the ratio of external finance to GDP being lower than vs. greater than 1. However the results are robust to the choice of other indicators such as financial markets' depth, efficiency, etc.

[3]I have used different samples and variable measures for this exercise and obtained qualitatively the same results for all cases: one using the entry cost data developed by Djankov et al (2002) matched with the same period's TFP series developed similar to Klenow and Rodriguez-Clare (2005) (as well as the TFP measure from Penn World Tables

reasonably high levels of financial development, there might be some issues with the causal argument made in the literature. This could be an issue with internal validity of the mechanism explained in the literature since the relationship does not apply to some subsets of the sample of countries. My proposed theory is consistent with different levels of financial development, and can explain the conflicting behavior we observe towards either extreme.

For now let's focus on different levels of financial development for a given level of business density. In the economy, firms and potential entrants can externally finance their entry costs and capital requirements, but they need to provide some collateral in order to access financing. The ratio of the collateral to the amount of the loan depends on how well the financial market is functioning. To explain my mechanism in a simple way, again I divide the countries to two categories based on their level of financial development.

Low financial development: Let's assume there is very little financing available to the agents in the economy. Higher entry costs combined with poor credit markets mean that, some highly productive but poor individuals will not be able to pay the costs in order to start their businesses. This is not the case for the wealthy individuals as they can start business if they want to, somewhat regardless of their productivity. As a result, wealthy but less productive entrepreneurs will face less competition from the highly productive ones since a larger portion of them will stay out of business due to high entry costs. This would result in the market becoming populated with wealthy producers many of whom with low productivity. This is what I call the *wealth-based selection* force, which translates to a lower level of aggregate productivity.

High financial development: This case is a little more complicated, so I split it into two phases. Phase 1: let's assume there is no entry cost initially and we start increasing it by some small amount. In this phase, any highly productive individual will manage to start her business thanks to the developed financial markets. A low productivity individual, on the other hand, will find it less interesting to enter production because of the increased entry cost. To put it differently, the individuals that are confident about the profitability

9.1). In another exercise I have used the same measures of TFP, and the startup cost as percent of GNI per capita from World Bank's doing business survey from 2005 to 2013. Also for the measure of financial development I have used both the ratio of external finance to GDP used by Buera et al (2011), as well as the financial development index developed by IMF.

of their ideas will finance their entry cost and capital to start their businesses, but those with poor ideas will not. This occupational decision is to a great extent related to individuals' productivity and has very little to do with their wealth. That is why I call this a *productivity-based selection* phase. This would raise the productivity bar for active firms or entrepreneurs and increase the TFP. Phase 2: After we keep increasing entry cost beyond some threshold, more and more of the poor but highly productive entrepreneurs will be left out because even a highly (but not 100%) developed financial market will not allow them to borrow very large amounts. This would decrease the business density and competition, and would make it appealing for wealthy but less productive individuals to enter and stay in production. This is the *wealth-based selection* phase which will reduce the TFP. The conclusion is that, for high levels of financial development, entry costs act like a productivity filter when they are low and act as a wealth filter when they are high.

The main point of the discussion above is that, given any initial state of the economy, when entry cost increases, two forces compete in opposite directions to influence TFP. What I argue is that for low levels of entry cost it is more likely that the *productivity-based selection* will dominate, and for higher levels of entry cost it is more likely that the *wealth-based selection* will dominate. The outcome of these competing forces will be a TFP that is a hump-shaped function of entry cost. It follows that, there is a level of entry cost (different from zero depending on the state of the economy) for which the TFP is maximized. In other words, as far as the aggregate productivity is concerned, increasing entry cost might be helpful, or similarly lowering the entry cost might hurt depending on the current state of the economy. This is in contrast with the developed theories and conventional intuitions about the function of entry costs, all of which consider any level of entry costs harmful for TFP.

I will go through my proposed mechanism using an analytical model of entrepreneurship, and will derive the corresponding conditions for productivity- and wealth-based selection phases. I will show using the sufficient conditions that, for any given level of financial development, there exists a non-zero level of entry cost for which the aggregate productivity is maximized. I will also show that the TFP maximizing level of entry cost increases as we improve financial markets.

This paper is organized as follows: After a review of related literature in the following, section 2 discusses some empirical considerations and reports the results of some exercises carried out using cross country data. Section 3 introduces the analytical model in continuous time. Section 4 provides analysis of the impact of entry cost on TFP as well as a brief policy discussion followed by quantitative model in section 5. Finally section 6 concludes.

Related Literature

This paper brings the literature studying the impact of entry costs on cross country income, output and TFP differences together with a broad literature that studies the importance of financial markets on economic development. there is a large theoretical and quantitative literature on the latter. Comprehensive surveys on earlier contributions are conducted by Levine (2005), Matsuyama et al (2007) and Townsend (2010). My paper is related to a series of recent works investigating the links between financial markets and TFP and implications of different policies: Cagetti and DeNardi (2006), Amaral and Quintin (2010), Buera et al (2011), Buera and Shin (2013), Midrigan and Xu (2014), Moll (2014), Buera and Nicolini (2017), Itskhoki and Moll (2019). None of these papers focus on entry costs. My paper is close to Moll (2014) in a theoretical sense, with the difference of having entry costs as well as the fact that workers in my economy can save and become entrepreneurs. This makes the dynamics of my model richer because it includes occupational choices which can affect agents' savings decisions. With regards to modeling, with the exception of entry costs, my paper is also closely related to Buera and Shin (2013). The difference is that they do a purely quantitative analysis using decreasing returns to scale technology which I abstract from[4] in order to gain advantage in tractability of my model. I have extended my constant returns to scale production function to a decreasing returns to scale one at the expense of analytical solution and used a pure numerical solution. See section 5.

Major theory of entry costs goes back to the extensive work by Hopenhayn (1992). In his model, increasing entry cost would result in a decrease in the number of entrants which would reduce the exit threshold for the incumbents. A lower exit threshold means a lower average productivity for active firms which would reduce the TFP and output.[5] Financial market is the

[4] I use constant returns to scale in my analytical model similar to Moll (2014)

[5] See Shaker-Akhtekhane (2017) for a version of Hopenhayn (1992) in continuous time which produces the same results both quantitatively and qualitatively.

main missing piece that makes my results different than Hopenhayn's. there is also a large empirical literature on the relevance of entry costs in explaining output and TFP differences across countries: Djankov et al. (2002), Nicoletti and Scarpetta (2003, 2006), Barseghyan (2008), Poschke (2010) Barseghyan and DeCecio (2011), Moscoso-Boedo and Mukoyama (2012). All of these papers conclude a negative relationship between entry costs (or entry regulations in general) and aggregate productivity.[6] Again, these papers are missing the financial development component which I argue is an important determinant of how entry costs impact TFP.

The only paper, to the best of my knowledge, that has both entry costs and financial frictions is Bah and Fang (2016). Their model is isomorphic to the one developed by Hopenhayn (1992), which would not produce different dynamics for entry costs despite the inclusion of financial development factor. That means the channel through which entry cost impacts TFP is unchanged and the financial frictions only act as a propagation mechanism. My model is similar to Buera and Shin (2013) where the occupational choice is made after agents observe their productivity,[7] in contrast with Hopenhayn (1992) and most of the papers listed above where agents enter production and then find out about their productivity.[8]

Policy-wise, my framework suggests that, for financially underdeveloped economies it is better to lower entry costs to encourage participation of especially highly productive individuals, whereas for financially developed economies it is better to retain some, probably low, level of entry cost to impose selection. This mechanism varies with respect to the level of the development of the economy, and in some sense acts similar to the investment vs. selection phases introduced by Acemoglu et al. (2006) or pro-business vs. pro-worker policies introduced by Itskhoki and Moll (2019).

[6] Djankov et al. (2002) do not discuss the productivity. They find a negative relation between GDP per capita and entry costs.

[7] There are other papers using a similar approach. See for instance Cagetti and DeNardi (2006), Buera et al. (2011) and Midrigan and Xu (2014).

[8] One could argue that neither of these limits could be true in real world examples. That is, the potential entrants may know about their productivity up to some degree, i.e. there is a chance that they could be wrong. This intermediate case can be accounted for by the persistence of the productivity shocks. That is a potential entrant knows her current level of productivity, but the future could be very uncertain given how persistent or transitory the idiosyncratic shocks are.

2 Empirical Considerations

I have done a few simple exercises using cross-country data to make my argument and motivate my research. The main point I want to clarify is that higher entry costs are not necessarily associated with lower TFP in the data. In the first exercise, I used cross country data on startup costs for 1999 constructed by Djankov et al. (2002),[9] and combined it with Klenow and Rodriguez-Clare's (2005) productivity measure for the same year[10]. The correlation of entry cost and TFP is -0.49 which is substantial.[11] Then I include the same measure of external finance to GDP used by Buera et al. (2011) and split my sample to countries where this measure is greater than 1 and those with this measure being less than 1.[12] For the countries with low financial development, the correlation of entry cost and TFP is -0.40, which is still large. But for the countries with higher financial development the correlation is +0.01. This suggests that the negative causal relation between entry cost and TFP discussed in the literature may not hold for the countries with highly developed financial markets. Figure 2.1 clarifies this point.

To reinforce my point, I conduct another exercise with a relatively richer data set. For this exercise I combine TFP measure from Penn World Table[13] with the entry cost as percent of GNI per capita from World Bank's Doing Business Survey which is the extension of Djankov et al. (2002) measure. I also add the financial development index of International Monetary Fund which includes measures of depth, accessibility and efficiency of financial markets and takes values between 0 and 1. This data set spans from 2005 to 2013 and includes data for more than 80 countries at each year, total of 742 observations. Such a large number of observations gives me the flexibility of obtaining the correlation of entry costs and TFP given different thresholds

[9] This is the data set that were used by many researchers studying the entry regulations afterwards, and became the basis for the World Bank's *Doing Business* survey.

[10] I have also used TFP data from Penn World Table and obtained the same results.

[11] The correlation is -0.62 when variables are measured in logs. The results are qualitatively the same whether we use variables in logs or not.

[12] The mean value of the external finance to GDP in this sample is 0.89. Because my data set for this exercise becomes limited to only 52 countries after combining the mentioned variables, This choice of threshold at 1, leaves 20 observations on the higher financial development set.

[13] I have also repeated the same exercise using the TFP measure I created following Klenow and Rodriguez-Clare (2005) approach since their measure of TFP is different than Penn World Tables'. For my exercise both measures of TFP provide the same results.

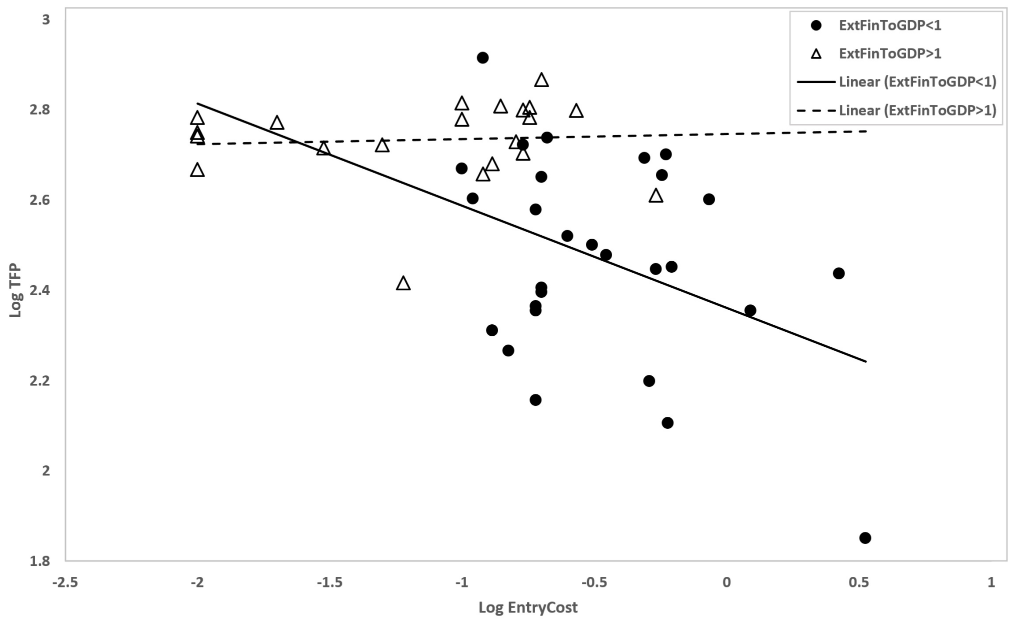


Figure 2.1: TFP vs Entry Cost for financially developed countries (triangles) and financially less developed ones (circles)

for financial development. The median, mean and 75th percentile of financial development index are 0.35, 0.40 and 0.60 respectively. So I will pick several thresholds for financial development index starting from 0.50 to 0.85.[14] Table 1 confirms that, as we limit our sample to countries with higher levels of financial development, the correlation of entry cost and TFP goes from negative to positive and increases as we continue. That means any causal analysis of entry costs on TFP should proceed with caution, and take financial development into account. I have repeated this exercise using different ranges for financial development and the results are qualitatively the same for all cases.

My third empirical exercise consists of a simple regression analysis. Using the same data set above, I simply regress log of TFP on log of entry cost including financial development index and its interaction term with entry cost. The results are shown in column (Spec1) of Table 2. As we can see from

[14]When I go beyond 0.85 the number of observations drop below 33 which would not give reliable estimates. However, the results show the same pattern even with very few observations. The results are also consistent when we choose different increments from 0.5 to 0.85.

Fin. dev. index >	Corr(entry cost, TFP)	Observations
> 0 (all sample)	-0.402	742
> 0.5	-0.385	259
> 0.6	-0.330	184
> 0.7	-0.107	136
> 0.8	+0.136	58
> 0.85	+0.338	33

Table 1: The correlation between startup costs and TFP given different levels of financial development.

this regression, the coefficient on log(entry cost) is negative, and the interaction term is positive. This means, for low levels of financial development, the coefficient of entry cost will be negative but for high levels of financial development, the sign of the coefficient will turn positive. More specifically, the coefficient of entry cost becomes positive when the financial development index becomes greater than 0.76 which is about the 90th percentile in the sample. The following equation makes it clear.

$$\begin{aligned} logTFP = & - 0.449 + 0.434 Fin.dev \\ & + (-0.161 + 0.211 Fin.dev) logEntryCost \end{aligned}$$

In a similar regression I add a new variable for business density defined as the total number of businesses divided by labor force. As we can see in column (Spec2) of Table 2, the addition of business density does not change the coefficients of the previous regression. The coefficient on the interaction term between business density and entry cost is also positive. This is consistent with the intuition that if the business density is already high, there is a higher chance that coefficient on entry cost becomes positive.

3 Model

3.1 Outline

The model is set up in continuous time. there is a continuum of individuals in the economy with measure normalized to 1. Agents can be either

Table 2

	Dependent variable:	
	log(TFP)	
	(Spec1)	(Spec2)
Constant	-0.449^{***}	-0.381^{**}
	(0.048)	(0.158)
Fin. dev.	0.434^{***}	0.437^{**}
	(0.076)	(0.187)
log(Entry Cost)	-0.161^{***}	-0.170^{***}
	(0.013)	(0.053)
Fin. dev $\times$ log(Entry Cost)	0.211^{***}	0.151^{*}
	(0.026)	(0.077)
Bus. dens		-0.005
		(0.697)
Bus. dens$\times$log(Entry Cost)		0.545^{*}
		(0.315)
Observations	742	102
R^2	0.590	0.579
Adjusted R^2	0.588	0.557
F Statistic	354^{***} (df = 3; 738)	26^{***} (df = 5; 96)
Note:		*p<0.1; **p<0.05; ***p<0.01

wage workers or entrepreneurs. At the beginning of every period, every individual makes an occupational choice with the knowledge of their productivity, z. They either choose to work for a wage where they inelastically supply their labor time and earn wage w, or run their own businesses as entrepreneurs and earn profits. Agents in the economy decide how much to save and consume, and the goal is to maximize their lifetime utility. Every period, the entrepreneurs hire labor and capital, and solve the following profit-maximization problem:

$$\begin{aligned} \Pi^E(a,z) &= \max_{k,l} \left\{ f(z,k,l) - wl - (r+\delta)k \right\}, \\ & s.t. \\ k &\leq \lambda a \end{aligned} \tag{3.1}$$

where all notation is as in standard models: a being the wealth of entrepreneur, w wage, r rental rate of capital, δ capital depreciation rate. Also λ is the level of financial development, ranging from 1 (no financial markets) to ∞ (perfect financial markets). The production function uses capital and labor as inputs and is assumed to be of the following constant returns to scale form:

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

Deviation from constant returns production is only possible at the expense of analytical solution, which will be considered in section 5. The constant returns to scale assumption implies that it would be optimal for entrepreneurs to produce using the maximum amount of the capital they can finance given the collateral constraint that depends on their wealth. This would give the following production decisions for an entrepreneur with wealth a and productivity z:

$$k(a,z) = \lambda a, \tag{3.2}$$

$$l(a,z) = \left(\frac{1-\alpha}{w}\right)^{\frac{1}{\alpha}} zk(a,z), \tag{3.3}$$

$$\Pi^E(a,z) = \eta(z)\lambda a, \tag{3.4}$$

where

$$\eta(z) = (z\pi - r - \delta), \quad \text{and} \tag{3.5}$$

$$\pi = \alpha \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}. \tag{3.6}$$

The evolution of wealth, or savings, for wage workers ($j = W$) and entrepreneurs ($j = E$) is given by:

$$\dot{a} = \Pi^j + ra - c,$$

where c is the consumption, $\Pi^W = w$, and Π^E is given by (3.1).

3.1.1 Entry to Entrepreneurship

Entry to entrepreneurship requires an upfront payment of entry cost, c_e.[15] An entrant can also borrow up to λa to finance the entry cost as well as the productive capital. That is,

$$k + c_e \leq \lambda a.$$

It follows that the new entrant's wealth evolves according to:

$$\dot{a} = \eta(z)(\lambda a - c_e) - (1 + r)c_e + ra - c.$$

This means at the very first period an entrant has access to all she can finance net of entry cost to use as productive capital, i.e. ($k = \lambda a - c_e$). After production she pays the loan back (that includes the entry cost) plus interest and decides how much to consume.

An alternative setup would be to assume as if there is a portion of wealth that entrants give up upon entry. Note that this entry cost equivalent bygone wealth will depend on entrant's productivity. Let's denote the entry cost equivalent part of wealth that entrant gives up by $\hat{c}_e(z)$. We have the following:

$$\dot{a} = \eta(z)\lambda(a - \hat{c}_e(z)) + r(a - \hat{c}_e(z)) - c.$$

Simple algebra gives the following expression for $\hat{c}_e(z)$:

$$\hat{c}_e(z) = \frac{\eta(z) + r + 1}{\eta(z)\lambda + r} c_e \tag{3.7}$$

As we can see in the formula above, other than technology parameters and prices, $\hat{c}_e(z)$ depends mainly on the financial market's parameter λ, and the productivity level of entrant z. As we will see later, this setup would be very helpful in the formulation of value function as a Hamilton-Jacobi-Bellman Quasi-Variational Inequality (HJBQVI).

[15] In this paper I do not include any exit cost or irreversibility cost in order to keep things simple and straightforward.

3.1.2 Standard Stopping Time Setup

Individuals, whether worker or entrepreneur, $j \in \{W, E\}$, maximize their lifetime utility which becomes a stopping time problem as follows:

$$V^j(a,z) = \max_{c_t,\tau} \left\{ E_0 \int_0^\tau e^{-\rho t} u(c_t) dt + e^{\rho \tau} V^{*-j} \right\} \tag{3.8}$$

Subject to

$$\begin{aligned} \dot{a}_t &= \Pi_t^j + r_t a_t - c_t \\ dz_t &= \mu(z_t)dt + \sigma^2(z_t)dW_t \end{aligned}$$

where $\Pi^W = w$, and Π^E is given by (3.1). Also τ is the occupation switch time; z can be some diffusion process with $z \in (z_{min}, z_{max})$;[16] and $-j$ means the occupation other than j. For now let's only mention that V^* is the value of switching occupation. Its arguments will be clarified in the next subsection. For tractability of the model I use a log utility function, $u(c) = \log(c)$.[17]

3.1.3 A Binary Choice Formulation of Stopping Time

Here I introduce an alternative formulation for the stopping time problems given by (3.8) from a different perspective. The new structure provides a unified value function for any agent type as a binary choice problem and does not involve stopping time (τ) which makes it easier to understand the formulation of value function. This setting may also be helpful in developing new ways of solving these stopping time problems merely as a binary choice optimization.[18]

Let's define the occupation state as a binary variable $o \in \{0, 1\}$, where 0 stands for wage worker, W, and 1 for entrepreneur, E. Also let's define the corresponding control for it as $x \in \{0, 1\}$, where 0 means continuing with

[16] Boundedness of the productivity shocks is not a strong assumption. See Moll (2014) for a discussion.

[17] The CRRA case is somewhat more involved compared to the log utility, and consumption and savings decision rules become a more complicated function of productivity shocks and wealth. I have used CRRA in the quantitative model and obtained consistent results.

[18] I will simply use the established quasi-variational-inequality formulation to solve the problem, because talking about the methods that deal with the binary choice equivalent is out of the scope of this paper.

the same occupation, and 1 means switching occupation. The equivalent problem can be written as the following:

$$V(\mathcal{O}, a, z) = \max_{c_t, x_t} \left\{ E_0 \int_0^\infty e^{-\rho t} u(c_t) dt \right\} \tag{3.9}$$

Subject to

$$\begin{aligned} \dot{a}_t &= \mathcal{O}_t \Pi_t(a_t, z_t) + (1 - \mathcal{O}_t) w_t + r_t a_t - c_t - (1 - \mathcal{O}_t) x_t \hat{c}_e(z) && (3.10) \\ d\mathcal{O}_t &= (1 - 2\mathcal{O}_t) x_t && (3.11) \\ dz_t &= \mu(z_t) dt + \sigma^2(z_t) dW_t && (3.12) \end{aligned}$$

In this setup the occupational choice is embedded in the problem as a control that looks like other choice variables in the model. In some cases the binary control can be treated as a continuous one if we assign a large penalty for deviation from binary choice. Note that the last term on the right hand side of (3.10) is the change in wealth associated with entry decision. $(1 - \mathcal{O}_t) x_t$ equals 1 only when the current occupation is W, ($\mathcal{O}_t = 0$), and a switch occurs at the same time, ($x_t = 1$).

3.2 HJBs and Decision Rules

The formulations of the problem provided above can be considered as a stochastic impulse control problem. The standard results of the impulse control theory indicate that the value functions defined in (3.8) and (3.9) can take the form of a quasi-variational inequality. See Bensoussan and Lions (1987) for a broad discussion on the subject. Also see Duckworth and Zervos (2001) for a discussion on a more related problem. We will have two recursive forms of value functions for entrepreneurs (denoted by E) and workers (denoted by W), known as Hamilton-Jacobi-Bellman quasi-variational inequalities (HJBQVI):

$$\begin{aligned} \text{E}: \quad \min\{\rho V^E(a, z) &- V_a^E(a, z) \dot{a}^E && (3.13) \\ &- V_z^E(a, z) \mu(z) - \frac{1}{2} V_{zz}^E(a, z) \sigma^2(z) \\ &- u^E(c) \ , \ V^E(a, z) - V^W(a, z)\} = 0. \end{aligned}$$

$$\begin{aligned}
\mathrm{W}: \quad \min\{\rho V^W(a,z) &- V_a^W(a,z)\dot{a}^W \\
&- V_z^W(a,z)\mu(z) - \frac{1}{2}V_{zz}^W(a,z)\sigma^2(z) \\
&- u^W(c), V^W(a,z) - V^E(a-\hat{c}_e(z),z)\} = 0,
\end{aligned} \tag{3.14}$$

Because of the entry costs, I will distinguish between incumbents, who were entrepreneurs last period and continue as entrepreneurs, $(E_{t-dt} \rightarrow E_t)$, and entrants, who were working for wage last period and decided to start their business, $(W_{t-dt} \rightarrow E_t)$. Since the technology is constant returns to scale (CRS), we have the simple optimal decision rules and profits given by (3.2, 3.3, 3.4) for incumbent entrepreneurs. We also have the following savings decision for incumbents.[19]

$$\dot{a} = (\lambda\eta(z) + r - \rho)\, a \tag{3.15}$$

We have the following decision rules and profits for entrants

$$k(a,z) = \lambda a - c_e, \tag{3.16}$$

$$l(a,z) = \left(\frac{1-\alpha}{w}\right)^{\frac{1}{\alpha}} zk(a,z), \tag{3.17}$$

$$\Pi(a,z) = \eta(z)\lambda(a - \hat{c}_e(z)). \tag{3.18}$$

where $\eta(z)$ is given by (3.5). For (3.15), we assume that the shock process is defined in a way that if you make 0 profits as an entrepreneur, you will exit to earn wage as a worker. This assumption means every active entrepreneur that does not exit, will produce a positive amount and there is no temporary shut down. The following propositions provide a simple presentation of the continuation value for entrepreneurs and wage workers.

Proposition 1. *Continuation value for entrepreneurs can be reduced to the following ordinary differential equation (ODE) form:*

$$V^E(a,z) = v^E(z) + \frac{1}{\rho}log(a),$$

[19]See Moll (2014) for a similar analysis of incumbents' decisions.

where $v^E(z)$ *solves*

$$\rho v^E(z) = \log \rho + \frac{1}{\rho}(\lambda\eta(z) + r - \rho) + \frac{dv^E(z)}{dz}\mu(z) + \frac{1}{2}\frac{d^2v^E(z)}{dz^2}\sigma^2(z).$$

Proposition 2. *Continuation value for wage workers can also be reduced to the following ODE form:*

$$V^W(a, z) = v^W(z) + \frac{1}{\rho}log(w + ra),$$

where $v^W(z)$ *solves*

$$\rho v^W(z) = log\frac{\rho}{r} + \frac{1}{\rho}(r - \rho) + \frac{dv^W(z)}{dz}\mu(z) + \frac{1}{2}\frac{d^2v^W(z)}{dz^2}\sigma^2(z).$$

These propositions show that the continuation values which are of a more complicated partial differential equation (PDE) forms can be simplified to ODEs which can be analytically solvable in many cases. Also these propositions show the link between two value functions. We see that ODEs for $v^E(z)$ and $v^W(z)$ differ only in constant terms as well as a multiply of z present in the former. This means they have similar particular solutions, and as a result solving one would lead to a general solution of the other. I do not go further in terms of the solution forms since it requires me to specify exact forms of $\mu(z)$ and $\sigma^2(z)$ which I do not intend to do here. The following proposition specifies the boundaries for entry and exit decisions.

Proposition 3. *Given prices, for any level of* z *there is a wealth threshold below which entrepreneurs exit:*

$$\underline{a}^{Exit}(z) = \frac{w}{\lambda\eta(z)},$$

and there is a wealth threshold above which workers enter business:

$$\underline{a}^{Enter}(z) = \frac{r\hat{c}_e(z) + w}{\lambda\eta(z)} + \hat{c}_e(z).$$

Proposition 3 specifies the entry and exit cutoffs, which also implies that there is an area in the space of (a, z) where all agents are entrepreneurs, and another area where all agents are wage workers. The area in between them can contain a mixture of entrepreneurs and workers. The following corollary is an obvious result of proposition 3.

Corollary 1. *We have*

$$\frac{d\underline{a}^{Exit}(z)}{dz} < 0, \quad \textit{and} \quad \frac{d^2\underline{a}^{Exit}(z)}{dz^2} > 0,$$

and similarly

$$\frac{d\underline{a}^{Enter}(z)}{dz} < 0, \quad \textit{and} \quad \frac{d^2\underline{a}^{Enter}(z)}{dz^2} > 0.$$

Corollary 1 characterizes the functional form (behavior) of $\underline{a}^{Exit}(z)$ and $\underline{a}^{Enter}(z)$. Using the entry and exit cutoffs provided in proposition 3, I define the entry and exit zones. See figure 3.1 for clarification.

Definition 3.1. Using the cutoffs in proposition 3, let's define the entry, exit and inaction zones as follows:

$$\mathcal{R}^{Enter} = \{(a, z) : z \in \{\underline{z}, \bar{z}\}, a \geq \underline{a}^{Enter}(z)\}, \tag{3.19}$$
$$\mathcal{R}^{Exit} = \{(a, z) : z \in \{\underline{z}, \bar{z}\}, a < \underline{a}^{Exit}(z)\}, \tag{3.20}$$
$$\mathcal{R}^{Inaction} = \{(a, z) : z \in \{\underline{z}, \bar{z}\}, \underline{a}^{Exit}(z) \leq a < \underline{a}^{Enter}(z)\}. \tag{3.21}$$

Now we can get the Kolmogorov Forward Equations (KFEs) which specify the distribution of agents in the economy. I will have two KFEs corresponding to two types of agents in the economy, workers and entrepreneurs. For occupation $j \in \{W, E\}$ we have:

$$\begin{aligned} \frac{\partial g^j(a, z, t)}{\partial t} &= \frac{1}{2}\frac{\partial^2}{\partial z^2}\left(\sigma^2(z)g^j(a, z, t)\right) - \frac{\partial}{\partial z}\left(\mu(z)g^j(a, z, t)\right) \\ &- \frac{\partial}{\partial a}\left[\dot{a}^j g^j(a, z, t)\right] - m(j, -j) + m(-j, j) \end{aligned} \tag{3.22}$$

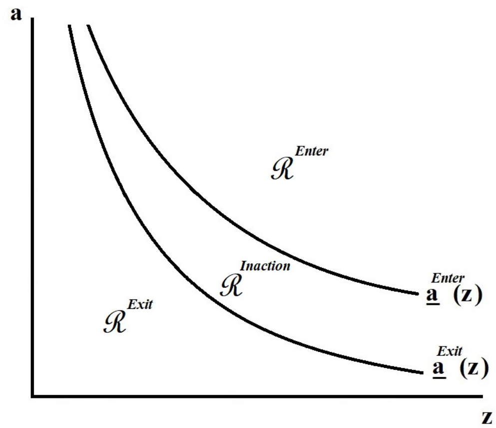


Figure 3.1: Entry and exit zones. $\mathcal{R}^{Exit}$ is the exit zone and only contains wage workers; $\mathcal{R}^{Enter}$ is the entry zone and only contains entrepreneurs; $\mathcal{R}^{Inaction}$ is the inaction zone and can contain both types.

where $m(j,-j)$ is the distribution of individuals switching away from j, and $m(-j,j)$ is the distribution of those switching into j.

$$\begin{aligned} m(W,E) &= g^W(\tilde{a},\tilde{z}), \text{ where } (\tilde{a},\tilde{z}) \in \mathcal{R}^{Enter}, & (3.23) \\ m(E,W) &= g^E(\tilde{a},\tilde{z}), \text{ where } (\tilde{a},\tilde{z}) \in \mathcal{R}^{Exit}. & (3.24) \end{aligned}$$

Also let's denote the aggregate distribution of all agents by:

$$g(a,z,t) = g^E(a,z,t) + g^W(a,z,t).$$

3.3 Market Clearing

In every period, the following market clearing conditions hold:
1. Financial market

$$\begin{aligned} \int a_t dG_t(a,z) &= \int k_t(a,z) dG_t^E(a,z) & (3.25) \\ &+ \int_{\mathcal{R}^{Enter}} c_e dG_{t-dt}^W(a,z), \end{aligned}$$

2. Labor market

$$\int l_t(a,z) dG_t^E(a,z) = \int dG_t^W(a,z) \qquad (3.26)$$

Now, before getting to aggregate analysis, I provide the following definitions for wealth shares that will be used throughout the paper.[20]

Definition 3.2. Let's define the wealth share held by individuals with productivity level z:

$$\varpi(z,t) = \frac{1}{A(t)} \int_0^\infty a g_t(a,z) da,$$

where $A(t)$ is the total wealth in the economy:

$$A(t) = \int a dG_t(a,z).$$

For entrepreneurs only, the wealth share is given by:

$$\omega(z,t) = \frac{1}{A(t)} \int_0^\infty a g_t^E(a,z) da. \tag{3.27}$$

Definition 3.3. In the same manner we can define two cumulative wealth distributions:

$$\Upsilon(z,t) = \int_{z_{min}}^{z} \varpi(x,t) dx$$

and

$$\Omega(z,t) = \int_{z_{min}}^{z} \omega(x,t) dx, \tag{3.28}$$

where the first one is the cumulative wealth share for everyone and the second one is for entrepreneurs only.

3.4 Aggregates

After occupational choices are made, new entrants pay the entry cost and entrepreneurs decide on the amount of capital and and labor to hire, then the financial and labor markets clear. Using the financial market clearing condition we have:

$$\begin{aligned} \int a dG_t(a,z) = \lambda \int a dG_t^E(a,z) \quad &- \quad \int_{\mathcal{R}^{Enter}} c_e dG_{t-dt}^W(a,z) \\ &+ \quad \int_{\mathcal{R}^{Enter}} c_e dG_{t-dt}^W(a,z), \end{aligned} \tag{3.29}$$

[20] I will use the wealth shares to specify aggregates. See Kiyotaki (1998) and Moll (2014), among others, that have used wealth shares to specify aggregates

This can be written as follows:

$$A(t) = \lambda\Omega(z_{max}, t)A(t), \tag{3.30}$$

or

$$\Omega(z_{max}, t) = \frac{1}{\lambda} \tag{3.31}$$

where $\Omega(z_{max}, t)$ is the wealth share held by all entrepreneurs. We have the following proposition on economic aggregates. Unless required, I drop the time index for notation simplicity.

Proposition 4. *Aggregate capital, output, and TFP are given by:*

$$\begin{aligned} K &= A - c_e\xi, & (3.32)\\ Y &= ZK^{\alpha}L^{1-\alpha}, \quad \textit{where} & (3.33)\\ Z &= \left(\frac{\lambda AM - c_e\hat{z}}{A - c_e\xi}\right)^{\alpha} & (3.34) \end{aligned}$$

where ξ is the number of entrants, or entry rate:

$$\xi = \int_{\mathcal{R}^{Enter}} dG^W_{t-dt}(a, z)$$

and

$$M = \int_{z_{min}}^{z_{max}} z\omega(z)dz \tag{3.35}$$

is the wealth-weighted productivity of all entrepreneurs, and

$$\hat{z} = \int_{\mathcal{R}^{Enter}} zdG^W_{t-dt}(a, z) \tag{3.36}$$

is the sum of productivity of entrants, and $\frac{\hat{z}}{\xi}$ is the average productivity of entrants.

Proof. Aggregate capital simply follows from capital market clearing condition, (3.25) and (3.30). First, let's obtain aggregate labor from the labor market clearing condition, (3.26), and optimal labor choices of entrepreneurs, (3.3, 3.17). We have the following:

$$L = \left(\frac{\pi}{\alpha}\right)^{\frac{1}{1-\alpha}} (\lambda AM - c_e\hat{z}) \tag{3.37}$$

where $\hat{z}$ and M are as defined in this proposition. To obtain Y integrate over entrepreneurs' production and use optimal choice of labor:

$$\begin{aligned} Y &= \int (zk)^{\alpha} l^{1-\alpha} dG^{E}(a,z) \\ &= \frac{\pi}{\alpha} \int zka dG^{E}, \end{aligned}$$

and using optimal capital choice for incumbents and entrants we have:

$$Y = \frac{\pi}{\alpha} \left(\lambda AM - c_e \hat{z} \right).$$

Now using (3.37) and (3.32) we get:

$$Y = \left(\frac{\lambda AM - c_e \hat{z}}{A - c_e \xi} \right)^{\alpha} K^{\alpha} L^{1-\alpha}.$$

which gives aggregate output and aggregate TFP as in (3.33) and (3.34). □

Corollary 2. *One result of proposition 4 is the following:*

$$Y = \left(\frac{w}{1-\alpha} \right) L.$$

That is, output per worker, Y/L only depends on equilibrium wage and factor share parameter.

4 Impact of Changing Entry Cost

4.1 TFP response to entry cost

We want to find the response of Z to changes in c_e. To simplify notation, for any variable x, let's denote $\frac{\partial x}{\partial c_e}$ by x'. We are interested in Z'. The following results and definitions will be useful in my analysis.

Definition 4.1. Let's denote by Υ the share of total wealth spent on entry cost:

$$\Upsilon = \frac{c_e \xi}{A}.$$

Lemma 1. *The following results hold:*

(1) $\xi' < 0$,

(2) $\lim_{c_e \to \infty} \xi = 0$,

(3) $(\underline{a}^{Enter}(z))' > 0$ *and* $(\underline{a}^{Exit}(z))' = 0$, $\forall z \in (z_{min}, z_{max})$,

(4) $\Upsilon' = \Upsilon \left(\frac{1}{c_e} + \frac{\xi'}{\xi} - \frac{A'}{A} \right)$,

(5) $\lim_{c_e \to 0} \Upsilon' = \frac{\xi}{A}$,

(6) $\lim_{c_e \to \infty} \Upsilon = 1$,

(7) $\lim_{c_e \to \infty} \Upsilon' = \frac{\xi'}{\xi}$.

Proof. (1) comes from workers' HJBVI, (3.14): with higher c_e higher shock is required to start a business. (2) follows from (1). (3) is a result of proposition 3: since the entry becomes difficult, the wealth requirement for running a business would rise at any productivity level. (4-7) simply follow from definition 4.1. □

Since I am mainly interested in the direction of the TFP response to changes in entry cost, I use the log(TFP) instead of TFP. This brings us to the following proposition.

Proposition 5. *The response of TFP to entry cost can be stated as:*

$$\mathcal{Z}' \equiv \frac{\partial \log(Z)}{\partial c_e} = \alpha \left(\frac{\Upsilon'}{1 - \Upsilon} + \frac{\lambda M' - \Upsilon' \frac{\hat{z}}{\xi} - \Upsilon \left(\frac{\hat{z}}{\xi} \right)'}{\lambda M - \Upsilon \frac{\hat{z}}{\xi}} \right) \tag{4.1}$$

The proof follows from the expression for Z. This proposition, as I will discuss in detail, indicates that M' plays a very crucial role in determining the sign of $\mathcal{Z}'$. Since M is defined as the wealth-weighted productivity of all entrepreneurs, M' would depict changes in productivity of entrepreneurs as well as in their wealth shares due to changes in entry cost. Therefore, it can be useful to take a more intuitive look into M and M'.

Lemma 2. *The wealth-weighted productivity of entrepreneurs, M, can be written as follows:*

$$M = \frac{z_{max}}{\lambda} - D$$

where

$$D = \int_{z_{min}}^{z_{max}} \Omega(z)dz.$$

Also

$$M' = -D'$$

Proof. This simply follows from integration of (3.35) by parts and using (3.31). □

In lemma 2, D is defined as the integration over the cumulative wealth shares of entrepreneurs for all productivity levels, that is the area under the cumulative wealth distribution curve. Since it is graphically more convenient to work with D, I will use it to provide more intuition for M. D' can be seen as the relative changes in D when c_e changes from some c_e^0 to $c_e^0 + \Delta$. The following lemma uses the notion of *second order stochastic dominance* (SOSD) to explain this.

Lemma 3. *Using the definition of D, we have the following equivalencies for the sign of D' when c_e goes from c_e^0 to $c_e^0 + \Delta$ for some arbitrarily small Δ:*

1. $D' < 0$ *iff*

$$\begin{cases} i.\ \Omega(z; c_e^0 + \Delta)\,(SOSD)\,\Omega(z; c_e^0) \\ or \\ ii.\ \frac{1}{\lambda} - \Omega(z_{max} + z_{min} - z; c_e^0)\,(SOSD)\,\frac{1}{\lambda} - \Omega(z_{max} + z_{min} - z; c_e^0 + \Delta) \end{cases}$$

2. $D' > 0$ *iff*

$$\begin{cases} i.\ \Omega(z; c_e^0)\,(SOSD)\,\Omega(z; c_e^0 + \Delta) \\ or \\ ii.\ \frac{1}{\lambda} - \Omega(z_{max} + z_{min} - z; c_e^0 + \Delta)\,(SOSD)\,\frac{1}{\lambda} - \Omega(z_{max} + z_{min} - z; c_e^0) \end{cases}$$

This lemma is a result of the definition of D as well as the concept of second order stochastic dominance.[21] Note for $D' < 0$ that we have either case (i) which means $\Omega(z; c_e^0 + \Delta)$ second order stochastically dominates $\Omega(z; c_e^0)$ or case (ii) which means that inverted mirror of $\Omega(z; c_e^0)$ second order stochastically dominates inverted mirror of $\Omega(z; c_e^0 + \Delta)$.[22]

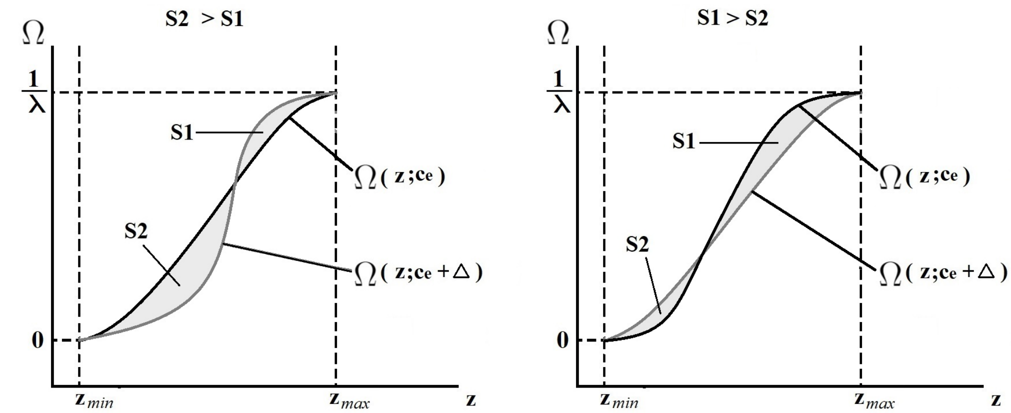
For $D' < 0$, Case (i): when entry cost increases, the wealth share of lower-productivity entrepreneurs shrinks and moves higher in the productivity ladder. Case (ii): when entry cost increases, the wealth share of mid level productivity entrepreneurs shrinks but shifts more upward than it does downward. In sum, $D' < 0$ means that wealth share gain (loss) by high productivity entrepreneurs is more (less) than that of the lower productivity entrepreneurs when we increase entry cost. The opposite can be said about $D' > 0$.

One important fact about the setting of lemma 3 is that cases (i) and (ii) correspond to both extensive and intensive margins of change in wealth shares. Let's assume entry cost increases. Again let's consider $D' < 0$. Both cases (i) and (ii) indicate that wealth shares move up along the productivity axis more than they move down. In other words, in general higher productivity entrepreneurs' wealth shares increase. One implication is that in the extensive margin there are more high productivity entrepreneurs in the economy, and as a result they will have greater portion of aggregate production than they had before. Another implication is in the intensive margin where the highly productive entrepreneurs become wealthier than they were before. Since the wealth of entrepreneurs is a determinant of their productive capital, the wealthy productive entrepreneurs mean that they will have greater weight in aggregate production.

This is depicted in figure 4.1 for $D' < 0$ including both cases where: (i) the wealth share of low-productivity entrepreneurs decrease and shifts up to higher productivity ones (panel a), and (ii) the wealth share of middle-productivity entrepreneurs decrease but shifts more upward than it does downward (panel b). A similar interpretation applies to $D' > 0$ where the

[21]The analysis in lemma 3 relies on the assumption that the density of the wealth distribution, $\omega(z)$ is uni-modal. I abstract from the irregular and multi-modal wealth distributions in order to carry out my analysis, however the quantitative version of my model does not rely on such distributional assumptions. Quantitative model is discussed in section 5.

[22]The distribution given in case (ii) of lemma 3 is the cumulative wealth distribution $\Omega(z)$ flipped in both axes' directions, which I call it inverted mirror.



(a) $D' < 0$, (case i)

(b) $D' < 0$, (case ii)

Figure 4.1: Equivalency conditions for $D' < 0$ in Lemma 3

low productivity entrepreneurs gain more (or lose less) wealth share compared to the high productivity ones. Now we turn to analysis of $\mathcal{Z}'$.

Lemma 4. *When c_e is arbitrarily small, $\mathcal{Z}'$ is given by*

$$\lim_{c_e \to 0} \mathcal{Z}' = \frac{1}{A}\left(\xi - \frac{\hat{z}}{\lambda M}\right) - \frac{D'}{M}. \tag{4.2}$$

Lemma 5. *When c_e is arbitrarily large, $\mathcal{Z}'$ is*

$$\lim_{c_e \to \infty} \mathcal{Z}' = \lim_{c_e \to \infty} \left(\frac{\Upsilon'}{1-\Upsilon} + \frac{\hat{z}'}{\hat{z}}\right). \tag{4.3}$$

Lemma 6. *We also have the following results when λ is arbitrarily large:*

(1) $\lim_{\lambda \to \infty} D = 0$,

(2) $\lim_{\lambda \to \infty} M = 0$,

(3) $\lim_{\lambda \to \infty} \lambda D = 0$,

(4) $\lim_{\lambda \to \infty} \lambda M = z_{max}$.

Proposition 6. *We have the following statements regarding the behavior of* $\mathcal{Z}'$.

1. $D' < 0$ *is a sufficient condition for existence of some* c_e *and* λ *where* $\mathcal{Z}' > 0$,

2. *For any given* λ, *there exist some* c_e *where* $\mathcal{Z}' < 0$.

Proof. For 1, I prove that given the sufficient condition, $\mathcal{Z}' > 0$ holds for some arbitrarily small c_e and large λ. Using the results from lemma 6 in the expression given in lemma 4, we see that the first term in right hand side is bounded and small, and the second term becomes larger (in absolute value) as λ increases. Therefore, the whole expression for $\lim_{c_e \to 0} \mathcal{Z}$ becomes positive if $D' < 0$.

For the second part of the proposition, we only need to prove that $\mathcal{Z}' < 0$ holds for some arbitrarily large c_e, which is the case from lemma 5. $\square$

Proposition 6 provides sufficient conditions for Z to be both increasing and decreasing when c_e changes. First part indicates that, for some combination of λ and c_e, TFP and entry cost move in the same direction if the wealth share of higher productivity entrepreneurs increases more (decreases less) than that of less productive entrepreneurs. In the proof I showed that, a high λ together with a low c_e is consistent with this result. The second part of the proposition states that TFP and entry cost move in opposite directions when entry costs are high. The following theorem gives one of the main results of this paper which is about the existence of TFP maximizing c_e given λ.

Theorem 1. *Let's define* $\lambda^* = \inf\{\lambda \in [1,\infty) : \mathcal{Z}'|_{c_e=0} > 0\}$.[23] *Given the sufficient condition in proposition 6, for any* $\lambda > \lambda^*$ *there exist some* $c_e^*(\lambda) \geq 0$ *that maximizes TFP. That is, if* Z *is differentiable in* c_e *then we have* $\mathcal{Z}'\Big|_{c_e^*(\lambda)} = 0$.

Proof. This follows from proposition 6. Given the sufficient condition: the first part of the proposition says that, for any $\lambda > \lambda^*$, we have $\mathcal{Z}'|_{c_e=0} > 0$. And the second part of the proposition says that for any λ (including λ^*) there exist a $c_e > 0$ such that $\mathcal{Z}' < 0$. These parts together prove the

[23]That means λ^* is the infimum value of λ that satisfies the first part of proposition 6, i.e, if $\lambda^* > 1$, we have $\lim_{\lambda \to \lambda^*} \mathcal{Z}'|_{c_e=0} = 0$.

existence of a TFP-maximizing entry cost for any given lambda, $c_e^*(\lambda)$. Also if Z is differentiable in c_e, we have $\mathcal{Z}' \Big|_{c_e^*(\lambda)} = 0$.[24] In the next subsection, I will discuss the sufficient conditions and how they are satisfied. $\square$

One simple interpretation of theorem 1 is that: there is a level of financial development above which zero entry cost is no longer what achieves the highest level of aggregate productivity. This is related to the main hypothesis of this paper, that if financial markets are developed enough in a country, decreasing entry costs will not necessarily lead to the higher levels of aggregate productivity.

4.2 A Discussion on Sufficient Condition

As we saw in proposition 6 and theorem 1, for high λ and low entry cost, $D' < 0$ is a sufficient condition for existence of such TFP-maximizing entry cost, $c_e^*(\lambda)$. Now I argue that, $D' < 0$ and $D' > 0$ correspond to productivity-based and wealth-based selection phases, respectively.

Given the equivalence of D' and second order stochastic dominance in lemma 3, we can see that $D' < 0$ means that the concentration of wealth moves away from lower productivity entrepreneurs toward the highly productive ones when we increase entry cost. This means that despite the decrease in entry rate, wealth shares move up the productivity ladder as a result of movement in two possible margins. It could be that, most of the entrants are highly productive which increases the number of high productivity entrepreneurs (extensive margin), or the highly productive incumbents have become relatively wealthier (intensive margin). Either of these two margins drive the wealth shares up toward higher productivity entrepreneurs, which I call the productivity-based selection phase.

In a similar fashion, $D' > 0$ means that the wealth shares move away from high productivity entrepreneurs toward lower productivity ones when entry cost increases. This means that most of the high productivity entrepreneurs cannot afford the entry cost and stay out while the wealthier individuals almost irrelevant of their productivity can enter and become entrepreneurs. Again both extensive and intensive margins are at work here, which push the wealth shares down toward lower productivity entrepreneurs. I call this

[24]Here I do not seek to prove the deferentiability of Z with respect to c_e. However, it should be noted that deferentiability has not been of any concern in my quantitative analysis given the functional form of Z.

the wealth-based selection phase. Note that at any given level of λ and c_e both productivity- and wealth-based selection forces are in play, and the dominating force determines the final impact.

Apart from the sufficient condition discussed above, it might be helpful to understand other conditions that would make it easier for theorem 1 to hold. These are the conditions that help satisfy $\mathcal{Z}' > 0$ for a given λ and some small c_e. Such conditions can be obtained from the first term on the right hand side of the equation (4.2) in lemma 4, $\frac{1}{A}\left(\xi - \frac{\hat{z}}{\lambda M}\right)$. This term provides the following additional conditions:

1. A high ξ: If current entry rate is high, it might imply that many of the entrants are not highly productive. Therefore, some increase in entry cost would decrease the entry rate and might actually stop some low productivity individuals from entering production.

2. A low $\frac{\hat{z}}{\lambda M}$: Here $\hat{z}$ is the productivity of entrants, and $\lambda M = \frac{M}{\Omega(z_{max})}$ is the wealth-weighted productivity of incumbents adjusted by their wealth share. A low $\frac{\hat{z}}{\lambda M}$ indicates that the current entrants are mostly low productivity individuals compared to the incumbents. In a case where $\frac{\hat{z}}{\lambda M}$ is very low, a higher entry cost would decrease the flow of lower productivity entrants into entrepreneurship, and would also protect the higher productivity incumbents.

4.3 Aggregate Output and Output-per-Worker

Aggregate output is given in corollary 2 by $Y = \left(\frac{w}{1-\alpha}\right) L$. This equation indicates that the impact of entry cost on output can be analyzed through the lens of employment, L and wages, w. So far I have ignored the impact of changes in entry costs on wage. Here I will argue that the productivity- and wealth-based selection phases can drive aggregate output in two opposite directions. So the aggregate output response would be similar to TFP's. From the first part of lemma 1, we know that the share of entrepreneurs decline as entry cost increases, which through the dynamics of my model means that L increases. A greater L would translate into lower wages. Therefore, given the equation in corollary 2 for aggregate output, the main point here is to see whether employment effect dominates wage effect or vice versa.

First, let's consider the productivity-based selection phase, where according to the results of lemma 3, an increase in entry cost leads to an increase in the wealth share of more productive entrepreneurs. In this phase, despite the decrease in the number of entrants which would put downward pressure on wages, the fact that more productive entrepreneurs are wealthier means that they will have more capital and therefore higher labor demand according to equation (3.3). This will put some upward pressure on wages and will offset some of the previous effect. As a result, a strong enough productivity-based selection force would imply a much less wage decrease compared to employment increase and as a result a direct relation between aggregate output and entry costs.

Now let's consider the wealth-based selection phase, where a higher entry cost means more wealth share for low-productivity entrepreneurs. In this phase when entry cost increases, the number of entrants decreases pushing wages down a bit. At the same time the wealth shares shift from high-productivity entrepreneurs toward low-productivity ones. This means even a greater decline in capital and labor demand pushing wages further down. The outcome is a greater drop in wages than the rise in employment as a result of increase in entry costs and therefore a reverse relation between aggregate output and entry costs.[25]

Since we have a constant population normalized to 1, the impact of entry cost change on output per capita will be the same as the impact on aggregate output which is also similar to the impact on TFP. As a result, with developed enough financial markets, if we increase entry costs up to some threshold, we will observe increase in aggregate output as well as in output per capita. But if we increase entry cost beyond such a threshold we will see a decline in both measures.

Output per worker is also given in corollary 2 by $\frac{Y}{L} = \left(\frac{w}{1-\alpha}\right)$. This means the entry cost impact is through the wage. The same arguments made above also hold here. In the productivity-based selection phase, wage changes mildly because of forces pushing in opposite directions[26], and in the wealth-based selection phase both forces put downward pressure on wages so it drops more significantly. As a result, although output per worker may decrease as we increase entry cost, it will not be at the same rate. The decline will be relatively small through the productivity-based selection phase and will be more significant through the wealth-based selection phase. Also as we

[25] These points are consistent with the results obtained from the quantitative model.

[26] In the quantitative model I did not come across a case where the upward pressure dominates the downward one which could mean a wage increase instead of a decrease. That is the reason I will presume that wages will decrease as we increase the entry cost.

will see in section 5, as the financial markets improve, the decline in output per worker becomes more insignificant.

4.4 Policy Analysis

I will do some policy exercises when it comes to the quantitative part of my analysis (will be added in the next version of the paper), but it might be useful to speculate on one particular exercise here. So far we saw that for any given level of financial development, λ, there exists a level of entry cost, $c^*(\lambda)$, where the aggregate productivity is at its maximum. Given the inter-relation of c_e and λ, one natural policy question that comes to mind after seeing the analysis in previous subsections is that: what should a policy maker do if her goal is to achieve highest possible TFP? Change the entry cost or improve the financial markets? From the perspective of the policy maker, TFP is a function of entry cost and financial development, $Z(\lambda, c_e)$. Every decision of the policy maker comes at a price. Let τ_λ and τ_{c_e} be the unit cost of improving financial markets and changing entry cost, respectively.[27] The best decision for the policy maker is to go in the direction of highest increase in TFP for every dollar spent. That is, the decision should be based on a cost-weighted gradient of function Z.

$$\nabla Z \cdot \tau^{-1} = \left(\frac{\partial Z}{\partial \lambda}, \frac{\partial Z}{\partial c_e}\right) \cdot \left(\frac{1}{\tau_\lambda}, \frac{1}{\tau_{c_e}}\right) \tag{4.4}$$

This is the same notion of marginal productivity per dollar spent for changes in λ and c_e. Here I use the same notation of x' for partial derivative of any variable x with respect ot c_e as used in the previous subsections, and I will denote by $\dot{x}$ the partial derivative of x with respect to λ. Without loss of generality, I will replace Z with $\log(Z)$ and use elasticity of Z with respect to λ and c_e. I get the following elasticity functions, denoted by ϵ:

$$\begin{aligned} \epsilon_{\mathcal{Z},\lambda} &= \lambda \frac{\partial \log(\mathcal{Z})}{\partial \lambda} \\ &= \lambda \left(\frac{\dot{\Upsilon}}{1-\Upsilon} + \frac{M + \lambda \dot{M} - \dot{\Upsilon}\frac{\hat{z}}{\xi} - \Upsilon\dot{(\frac{\hat{z}}{\xi})}}{\lambda M - \Upsilon \frac{\hat{z}}{\xi}} \right) \end{aligned} \tag{4.5}$$

[27] These costs can themselves be functions of multiple factors related to the current state of the economy. I will abstract from such cases to avoid complications in my analysis. However, when dealing with elasticities, we can think about these costs as costs of percentage improvements instead of unit improvements.

and

$$\begin{aligned}\epsilon_{\mathcal{Z},c_e} &= c_e\frac{\partial\log(\mathcal{Z})}{\partial c_e} \\ &= c_e\left(\frac{\Upsilon'}{1-\Upsilon}+\frac{\lambda M'-\Upsilon'\frac{\hat{z}}{\xi}-\Upsilon(\frac{\hat{z}}{\xi})'}{\lambda M-\Upsilon\frac{\hat{z}}{\xi}}\right). \qquad (4.6)\end{aligned}$$

The policy choice breaks down to the following:

P1. Increase c_e if $\frac{\epsilon_{\mathcal{Z},\lambda}}{\tau_\lambda}<\frac{\epsilon_{\mathcal{Z},c_e}}{\tau_{c_e}}$,

P2. Decrease c_e if $\frac{\epsilon_{\mathcal{Z},\lambda}}{\tau_\lambda}<-\frac{\epsilon_{\mathcal{Z},c_e}}{\tau_{c_e}}$,

P3. Increase λ if $\frac{\epsilon_{\mathcal{Z},\lambda}}{\tau_\lambda}>\left|\frac{\epsilon_{\mathcal{Z},c_e}}{\tau_{c_e}}\right|$.

Equations (4.4), (4.5) and (4.6) can provide the conditions required for choosing a policy. I will talk about some specific policy analysis in the quantitative model.

5 A Quantitative Analysis

5.1 Extended Model

In this part I will discuss the model in quantitative framework while relaxing some of the assumptions made in the previous sections. I use a more general CES form of utility function instead of log form used so far.

$$u(c)=\frac{c^{1-\sigma}}{1-\sigma}.$$

I also replace the constant returns to scale production function with a decreasing returns to scale one.

$$f(z,k,l)=zk^\alpha l^\theta, \qquad \alpha+\theta<1.$$

Given these two changes, the entrepreneurs will solve the same problem given by (3.1), but the choice of capital and labor will be different because of the DRS technology:

$$k(a,z)=\min\{k^*(z),\lambda a\},$$

where

$$k^*(z) = \left(z \left(\frac{\theta}{w} \right)^{\theta} \left(\frac{\alpha}{r+d} \right)^{1-\theta} \right)^{\frac{1}{1-\alpha-\theta}},$$

and

$$l(a,z) = \left(z \left(\frac{\theta}{w} \right) k^{\alpha}(a,z) \right)^{\frac{1}{1-\theta}}.$$

Entrants can borrow up to λa to finance the entry cost as well as the starting capital. The entry cost is paid upfront but the entrants' productivity should be high enough so that they are able to pay back the interest on their loan. This puts lower productivity entrants in a disadvantaged position as if their entry cost is higher compared to more productive ones. We will have a similar expression as given by (3.7) for the effective entry cost paid by entrants with different productivity levels. The exact effective entry cost is provided in the appendix.

The value functions for entrepreneurs and workers will be given by the same HJB equations in (3.13) and (3.14) respectively. The quantitative results support the properties of entry and exit boundaries provided in corollary 1, and defines similar entry, exit and inaction zones given by definition 3.1. Figure 5.1 illustrates this for different values of entry costs as well as two different levels of λ. As we can see in figure 5.1 a higher entry cost means a smaller entry region. Also a greater λ means larger entry zone at every given entry cost, as well as a shrunk exit region.

Similarly, the stationary distribution of agents are given by Kolmogorov Forward equations in (3.22). I have used finite difference method to solve for both value functions and distributions. The market clearing conditions are also the same as (3.25) and (3.26). Using the solutions of HJBQVI and KFEs as well as market clearing conditions I solve for the stationary equilibrium of the economy. In the quantitative framework I have used Ornstein-Uhlenbeck process for agents' idiosyncratic shocks.[28] It is worth noting that the setup of the problem makes sure that every entrepreneur will at some point exit because there is a positive probability of receiving a shock that will put individuals in the region $\eta(z) < 0$ where they will exit regardless of their wealth. This will ensure that no entrepreneur will become too wealthy to take over the economy and that there exists a stationary equilibrium. The algorithm for solving the stationary equilibrium is described below.

[28]See Moll (2014).

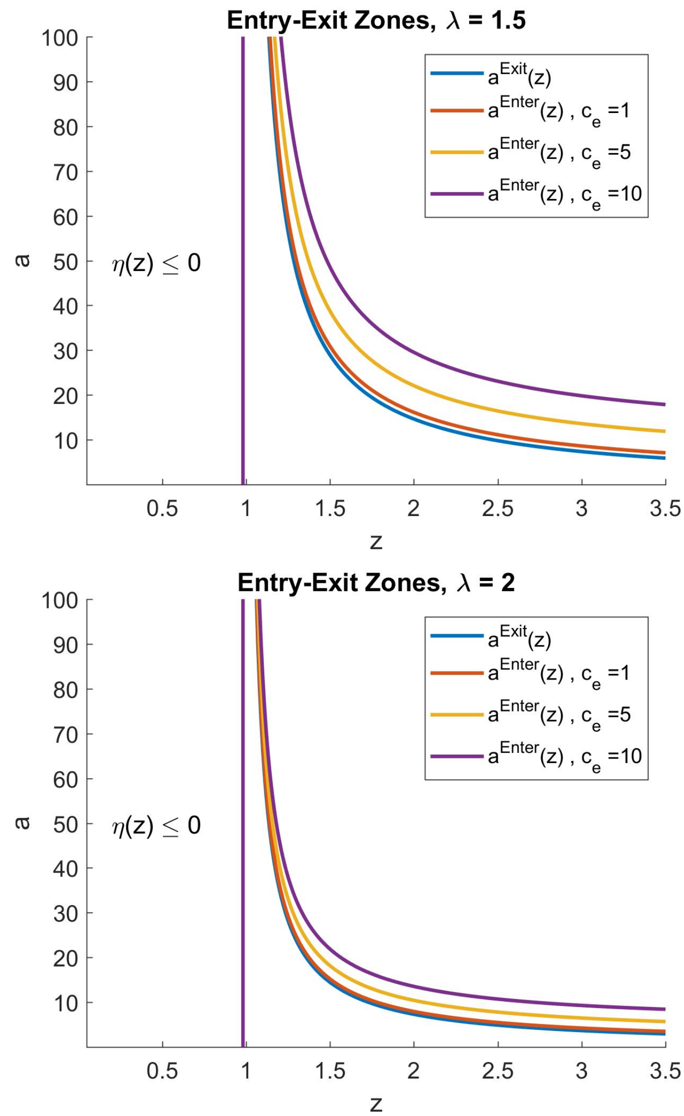


Figure 5.1: Exit and Entry boundaries

Algorithm:

Start with an initial guess on wages, w^0, and interest rate, r^0. Then, for $s = 0, 1, 2, ...$ repeat the following:

1. Given the prices, for any occupation $\mathcal{O} \in \{W, E\}$ solve for the optimal consumption, $c^{\mathcal{O}}$ using the first order conditions from HJBs.

2. For the entrepreneurs solve for the optimal capital and labor, using (3.2, 3.3) as well as using the borrowing limit. Then solve for the profits Π^E for any point in the state space.

3. Solve for the entrants' equivalent wealth upon entry, a', considering the effective entry cost $\hat{c_e}(z)$. Given the wealth upon entry we can solve for the value of entry, $V^{*E}(a, z) = V^E(a', z)$.

4. Also since there is no irreversibility costs, the wealth of exiting entrepreneur will not change, and therefore $V^{*W}(a, z) = V^W(a, z)$.

5. Given the outside options from steps 3 and 4, solve for the value function as a linear complementarity problem (LCP), and obtain $V^{\mathcal{O}}$.

6. Repeat steps 1-5 until both value functions converge to their respective values, $(V^W)^s$ and $(V^E)^s$, and obtain entry, exit and inaction zones.

7. Start with initial some distributions for $j = 0$, $(g^W)^j$ and $(g^E)^j$, and use the entry and exit zones obtained in step 6 to obtain the distribution of entering and exiting individuals given by equations (3.23) and (3.24).

8. Solve for the next distributions, $(g^W)^{j+1}$ and $(g^E)^{j+1}$, from KFEs (3.22) using finite difference method, and repeat steps 7 and 8 until the distributions converge.

9. Solve for the aggregate outcomes using the distribution obtained in step 8. Use the market clearing conditions (3.25 and 3.26) to update the prices.[29]

10. Given the newly obtained values for prices and aggregate quantities, return to step 1. Stop iterating if the changes in prices (or aggregate

[29] I have used Broyden's method to update prices using aggregates and market clearing conditions

quantities) are very small. That is, stop if the following condition holds for some desirably small ϵ:

$$\left|w^{s+1} - w^{s}\right| + \left|r^{s+1} - r^{s}\right| < \epsilon.$$

This provides the stationary equilibrium of the economy which is given by $\left(w^s, r^s, (V^{\mathcal{O}})^s, (g^{\mathcal{O}})^s\right)$ and the aggregate outcomes. One example of stationary distributions is illustrated in figures 5.2 and 5.3.

Regarding the main result of this paper in theorem 1, I have run a quantitative exercise to see if I can get compatible results in the less restricted model. The only goal of this exercise is to compare the results with those derived in section 4. Figure 5.4 basically confirms the findings of theorem 1. It is clear from figure 5.4 that as financial markets improve, i.e. as λ increases, the optimal level of TFP improves and is reached at a higher level of entry cost. This is in agreement with the results discussed in section 4. In other words, if financial markets are highly developed, it is better to have some sort of filter to keep unproductive individuals from entering business and that filter needs to be stronger as financial markets become more and more accessible for everyone in the economy.

5.2 Calibration

This section is to be completed in the next version of the paper.

6 Conclusion

In this paper I revisited a question on the impact of entry costs on aggregate productivity, and analyzed the impact of entry costs on aggregate outcomes. I emphasized the role of financial development in the dynamics of my model, a role which has been ignored to a certain extent in previous analyses of entry costs. My theory suggests that for highly developed financial markets, some non-zero level of entry cost might achieve optimal level of TFP. This is in contrast with developed intuitions on the function of entry costs, which I show could only be consistent with the economies with poor financial markets. Cross country observations also indicate that for countries with less developed financial markets, higher entry costs are tied to lower TFP. However, for the countries with well-developed financial markets, high entry costs

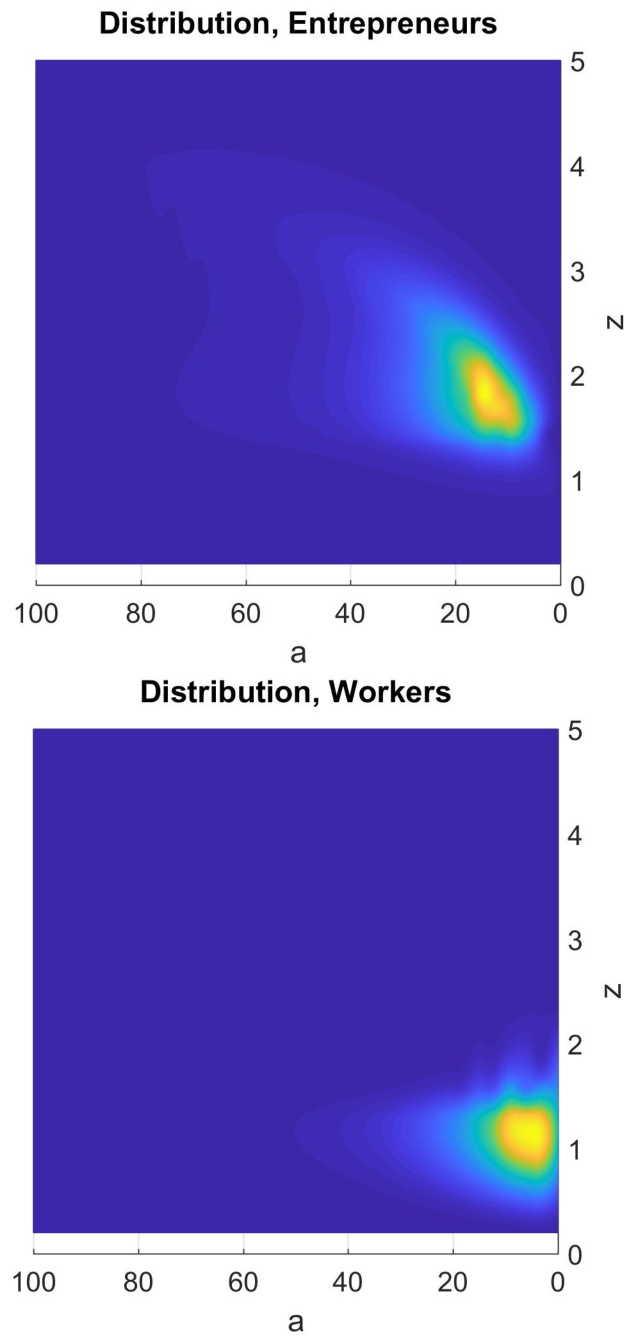


Figure 5.2: Stationary Distributions

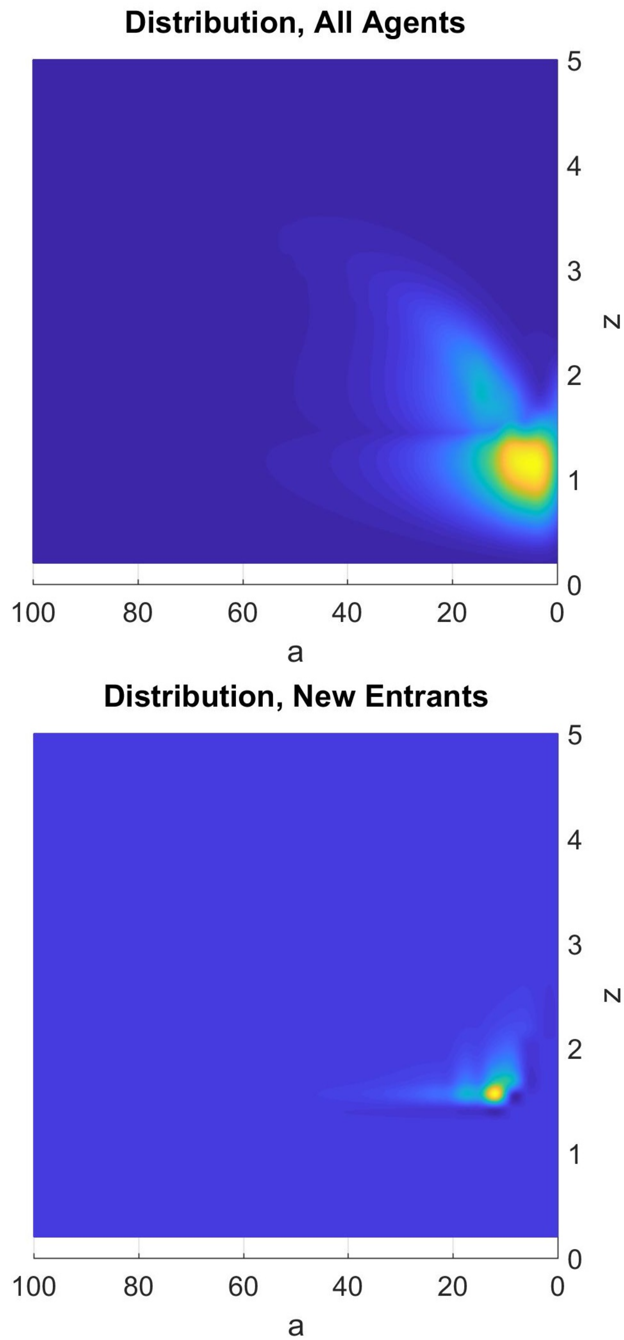


Figure 5.3: Stationary Distributions

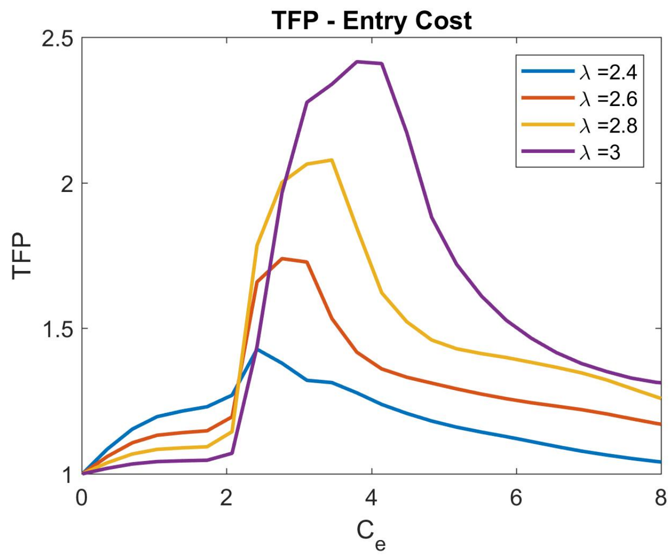


Figure 5.4: TFP vs entry cost for different levels of financial development.

are associated with high TFP. The results hold for different proxies for key variables even when controlling for some certain effects. I finally relaxed some model assumptions used in sections 3 and 4, and developed a less restrictive model and solved it in a quantitative setting. The preliminary results of my quantitative analysis are also consistent with the findings of the analytical model as well as the cross-country observations.

It is worth mentioning that my analysis only suggests such a hump shaped relationship for TFP, aggregate output and output per capita. However, if a policy maker cares the most about output per worker, my analysis shows no evidence for such a hump shape relationship, and the best we can say is that the negative impact of entry cost on output per worker becomes smaller for economies with highly developed financial markets.

An important missing piece in my analysis is the sectoral dynamics. Adding such a feature to my model would make it impossible to proceed with an analytical solution, and there can only be a quantitative analysis. In a parallel work I have developed a two sector model similar to the one by Buera et al (2011) and looking into cross-industry data sets to see if there is evidence for the role of entry costs across different sectors and industries. I am also considering adding different types of adjustment costs for employment and investment and see how they impact the dynamics of my model and the channel through which entry costs impact aggregate outcomes.

References

[1] Acemoglu, D., Aghion, P. and Zilibotti, F., 2006. Growth, development, and appropriate versus inappropriate institutions. EBRD-Japan Research Programme Papers, Banque européenne pour la reconstruction et le développement, Londres.

[2] Amaral, P.S. and Quintin, E., 2010. Limited enforcement, financial intermediation, and economic development: a quantitative assessment. International Economic Review, 51(3), pp.785-811.

[3] Bah, E.H. and Fang, L., 2016. Entry costs, financial frictions, and cross-country differences in income and TFP. Macroeconomic Dynamics, 20(4), pp.884-908.

[4] Barseghyan, L., 2008. Entry costs and cross-country differences in productivity and output. Journal of Economic Growth, 13(2), pp.145-167.

[5] Barseghyan, L. and DiCecio, R., 2011. Entry costs, industry structure, and cross-country income and TFP differences. Journal of Economic Theory, 146(5), pp.1828-1851.

[6] Buera, F.J., Kaboski, J.P. and Shin, Y., 2011. Finance and development: A tale of two sectors. American economic review, 101(5), pp.1964-2002.

[7] Buera, F.J. and Shin, Y., 2013. Financial frictions and the persistence of history: A quantitative exploration. Journal of Political Economy, 121(2), pp.221-272.

[8] Buera, F. and Nicolini, J.P., 2017. Liquidity Traps and Monetary Policy: Managing a Credit Crunch: Online Appendix", Staff Report, No. 541, Federal Reserve Bank of Minneapolis.

[9] Cagetti, M. and De Nardi, M., 2006. Entrepreneurship, frictions, and wealth. Journal of political Economy, 114(5), pp.835-870.

[10] Djankov, S., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A., 2002. The regulation of entry. The quarterly Journal of economics, 117(1), pp.1-37.

[11] Hopenhayn, H.A., 1992. Entry, exit, and firm dynamics in long run equilibrium. Econometrica: Journal of the Econometric Society, pp.1127-1150.

[12] Hsieh, C.T. and Klenow, P.J., 2009. Misallocation and manufacturing TFP in China and India. The Quarterly journal of economics, 124(4), pp.1403-1448.

[13] Itskhoki, O. and Moll, B., 2019. Optimal development policies with financial frictions. Econometrica, 87(1), pp.139-173.

[14] Klenow, P.J. and Rodriguez-Clare, A., 2005. Externalities and growth. Handbook of economic growth, 1, pp.817-861.

[15] Levine, R., 2005. Finance and growth: theory and evidence. Handbook of economic growth, 1, pp.865-934.

[16] Matsuyama, K., Gertler, M. and Kiyotaki, N., 2007. Aggregate implications of credit market imperfections. NBER Macroeconomics Annual, 22, pp.1-81.

[17] Midrigan, V. and Xu, D.Y., 2014. Finance and misallocation: Evidence from plant-level data. American economic review, 104(2), pp.422-58.

[18] Moll, B., 2014. Productivity losses from financial frictions: Can self-financing undo capital misallocation?. American Economic Review, 104(10), pp.3186-3221.

[19] Moscoso-Boedo, H.J. and Mukoyama, T., 2012. Evaluating the effects of entry regulations and firing costs on international income differences. Journal of Economic Growth, 17(2), pp.143-170.

[20] Nicoletti, G. and Scarpetta, S., 2003. Regulation, productivity and growth: OECD evidence. Economic policy, 18(36), pp.9-72.

[21] Nicoletti, G. and Scarpetta, S., 2006. Regulation and economic performance: product market reforms and productivity in the OECD. Institutions, Development, and Economic Growth, 13, p.81.

[22] Poschke, M., 2010. The regulation of entry and aggregate productivity. The Economic Journal, 120(549), pp.1175-1200.

[23] Shaker Akhtekhane, S., 2017. Firm Entry and Exit in Continuous Time. Working paper, Ohio State University.

[24] Townsend, R., 2010. Financial structure and economic welfare: Applied general equilibrium development economics. Annu. Rev. Econ., 2(1), pp.507-546.