

CSAE WPS/2010-07

The Dynamics of the Informal Economy*

Roxana Gutiérrez-Romero

Oxford University

Department of International Development

roxana.gutierrezromero@qeh.ox.ac.uk

Abstract

This paper analyses the factors that give rise to the existence of the informal economy and how it evolves over time. Using an occupational-choice model the paper shows that at early stages of development, informal and formal markets coexist, but in the long-run the size of the informal economy can decline depending on the initial distribution of wealth. The model shows that the higher the initial wealth inequality the larger the size of the informal economy and the higher the wealth inequality will be in the long-run. The paper calibrates the model using numerical simulations.

JEL Classification Numbers: D31, K4

Keywords: informal economy, occupational choice and inequality.

*I am grateful to James Mirrlees for his valuable comments and advice. I also wish to thank Tim Besley, Partha Dasgupta, Adam Pepelasis and the participants of the seminars at Trinity College Dublin, University of York, University of Cambridge, University of Oxford and the World Bank Research Network on Inequality and Poverty for their comments. I acknowledge funding from CONACYT and from the UK Department for International Development (DFID) as part of the Improving Institutions for Pro-Poor Growth (iiG), a research consortium aimed at studying how to improve institutions in Africa and South-Asia. The views expressed are not necessarily those of DFID.

1 Introduction

In developing countries the poor and even the middle classes often produce and satisfy most of their financial needs from informal markets. Although informal markets represent an alternative form of employment and production for many, they too impose large costs. Tax evasion, disrespect for the rule of law, unfair competition, among others, impose a big challenge for governments to face. To adopt adequate policies to foster development a better understanding is needed of the factors that drive the dynamics of the informal economy and the impact that informal markets have on development.

There has been a growing interest on studying the determinants of informality. In particular, the efforts have focused on empirically assessing the factors that might increase informality such as start up costs (Antunes et. al., 2007; Djankov et. al., 2002; Strabu, 2005), high tax burden (Cebula, 1997; Giles and Tedds, 2002; Ihrig and Moe, 2004; Schneider and Enste, 2000), excessive regulation, especially in the labor market (Albrecht et. al., 2009; Friedman et. al., 2000; Fugazza and Jacques, 2004; Maloney, 2004; Perry et. al., 2007), and institutional quality such as bureaucracy and corruption (Friedman et. al., 2000; Chong and Gradstein, 2007; Johnson et. al., 1997, 1998; Loayza, 1996). Most of these studies have focused on analyzing the informal economy in the short run, and the minority of studies that have analyzed the informal economy from a dynamic perspective suggest that it will naturally tend to decline over time. Indeed most of the already developed countries show a reduction of the informal economy. However, in developing countries such as Latin America and the Sub-Saharan Africa, the size of the informal economy, employing over half of the urban labor force, hardly suggests that the informal economy is declining. If anything, the trend of informality in these countries suggests that it is on the rise (Perry et. al. 2007).

The objective of this paper is to contribute to the literature by assessing the main determinants of the informal economy, and most importantly, how the informal economy evolves over time and its impact on development in the long-run. For this purpose a dynamic model is developed taking into account the imperfections in credit markets found in developing countries. In the model each individual lives for two periods. At the beginning of the first period each individual faces an occupational choice problem. Individuals decide whether to become entrepreneurs or workers, either in the formal or in the informal economy. The entrepreneurs maximize the size of their firm according to their initial wealth and the credit they can get. Credit can be acquired in the formal credit market, if entrepreneurs are registered

with the tax authorities, which involves paying a registration fee and taxes on profits and labor. Otherwise, credit can be obtained in the informal credit market, but at higher interest rate.

Following the dynasty of each individual over time, the model shows that the occupation-choice of each offspring depends on the initial distribution of wealth, and the structure of occupational choice (in informal and formal sectors) in turns determines people's rents and savings giving rise to a new distribution of wealth. The paper focuses on the dynamics of the model which are non-linear and which can reach a number of distinct long-run equilibria. The model is calibrated with numerical examples which give some insight into the four types of equilibria the model can reach. One of these equilibria shows that, the economy at early stages of development can simultaneously have formal and informal economies, but in the long-run the size of the informal economy can decline to a negligible size, as is the case with developed countries. This depends crucially on the initial wealth distribution and parameters of the economy. If there are enough entrepreneurs with wealth such that they set up businesses in the informal sector at medium scale, in the long-run, the expansion of their scales of production will lead to higher demand for labor, higher salaries and eventually the formalization of their businesses. Given the higher costs of credit in the informal market, the incentive to set up business in the formal sector increases with the size of the firm. The model also explains why at early stages of development, the higher the inequality is the larger the informal economy will be. At high levels of inequality there is a smaller number of entrepreneurs in the formal sector and therefore little competition in the labor market, which is reflected in low salaries. Potential workers prefer to set up business in the informal sector than earn the low salaries offered. It is only when entrepreneurs (in the formal and informal sector) start expanding their scales of production that salaries will rise, reducing the gap between the earnings of a worker and an entrepreneur. This will be reflected in a decrease in inequality. If the expansion in scale of production is such that it becomes more profitable for entrepreneurs to formalize business then eventually the informal markets will shrink. However, informal and formal sectors will coexist in the long-run if there were originally high levels of inequality such that only a few entrepreneurs manage to formalize businesses and stay afloat in that sector.

The paper contributes to the literature by identifying inequality in wealth, high start up-front costs, inadequate tax policies and imperfect credit markets as the main determinants of

the existence and size of the informal economy. Furthermore, the paper shows that although in early stages inequality in wealth is one of the main causes for the existence of the informal sector, informality later on could increase inequality and polarization of rents between people engaged in both sectors over time. The size of the informal economy will not necessarily decline in the long-run, as is generally assumed in the literature, but an equilibrium could be reached where there is low aggregated wealth and formal and informal economies coexist. The current high levels of inequality in wealth in Latin America and the red tape associated with formalizing businesses in the region might well explain why the informal sector currently absorbs over half of the urban labor force (Perry et. al. 2007). The size of this informal sector, might in the future increase inequality further thereby increasing informality over time if no adequate policies are taken. If that is the case the long-run equilibrium predicted will be one where the formal and informal sectors coexist with low aggregated levels of wealth and wealth concentrated in the hands of a few big entrepreneurs.

The paper proceeds as follows. Section 2 discusses the definition of the informal economy and presents some data on the extent of the informal economy world-wide. Section 3 presents the model. Section 4 presents numerical examples and discusses the main predictions of the model. Section 5 summarizes the conclusions of the paper.

2 What is the Informal Economy?

The United Nations in 1993 suggested an international definition of the informal economy, with the view to incorporate it into the national account statistics.¹ This market was described as consisting of units that provide goods or services with the aim of creating jobs and income for the individuals implicated. These units are characterized by very low levels of organization, with little or no division between labor and capital, working at small scale. The labor relationships, if present, are based generally on an occasional job basis, and a family, personal or social relationship, rather than on labor contracts with any sort of formal warranties.

Although this definition is a useful attempt to describe the activities included in the informal economy, these activities have many faces that vary across countries and over time. For this reason these activities have been described in many ways, such as parallel, residual,

¹UNDP (1993), National Accounts Systems, Brussels.

illegal, informal, unofficial, black, curb, shadowed, subterranean and informal activities. As a result, there is yet no international consensus on whether activities such as street vending, common in developing countries, should be conceived as part of the informal economy or as illegal activities such as drug dealing, trading with stolen goods and the piracy of products. In addition to the difficulty of grasping the range of activities of the informal economy, further problems arise; current surveys show that the informal economy has managed to organize its operations under different kinds of labor and credit contracts and that for some people the informal economy no longer represents merely a temporary job, but a secure and permanent source of income.² In this paper the informal economy is understood as including economic activities that do not comply with the government-imposed taxes and regulations. This definition has been used in other work by De Soto (1989), Portes et. al. (1989) and Loayza (1996).

There is a large literature focused on explaining the causes of the existence of the informal economy.³ Currently there are two main views on the subject. The first one suggests that the most important factors that make more likely the existence of the informal economy are large start up costs, rigid labor legislation, inefficient tax systems, corruption and high registration fees (Loayza, 1996; Johnson et. al., 1997; Johnson et. al., 1998). The second view argues that entrepreneurs have less of an incentive to formalize their businesses if they are not constrained in the informal economy from the public goods and services available to the formal sectors. Otherwise, entrepreneurs might have more incentives to formalize to improve their profits and productive infrastructure (Johnson et. al., 1998; Loayza, 1996; Montiel et. al., 1993).

Although the theoretical literature has found empirical support, there is still more to be done particularly on two issues. Firstly, it is necessary to assess what the real costs and penalties are that individuals face working informally and why despite these costs, the informal economy represents an alternative source of income for a large proportion of the population in developing countries. Secondly, it is essential to study the dynamics of the informal economy and specifically whether its dynamics are related to the dynamics of development. For instance, with respect to what sort of services people engaged in the informal economy will

²Heikki and Jose A. Pegan (2001) using the National Mexican Survey of Microenterprises conducted by INEGI, found that 66% of microenterprise owners entered the sector to stay permanently and 21.2% plan to expand their current size of operations. About 54% of these microenterprises can be considered as part of the underground economy, by not complying with tax regulations.

³For a literature review on the topic see Schneider and Enste (2000).

be constrained from, table (1) presents some of the results on the 2003 Global Survey of Public Credit Registries in 64 countries carried out by the World Bank. The data show that banks require information from firms that would be very difficult for entrepreneurs working in the informal market to present. For example, in 60.9% of the countries analyzed banks collect the taxpayer ID from the firm and in 46% of the countries analyzed the banks may distribute to individuals, firms or governmental institutions the taxpayer ID.

Given that entrepreneurs in the informal sector are constrained from formal sources of credit, what would it cost them to formalize their business? The 2003 “Doing Business” database of the World Bank includes information on the bureaucratic and legal procedures involved in incorporating and registering an average small, or medium-sized company into the formal economy across 133 countries.

The data show that the economies surveyed differ significantly in the way they regulate the entry of new businesses. For instance, in Latin America and the Caribbean entrepreneurs can expect to go on average through 12 steps to launch a business over 74 days, and the amount that entrepreneurs must deposit in a bank before registration is equal to 68% of gross national income (GNI) per capita. These regions happen to have the largest informal economies, and highest levels of inequality.⁴ In addition to these procedures, entrepreneurs face up-front fees such as notaries fees costing approximately \$1,000 in some countries such as Mexico and Uruguay. Similarly, Sub-Saharan Africa, Europe and Central Asia the notaries’ fees account for up to 50% of the total starting cost for setting up a business.

⁴The size of the informal economies in Africa, Asia, Transition countries and Latin America were estimated according to the physical input (electricity) method, the currency demand and the DYMIMIC (dynamic multiple-indicators multiple-causes) approach. In the OECD countries the size of the informal economies was estimated using the currency approach. The physical input estimates the size of the informal economy by subtracting the estimates of official GDP from the overall electricity consumption. The DYMIMIC model measures the informal economy as an unobserved variable. The currency demand method estimates econometrically the demand for currency, the excess increase in currency is attributed to an increase in the size of the informal economy.

Economy	Informal Economy (%GNI)	Gini	Number of Procedures	Duration (days)	Min. Capital (% GNI)	Employment Law Index	Private Bureau Coverage (borrowers per 1000 capita)
Sub-Saharan Africa	42.3	45.9	11	73	237.6	52	36.3
Latin America & Caribbean	41.5	50.9	12	74	85.6	61	196.6
Europe & Central Asia	37.7	32.5	10	48	123.9	57	38.6
South Asia	35.7	34.7	8	44	86.1	49	1.8
Middle East & North Africa	27.5	36.1	11	50	410.2	48	14.9
East Asia & Pacific	24.3	41.7	8	66	68	45	107.8
OECD: High Income	16.8	31.1	6	30	61.2	45	443.5

Source: 2003 Doing Business Database, World Bank (2003). Gini taken from World Development Indicators, World Bank (2003).

Table 1: Regional Statistics for the Informal Economy, Inequality and Business Environment

The next section presents a model that aims at capturing the main constraints that entrepreneurs face in the formal and informal markets, studying the dynamics of the informal economy and its impact on development over time.

3 The Model

There is a constant population, N , of two-period lived individuals belonging to generations of altruistic families. Individuals in each generation receive a bequest from their parents, b , and choose an occupation during the first period of life. During the second period of life individuals consume, x . The proportion consumed from their wealth, $(1 - \beta)$, is assumed to lie between 0 and 1. The utility preferences over consumption and bequest, $U(x, b)$, are expressed in eq.(1).

$$U(x, b) = (1 - \beta) \log x + \beta \log b \quad (1)$$

Individuals choose the occupation that maximizes their utility subject to their wealth constraint. The occupational choices are to become either an entrepreneur, or a worker, or unemployed. It is assumed that except for the entrepreneurial activity, the rest of the occupations do not require any starting up investment to enrol in economic activity.

The economy is a one good world that can be produced either in the formal sector, A , or in the informal sector, B . In sector A , the government requires entrepreneurs to register with the tax office, which implies paying a fixed registration fee⁵ F . It also implies paying

⁵This fee could be interpreted as an administrative registration fee or a fee charged by notaries.

a tax rate τ^p proportional to the returns of the firm and a tax rate τ^s proportional to the salary of workers.

A worker earns a net salary $s^A(1 - \tau^s)$ if hired in sector A . In sector B , a worker earns a salary s^B and evades taxes. Nevertheless, entrepreneurs in sector B offer competitive salaries that are less or equal the net salary in sector A , $s^B \leq s^A(1 - \tau^s)$ depending on the total demand for labor as it will be explained later. An unemployed person does not receive an unemployment benefit. However, it is assumed that unemployed people using their labor are able to produce some fixed amount γ , to remain in subsistence. This amount is not liable to taxes.

Entrepreneurs at time t use all their initial wealth w_i , to cover the costs associated with acquiring the machinery for setting up business. It is assumed that the only technology available in the economy is a Leontief production function in the sense that it uses two input factors, capital and labor which are not substitutable. The rigidity of this production function is given by the fact that although in the aggregate the economy might choose to specialize in labor or capital, for an entrepreneur there is a limit to this specialization in one of these inputs.

In sector A , machinery must be acquired at time t , at a fixed cost σ plus a variable cost $\alpha_i^A \sigma$ that depends on the scale of production, α_i^A . The machinery depreciates, such that the net value of the machinery at time $t + 1$ is $c(\sigma + \alpha_i^A \sigma)(1 + r)$, where c is the proportion left after depreciation, and $(1 + r)$ is the discount factor. The project at time $t + 1$ yields a return ρ times the scale of the firm α_i^A . Given the initial wealth, the maximum scale of production an entrepreneur can set up is $\widehat{\alpha}_i^A$ in eq.(2). An entrepreneur does not have incentives to invest in machinery an amount smaller than his initial wealth, given that there is no risk in this economy and it is assumed that the returns of running the project ρ are higher than the returns to saving $(1 + r)$.

$$w_i = F + \sigma + \widehat{\alpha}_i^A \sigma \quad (2)$$

On the other hand, entrepreneurs in sector B need to acquire machinery at a fixed cost σ , plus a variable cost $\alpha_i^B \sigma$ that depends on the scale of production, α_i^B . The machinery depreciates at the same rate as in sector A . The project at time $t + 1$ yields a return ρ times the scale of the firm α_i^B . The maximum scale of production when the entrepreneur invests all his wealth is $\widehat{\alpha}_i^B$.

$$w_i = \sigma + \widehat{\alpha}_i^B \sigma \quad (3)$$

In both sectors, the scale of production can be increased if the entrepreneur gets a loan.

3.1 Credit Markets

Banks in the formal credit sector require borrowers to be registered with the tax office. Banks offer a loan contract conditional on borrowers providing a collateral, which can only be machinery. Wealth is not collateralisable, given that it is a liquid asset and that entrepreneurs can easily run away with it. Then an entrepreneur will offer as a collateral the property rights of his machinery, $\sigma + \widehat{\alpha}_i^A \sigma$, which is in fact the value of his initial wealth less the cost of the registration fee F . It is assumed that the interest rate, r is fixed over time. The underlying assumption is that the economy analyzed is a small economy subject to international interest rates that remains fixed over time.

$$w_i - F = \sigma + \widehat{\alpha}_i^A \sigma \quad (4)$$

At time t the maximum value of the loan L_i^A is equal to or smaller than the value of the collateral, eq.(5).

$$L_i^A \leq \sigma + \widehat{\alpha}_i^A \sigma \quad (5)$$

Once the project is set up and makes profits the entrepreneur could try to default on the repayment of the loan $(1+r)L_i^A$, given that the collateral has depreciated and at $t+1$ is just worth $c(\sigma + \widehat{\alpha}_i^A \sigma)(1+r)$.

To prevent this, banks impose an ex post liability constraint to increase the expected cost of default. It is assumed that banks will seize the total value of the depreciated collateral. An entrepreneur will honour the loan if the cost of the loan is less than or equal to the net value of the collateral at time $t+1$.

$$(1+r)L_i^A \leq c(\sigma + \widehat{\alpha}_i^A \sigma)(1+r) \quad (6)$$

From eq.(4) $\sigma + \widehat{\alpha}_i^A \sigma$ is equal to $w_i - F$, hence the loan L_i^A is equal to,

$$L_i^A = c(w_i - F) \quad (7)$$

An entrepreneur with an initial wealth, w_i and the loan L_i^A will run the project at a scale α_i^A , which is greater than $\widehat{\alpha}_i^A$, the scale of the project the entrepreneur would have run without the loan.

$$w_i - F + L_i^A = \sigma + \alpha_i^A \sigma \quad (8)$$

Solving for the scale α_i^A , the optimum scale of the project is given by,

$$\alpha_i^A = \frac{w_i - F + L_i^A - \sigma}{\sigma} \quad (9)$$

Substituting L_i^A , the optimal scale is,

$$\alpha_i^A = \frac{(w_i - F)(1 + c) - \sigma}{\sigma} \quad (10)$$

Banks will advance loans to individuals whose collateral is positive, regardless of the size of the scale of production. This is because it is assumed that even if the entrepreneur makes losses banks will always get hold of the collateral provided.

The net returns η_i^A of entrepreneurs with loan L_i^A in sector A are determined by the net returns of the scale of production $\alpha_i^A \rho(1 - \tau^\rho)$ plus the value of the depreciated machinery $c(\sigma + \alpha_i^A \sigma)(1 + r)$, minus the salary bill $\alpha_i^A s^A$ and minus the repayment of credit $(1 + r)L_i^A$.

$$\eta_i^A = \alpha_i^A \rho(1 - \tau^\rho) + c(\sigma + \alpha_i^A \sigma)(1 + r) - \alpha_i^A s^A - (1 + r)L_i^A \quad (11)$$

Entrepreneurs in sector A , will require loans only if the project yields higher returns than the net return of other possible occupation. For instance the returns η_i^A need to be equal to or higher than the earnings of a worker, given by his disposable salary $s^A(1 - \tau^s)$ plus savings $w_i(1 + r)$,

$$\eta_i^A \geq w_i(1 + r) + s^A(1 - \tau^s) \quad (12)$$

Substituting η_i^A eq.(11), α_i^A eq.(10) and L_i^A eq.(7) into eq.(12), the initial wealth w_i , has to be equal to or greater than a wealth threshold w^A (see eq.(13)) to obtain returns η_i^A higher than workers' return,

$$w_i \geq \frac{S - F\left(\frac{1+c}{\sigma}\right)s^A - s^A\tau^s}{Q - \left(\frac{1+c}{\sigma}\right)s^A} = w^A \quad (13)$$

where $Q = \frac{1+c}{\sigma}\rho(1-\tau^\rho) + (1+r)(c^2-1)$, and $S = F\{(1+r)c^2 + \rho(1-\tau^\rho)\left(\frac{1+c}{\sigma}\right)\} + \rho(1-\tau^\rho)$.

Another way to interpret eq.(12) is by solving for the maximum salary \bar{s}_i^A that the entrepreneur i can afford to pay in sector A and still have a higher return than workers.

$$s^A \leq \frac{w_i Q - S}{w_i\left(\frac{1+c}{\sigma}\right) - F\left(\frac{1+c}{\sigma}\right) - \tau^s} = \bar{s}_i^A \quad (14)$$

Entrepreneurs can evade the registration fee and taxes by not registering with the tax authorities, which will constrain them from access to formal credit markets. Nevertheless, entrepreneurs can borrow from the informal credit market.

The distinction between the formal and the informal credit markets is that the former has a competitive interest rate equal to r , whilst a lender in the informal credit market faces a positive extra cost, z , in keeping track of each borrower⁶. This “tracking” cost depends on the amount borrowed L_i^B . The lender will charge an interest rate r^B , where $r^B = r + z$.

$$r^B L_i^B = r L_i^B + z L_i^B \quad (15)$$

The lender will also ask for an illiquid collateral, such as machinery. The reason for not accepting wealth as a collateral is the same as that for banks in the formal credit sector; wealth is liquid and hence the entrepreneur could easily default paying back the loan. Therefore, at time t the informal lender will accept as a collateral the value of the machinery $\sigma + \hat{\alpha}_i^B \sigma$, where $\hat{\alpha}_i^B$ is the scale of the project the entrepreneur would have run without the loan.

$$L_i^B \leq \sigma + \hat{\alpha}_i^B \sigma \quad (16)$$

At time $t+1$ the cost of the loan is $(1+r+z)L_i^B$ and the discounted value of the collateral, taking into account depreciation, is $c(\sigma + \hat{\alpha}_i^B \sigma)(1+r)$. To prevent defaults the informal lender

⁶Another explanation for this additional cost z could be that lending aside regulations implies a risk of being caught and therefore the extra cost reflects a risk premium.

will collateralise a fraction e , from the starting up cost $\sigma + \widehat{\alpha}_i^B \sigma$, where $e < c$. The fraction e is less than c reflecting the fact that given the illegality of the market, if the entrepreneur defaults, to seize the collateral the lender might incur an additional cost compared to banks, such as that of hiring bodyguards to ensure borrowers give up the collateral.

$$(1 + r + z)L_i^B \leq e(\sigma + \widehat{\alpha}_i^B \sigma)(1 + r) \quad (17)$$

To simplify notation in eq.(17), define the ratio $\frac{1+r}{1+r+z}$ as κ . Then substituting $\sigma + \widehat{\alpha}_i^B \sigma$ equal to w_i from eq.(3) and solving for L_i^B ,

$$L_i^B = ew_i \kappa \quad (18)$$

With the loan L_i^B , the entrepreneur will run the project at a scale α_i^B , which is greater than $\widehat{\alpha}_i^B$, the scale of the project the entrepreneur would have run in sector B without a loan.

$$w_i + L_i^B = \sigma + \alpha_i^B \sigma \quad (19)$$

Substituting L_i^B , then the optimal scale α_i^B is expressed as follows,

$$\alpha_i^B = \frac{w_i(1 + e\kappa) - \sigma}{\sigma} \quad (20)$$

The lender in the informal sector will advance loans to entrepreneurs who provide a positive collateral, regardless of the scale α_i^B .

The net returns η_i^B of entrepreneurs with loan L_i^B in sector B are determined by the net returns of the scale of production $\alpha_i^B \rho$ plus the value of the depreciated machinery $c(\sigma + \alpha_i^B \sigma)(1 + r)$, minus the salary bill $\alpha_i^B s^B$ and the repayment of credit $(1 + r + z)L_i^B$.

$$\eta_i^B = \alpha_i^B \rho + c(\sigma + \alpha_i^B \sigma)(1 + r) - \alpha_i^B s^B - (1 + r + z)L_i^B \quad (21)$$

The individual will prefer to become an entrepreneur in sector B than a worker, if the project yields returns higher than the returns earned by a worker.

$$\eta_i^B \geq w_i(1 + r) + s^B \quad (22)$$

Substituting η_i^B eq.(21), α_i^B eq.(20) and L_i^B eq.(18) into eq.(22), to obtain returns η_i^B higher than workers' return, the initial wealth w_i has to be equal to or greater than a wealth

threshold denoted by w^B in eq.(23).

$$w_i \geq \frac{\rho}{V - \frac{(1+e\kappa)s^B}{\sigma}} = w^B \quad (23)$$

where $V = (\frac{1+e\kappa}{\sigma})\rho + (1+r)[c - e + ce\kappa - 1]$.

Another way to interpret eq.(22) is by solving for the maximum salary \bar{s}_i^B that the entrepreneur i can afford to pay in sector B to get higher returns as an entrepreneur than as a worker.

$$s_i^B \leq \frac{(w_i V - \rho)\sigma}{(1+e\kappa)w_i} = \bar{s}_i^B \quad (24)$$

An additional consideration for potential entrepreneurs is the comparison of entrepreneurial returns in sector A and in sector B . Entrepreneurs will prefer to set up business in sector A only if they get equal or higher returns, hence only if $\eta_i^A \geq \eta_i^B$.

Substituting α_i^A , eq.(2) and L_i^A , eq.(7) into η_i^A in eq.(11) then substituting α_i^B , eq.(3) and L_i^B , eq.(18) into η_i^B in eq.(21) and solving for the initial wealth w_i , the entrepreneur will get higher profits in sector A , if the initial wealth is greater than or equal to a wealth threshold denoted by w^{AB} in eq.(25) ,

$$w_i \geq \frac{H - s^A[\frac{F(1+c)}{\sigma} + \tau^s]}{M + s^A(\frac{e\kappa - c - \tau^s - \tau^s e\kappa}{\sigma})} = w^{AB} \quad (25)$$

where $H = F\{\frac{(1+c)}{\sigma}[\rho(1 - \tau^\rho) + \sigma c(1+r)] - (1+r)c\} - \rho\tau^\rho$ and $M = (\frac{1+c}{\sigma})(\rho(1 - \tau^\rho) + (1+r)c\sigma) - (\frac{1+e\kappa}{\sigma})(\rho + c\sigma(1+r)) - (1+r)(c - e)$.

The threshold of wealth w^{AB} in eq.(25) shows that if the registration fee F and the tax rates τ^ρ and τ^s are equal to zero, entrepreneurs will always get higher profits in sector A . Depending on the parameters of the economy, w^{AB} could be such that for no individual in the economy is it profitable to become an entrepreneur in sector A , for instance if taxes on entrepreneurial returns are such that entrepreneurs get no profits or lower profits than if they were in another sector or occupation.

The order of the different wealth thresholds in this occupational choice, w^A , w^B and w^{AB} , will depend upon the different parameters of the economy. It is more likely that in an economy the highest wealth threshold will be w^{AB} , since it would require that only the wealthiest entrepreneurs make more profits in sector A . The second highest wealth threshold is likely to be w^B , since individuals that were not as wealthy as to set up business in sector

A still have enough wealth to set up a small or medium size business in sector B and still earn more returns than workers. Then, the lowest wealth threshold is likely to be w^B , since people with wealth below w^B do not have enough wealth to set up business in either sector.

If the order of wealth thresholds is the above mentioned, then to become an entrepreneur in sector B the initial wealth must lie in the range $w^B \leq w_i < w^{AB}$ and to become an entrepreneur in sector A , initial wealth must lie in the range $w_i \geq w^{AB}$.

3.2 Labor Market

It is assumed that workers supply inelastically one unit of labor and they can work for as many firms as they want as long as the total supply of labor does not exceed their labor unit endowment. The total demand for labor in sector A is denoted by D^A and is given by n^A number of firms. Equally, the total demand for labor in sector B is denoted by D^B and given by n^B number of firms.

The demand for labor D_i^A of entrepreneur i , will be equal to zero if the salary s^A is greater than \bar{s}_i^A , the maximum salary an entrepreneur can afford to pay in sector A . If the salary s^A is less than or equal to \bar{s}_i^A the demand for labor will be equal to the scale of production α_i^A . The total demand for labor in sector A , D^A will be equal to the sum of the individual demands of those entrepreneurs whose initial wealth is greater than w^{AB} .

$$D^A = \left\{ \begin{array}{ll} 0 & \text{if } s^A > \bar{s}_i^A \\ \int_{w^{AB}}^{\bar{w}} \alpha_i^A(w) W_t'(w) dw & \text{if } s^A \leq \bar{s}_i^A \end{array} \right\} \quad (26)$$

where $W_t'(w)$ is the density function of the wealth distribution at time t and \bar{w} is the maximum level of wealth in the economy.

The demand D_i^B of entrepreneur i is zero if the salary s^B , is higher than the maximum salary an entrepreneur can afford to pay in sector B , \bar{s}_i^B . It is assumed that in sector B the maximum net salary offered s^B is equal to the exit option of workers which is the net salary in sector A . If the salary s^B is less than or equal to $s^A(1 - \tau^s)$ the total demand D^B for labor will be the sum of the individual demands for labor of the entrepreneurs with wealth greater than w^B but lower than w^{AB} . Entrepreneurs with initial wealth higher than w^{AB} obtain more profits in sector A , and individuals with initial wealth lower than w^B will obtain higher profits if they become workers.

$$D^B = \begin{cases} 0 & \text{if } s^B > \bar{s}_i^B \\ \int_{w^B}^{w^{AB}} \alpha_i^B W_t'(w) dw & \text{if } s^B \leq \bar{s}_i^B \end{cases} \quad (27)$$

The total supply of labor denoted by Z , will take the value of zero if the current net salary is less than γ , the return to being unemployed. The supply will be positive as long as the net salary in either sector is equal to or higher than γ , and everyone in the economy will be willing to supply labor if the salaries are above \bar{s}_j , which is the highest salary that the wealthiest entrepreneur in the economy is willing to pay.

$$Z = \begin{cases} 0 & \text{if } s^A(1 - \tau^s) < \gamma \\ \int_{\underline{w}}^{w^B} W_t'(w) dw & \text{if } s^A(1 - \tau^s) \geq \gamma \end{cases} \quad (28)$$

where $W_t'(w)$ is the density function of the wealth distribution and N is the population in the economy.

As the following figure shows, in equilibrium $s^B = s^A(1 - \tau^s)$. As long as there are entrepreneurs in sector A , the salary paid in sector B will be equal to the net salary paid in sector A . If there are no entrepreneurs in sector A , the prevailing salary in sector B will be determined purely by the demand D^B and the labor supply Z . In order to guarantee that there will be individuals willing to become workers in either of the sectors the following assumptions are introduced

Assumption 1 *The returns of unemployed individuals γ are strictly smaller than the maximum equilibrium salary entrepreneurs are willing to offer, \bar{s}_i^A or \bar{s}_i^B .*

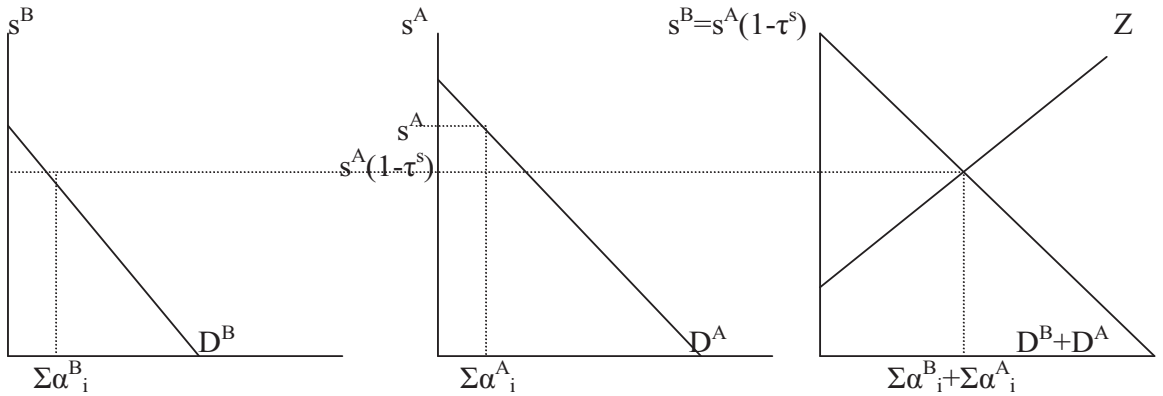


Figure 1: The Labour Market

Assumption 2 *If the demand for and supply of labor are equal the equilibrium salary will make individuals indifferent between becoming workers and entrepreneurs.*

If it is the case that the supply of labor is greater than the demand for labor, $Z \geq D^A + D^B$ there is an excess supply at the current salary. Given that the salary s_t^A and hence s_t^B are endogenously determined in the model, the salaries will adjust to reduce or clear if possible, the excess supply or demand for labor.

3.3 Dynamics of Distribution of Wealth

The idea of the dynamics of wealth distribution in this model is that if we were to measure the change in the number of individuals in each occupation and the evolution of their earnings, we could analyze how these two look given an initial distribution of wealth $W_t(w)$ and then follow the dynasties of each individual.

One can analyze the distribution of wealth of dynasties by looking at the distribution of bequests. The bequest or the future starting wealth of the offspring of unemployed individuals is a proportion β of the sum of savings $w_{u,t}(1+r)$ and the subsistence earning γ .

$$w_{u,t+1} = \beta\{w_{u,t}(1+r) + \gamma\} \quad (29)$$

The future starting wealth of the offspring of workers in either sector A or B is denoted by $w_{w,t+1}$ which is the proportion β of the sum of savings $w_{w,t}(1+r)$ and the net salaries obtained in either of the sectors $s_t^A(1-\tau^s) = s^B$, and provided that provided that $w_t \leq w^B$.

$$w_{w,t+1} = \beta\{w_{w,t}(1+r) + s_t^A(1-\tau^s)\} \quad (30)$$

The future starting wealth of the offspring of entrepreneurs in sector A is the proportion β of the returns of being an entrepreneur in sector A , η^A .

$$w_{e,t+1}^A = \beta\{\alpha_t^A(\rho(1-\tau^\rho) - s_t^A) + c(\sigma + \alpha_t^A\sigma)(1+r) - (1+r)L_t^A\} \quad (31)$$

The future starting wealth of the offspring of an entrepreneur in sector B is the proportion β of the returns of being an entrepreneur in sector B , η^B .

$$w_{e,t+1}^B = \beta\{\alpha_t^B(\rho - s_t^B) + c(\sigma + \alpha_t^B\sigma)(1+r) - (1+r+z)L_t^B\} \quad (32)$$

Substituting the values of α_t^A , L_t^A , α_t^B and L_t^B in eq.(31) and eq.(32) respectively, the future starting wealth of the offspring of entrepreneurs in sector A is,

$$w_{e,t+1}^A = \beta\{w_t\mu_t - \delta_t\} \quad (33)$$

$$\mu_t = \left(\frac{1+c}{\sigma}\right)(\rho(1 - \tau^\rho) - s_t^A + c\sigma(1 + r)) - (1 + r)c$$

$$\delta_t = F\left[\left(\frac{1+c}{\sigma}\right)(\rho(1 - \tau^\rho) - s_t^A + c\sigma(1 + r)) - (1 + r)c\right] + (\rho(1 - \tau^\rho) - s_t^A).$$

The term $\mu_t > 0$ and $\delta_t > 0$, since $s_t^A < \rho(1 - \tau^\rho)$ given that the salaries are endogenously determined to have a positive return in sector A and that $c > 0$. In addition, $w_t \geq w^{AB}$.

The future starting wealth of the offspring of entrepreneurs in sector B is,

$$w_{e,t+1}^B = \beta\{w_t\widehat{\mu}_t - \widehat{\delta}_t\} \quad (34)$$

$$\widehat{\mu}_t = \left(\frac{1+e\kappa}{\sigma}\right)(\rho - s_t^B) + (1 + r)(c - e + ce\kappa)$$

$$\widehat{\delta}_t = \rho - s_t^B.$$

The term $\widehat{\mu}_t > 0$ and $\widehat{\delta}_t > 0$, since $s_t^A(1 - \tau^s) \geq s_t^B < \rho$ to have a positive return in sector B , and it has been assumed that $c > e$, where $c > 0$ and $e > 0$. In addition, $w^{AB} > w_t \geq w^B$.

The problem in analyzing the dynamics of wealth in this model is to find whether the wealth of the dynasties in each occupation converges to a stable fixed point or whether wealth continuously expands over time. If we think of the relationship between w_t and w_{t+1} as $w_{t+1} = f(w_t)$ then the steady state occurs when w_{t+1} equals w_t , and we can call their common value \widehat{w} . It is known that the conditions for the stability of the dynamics of linear recurrence relations with **constant coefficients** can be expressed by the following *Theorem 1*.

Theorem 1 *If \widehat{w} is a fixed point of the first order recurrence equation $w_{t+1} = f(w_t) = R w_t + a$, then \widehat{w} is a stable fixed point if $-1 < R < 1$ and an unstable fixed point if $R > 1$.*

Theorem 1 Proof. *Let v_t be the difference between w_t and \widehat{w} . Then $v_t = w_t - \widehat{w}$ and $v_{t+1} = w_{t+1} - \widehat{w} = f(w_t) - \widehat{w} = f(\widehat{w} + v_t) - \widehat{w}$.*

By Taylor's theorem it follows that

$$v_{t+1} \approx f(\widehat{w}) + f'(\widehat{w})v_t - \widehat{w}.$$

But \widehat{w} is a fixed point so $\widehat{w} = f(\widehat{w})$ and $f'(\widehat{w}) = R$. Thus $v_{t+1} \approx R v_t$.

Since R is a constant, the error v_t decays to zero if $-1 < R < 1$. However if $R > 1$ the error v_t continuously increases. ■

To find the fixed points of earnings by occupation in this model, w_{t+1} is set equal to w_t . This process is carried out in the recursive relationship for the unemployed eq.(29), for workers, eq.(30), for entrepreneurs in sector A , eq.(33) and for entrepreneurs in sector B , in eq.(34). The fixed point for the unemployed is \widehat{w}_u , for the workers \widehat{w}_w , for the entrepreneurs in sector A \widehat{w}_e^A and \widehat{w}_e^B for the entrepreneurs in sector B .

$$\widehat{w}_u = \frac{\beta\gamma}{1 - \beta(1+r)} \quad (35)$$

$$\widehat{w}_w = \frac{\beta s^A(1 - \tau^s)}{1 - \beta(1+r)} \quad (36)$$

$$\widehat{w}_e^A = \frac{-\beta\delta_t}{1 - \beta\mu_t} \quad (37)$$

$$\widehat{w}_e^B = \frac{-\beta\widehat{\delta}_t}{1 - \beta\widehat{\mu}_t} \quad (38)$$

Note that the fixed points \widehat{w}_u , \widehat{w}_w , \widehat{w}_e^A or \widehat{w}_e^B depend on the salaries, which depend on the distribution of wealth. Therefore the coefficients associated to how wealth changes over time, i.e. $(w_{i,t+1} - w_{i,t})$, are not constant and Theorem 1 therefore cannot be applied directly to analyze the dynamics of wealth distribution in this model. Therefore depending on the initial distribution of wealth W_t , the long-run wealth distribution could converge to numerous different types of long-run distributions. Nevertheless, under certain assumptions, Theorem 1 can shed light on the type of stability that each of the difference equations in this model can have.

Assumption 3 *The returns of individuals as entrepreneurs in either sector A or B are strictly higher than the returns they would earn being unemployed $\delta_t > \gamma$, $\widehat{\delta}_t > \gamma$, $\mu_t > (1+r)$, $\widehat{\mu}_t > (1+r)$, $\delta_t > 0$ and $\widehat{\delta}_t > 0$.*

Assumption 4 *A dynasty cannot become rich over time just by saving a fraction of its wealth $\beta(1+r)w_{i,t}$. Therefore it will be assumed that $\beta(1+r) < 1$.*

Assumption 5 *A dynasty of unemployed people cannot accumulate wealth over time sufficient to set up business in either sector A or B . Therefore, it is assumed that $\frac{\beta\gamma}{1-\beta(1+r)} < w^B$ $\frac{\beta\gamma}{1-\beta(1+r)} < w^{AB}$.*

Assumption 6 *The fixed point for workers' earnings lies below the wealth threshold to*

set up a business $w^B > \widehat{w}_w$ and $w^{AB} > \widehat{w}_w$, if salaries remain low over time $s_t^A(1 - \tau^s) = \gamma$.

Assumption 7 *The fixed points for entrepreneurs' earnings in both sectors A and B are equal to or greater than the wealth threshold to set up business, $w^B \leq \widehat{w}_e^B$ and $w^{AB} \leq \widehat{w}_e^A$.*

If assumptions 1-6 are fulfilled then the long-run distribution for each occupation can be characterized in the following propositions.

Proposition 1 *The long-run earnings of workers and the unemployed are positive and stable if $\beta(1 + r) < 1$.*

Proof. *The earnings of unemployed and workers at time $t + 1$ are given by eq.(35) and eq.(36) respectively. Since the numerators of these equations are positive $\beta\gamma > 0$ and $\beta s^A(1 - \tau^s) > 0$, to have positive long-run earnings, the denominator $1 - \beta(1 + r)$ must also be positive, therefore $\beta(1 + r) < 1$. Since the recurrence relation between w_{t+1} and w_t decays to zero only if $-1 < \beta(1 + r) < 1$, thus the fixed points are positive and stable if $\beta(1 + r) < 1$.*

■

Proposition 2 *The long-run earnings of entrepreneurs in sector A have a positive but unstable fixed point if $\beta\mu_t > 1$.*

Proof. The long-run earnings for entrepreneurs in sector A are given by eq.(37). The numerator of this equation is negative since $-\beta\delta_t < 0$, $\beta > 0$ and $\delta_t > 0$. To secure a positive fixed point the denominator must also be negative which implies $\beta\mu_t > 1$, hence the fixed point for formal entrepreneur earnings is unstable. ■

Proposition 3 *The fixed point for earnings of entrepreneurs in sector B is positive, but unstable if $\beta\widehat{\mu}_t > 1$.*

Proof. The long-run earnings for entrepreneurs in sector B are given by eq.(38). The numerator of this equation, $-\beta\widehat{\delta}_t < 0$ since $\beta > 0$ and $\widehat{\delta}_t > 0$. To secure a positive fixed point the denominator must also be negative, which implies $\beta\widehat{\mu}_t > 1$, hence the fixed point for entrepreneurs' earnings in sector B is unstable. ■

Propositions 1, 2 and 3 guarantee that the fixed points are positive and analyze their stability. In summary, those whose initial wealth is below $w_i < \widehat{w}_e^B$ will become workers and their wealth will converge to $\widehat{w}_w = \widehat{w}_u$, that is assuming that $s^A(1 - \tau^s) = \gamma$. For entrepreneurs with initial wealth is in the range $w^B \leq w_i < \widehat{w}_e^B$ being an entrepreneur in the informal sector represents merely a temporal occupation, as the United Nations has explained. In contrast for those entrepreneurs with wealth in the range $w_i > \widehat{w}_e^B$ the informal sector represents a permanent occupation and these entrepreneurs will manage to expand

their scales of production and eventually formalize their businesses.

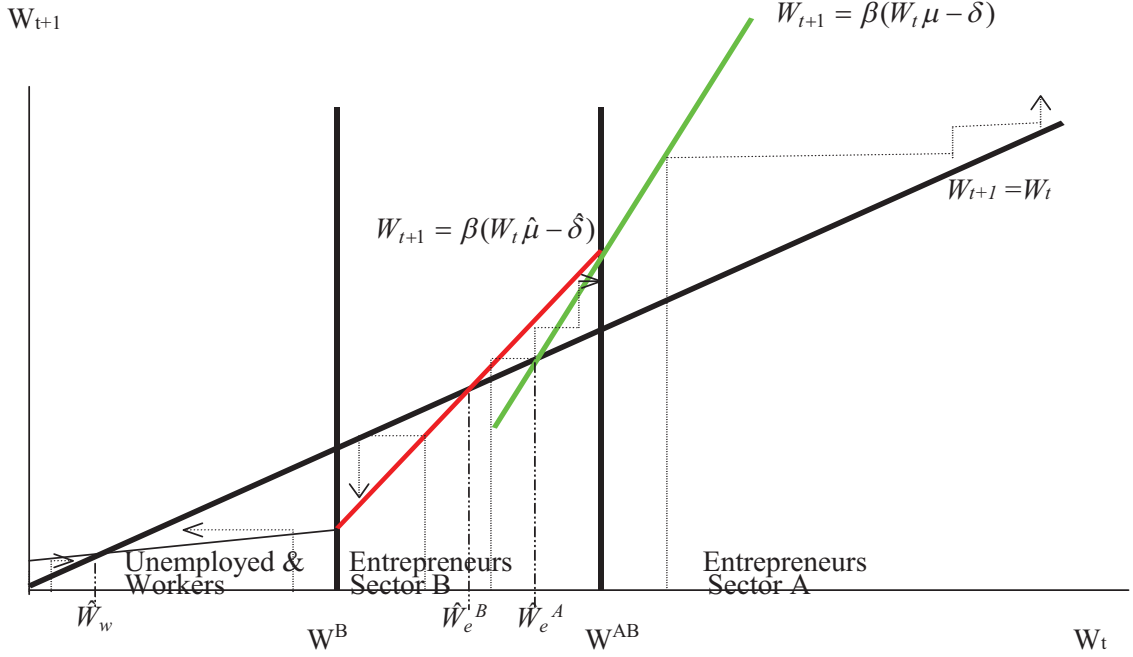


Figure 2: Dynamics of Wealth (Assuming $s^A(1 - \tau^s) = \gamma$)

Given that people whose wealth $w_i > \hat{w}_e^B$ expand their scale of production over time, this means that every period the offspring of these generations will receive larger bequests and will therefore further expand their scales of production, raising labor demand. Eventually this will lead to an excess demand for labor. If this is the case, the salaries will increase and for certain individuals that were entrepreneurs in the past, the rise in salaries will make it more attractive to become workers.

Then, the increase in salaries will decrease the incentives to become an entrepreneur in either of the sectors. This will be reflected in a higher wealth threshold at which it is more profitable to become an entrepreneur. This will also affect the profitability of entrepreneurship, and fewer people will choose to remain entrepreneurs.

The dynamics of wealth will not follow a linear path, but the one presented in figure (3). The wealth of entrepreneurs in sector *A* will increase towards ∞ . This is denoted by the left curve in figure (3). Similarly, the wealth of the middle sized entrepreneurs in sector *B* will continuously expand, such that it becomes more profitable, at some point, for them to formalize their business and switch to sector *A*. Therefore their wealth will increase towards ∞ . However, the wealth of small entrepreneurs in sector *B* will decrease towards $-\infty$. The

scale of production of these entrepreneurs is too small to keep financing the cost of replacing the depreciated machinery. Therefore, at some point they will be forced to shut down their business and either become workers or unemployed. The dynamics of wealth for entrepreneurs in sector B are denoted by the right curve in figure (3).

There is growth in this economy as long as there are middle size entrepreneurs in sector B that can keep expanding their scale of production. Growth will stop once the expansion of scales of production is constrained by the labor capacity of the economy. Hence, the wealth of entrepreneurs instead of tending towards ∞ will be constrained by an upper threshold (\bar{w}^A), given the economy's labor capacity. This upper threshold \bar{w}^A is denoted by the dot in the upper curve that intersects the 45° line in figure (3).

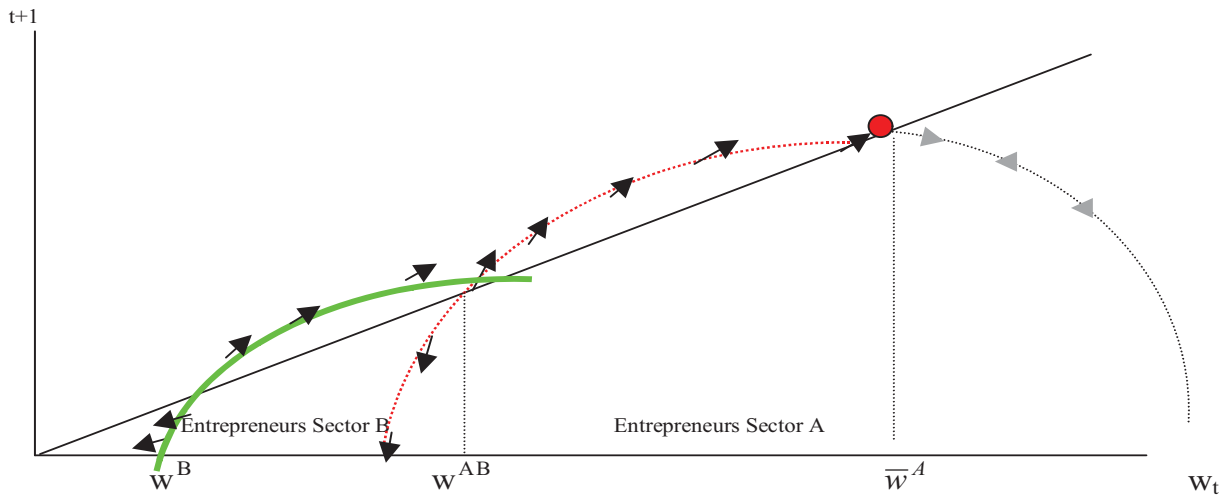


Figure 3: Long Run Stationary Distribution

Figure (3) shows that some entrepreneurs in sector A and B keep expanding their scales of production, while others in sector B will shut down their business. However, it is not certain what happens to the dynamics of salaries. This depends on the distribution of wealth, how wealthy the entrepreneurs are in each sector and the inequality among entrepreneurs.

Proposition 8 *The initial distribution of wealth converges to either of four types of stationary distributions*

Stationary Distribution Type 1: The informal market disappears and the formal economy prevails. This stationary distribution is achieved when all entrepreneurs in sector B either expand their scales of production such that they formalize their business or they become workers.

Stationary Distribution Type 2: The informal and the formal economy coexist. This type of distribution can be achieved if the entrepreneurs in sectors A and B are hiring all the population and no entrepreneur can increase salaries in order to expand his scale of production.

Stationary Distribution Type 3: The only existing sector is the informal market. This stationary distribution is achieved when entrepreneurs in sector B can no longer expand their scales of production by increasing salaries. Since full employment is reached there is no possibility for them to expand and formalize their business.

Stationary Distribution Type 4: Full unemployment prevails. This stationary distribution will be achieved if the initial wealth of all individuals is below the wealth threshold necessary for setting up a business in either sector at the lowest possible salary γ , $w_{i,t} < \hat{w}_e^B$. This stationary distribution can also be reached, if all the entrepreneurs in the economy are engaged in sector B and their scales of production are so small that in the long-run all firms shut down.

The dynamics of inequality depend on what happens to unemployment and salaries over time. The larger is the proportion of wealthy people, understood as those who can set up a business or remain in business, the more jobs will be available and the higher the salaries will be. Salaries will increase faster, the less poor the non-entrepreneurs are. This is because the reservation salary is not given only by γ , but actually by the maximum salary that every individual is willing to pay.

4 Numerical Examples

Four numerical examples are presented to illustrate the types of long-run stationary distributions the model can reach. The objective of the numerical analysis is to contrast the effects that both initial wealth distribution and the economy parameters can have in the long-run. Three initial wealth distributions are used. The first two distributions (defined as type I and type II) are depicted in figure(4). The third distribution used has the same aggregate wealth as the second distribution but has marginally less inequality in wealth. Table (2) provides the parameters used through the examples presented, while table (3) shows the main statistics of the initial and long-run distribution of each of the four examples presented.

N	β	σ	r	ρ	F	c	e	z	τ^P	τ^S
100	0.1	3.5	0.2	110	0.421	0.4	0.3	0.3	0.01	0.01

Table 2: Initial Parameters of Examples I-IV

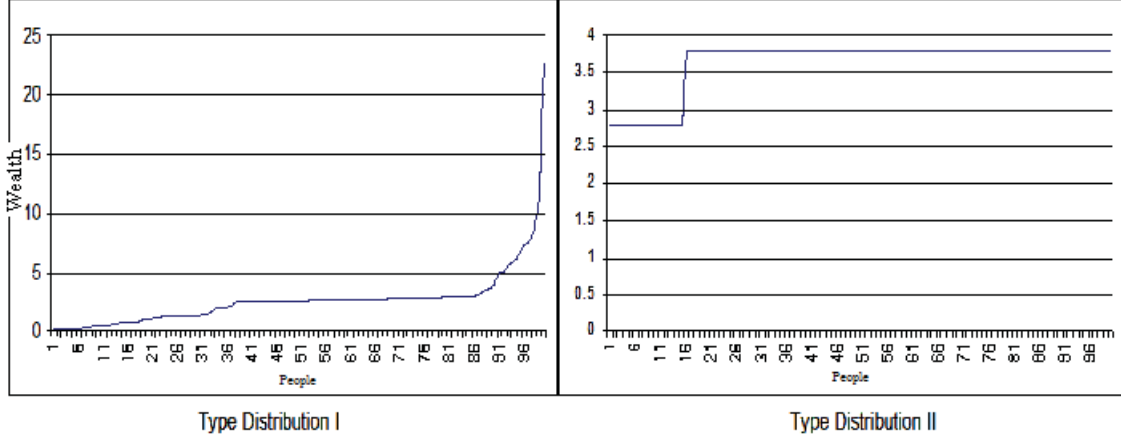


Figure 4: Initial Wealth Distributions Used in Examples I-III

Example I (Formal economy prevails with two entrepreneurs)

A first example is presented using the initial wealth distribution type I, the parameters shown in table (2) and setting the parameter $\gamma = 0.05$ (what unemployed people can produce with their labor to remain in subsistence). In the first period there are 4 entrepreneurs in sector A and 8 entrepreneurs in sector B . During the first six periods the entrepreneurs increase their scale of production without raising salaries. The wider gap between the earnings of entrepreneurs and workers increased inequality (as measured by the Gini coefficient). At time $t = 7$ there is an excess of demand for labor that leads to an increase in salaries such that the labor market clears. With the rise in salaries only two entrepreneurs in sector A remain in business, the rest preferring to become workers. Inequality decreases given the rise in salaries and from this time onwards inequality settles down to 0.2.

In the long-run the level of inequality is reduced considerably, as shown in figure (5). Moreover, the dynamics of inequality in this example followed an inverted-U pattern.

The economy converges to a state with two entrepreneurs in sector A that hire the rest of the population. These two entrepreneurs keep salaries high, given that if one of them lowers salaries workers will prefer to work for the other entrepreneur who still offers higher salaries. Therefore, these two entrepreneurs keep the competition in the labor market.

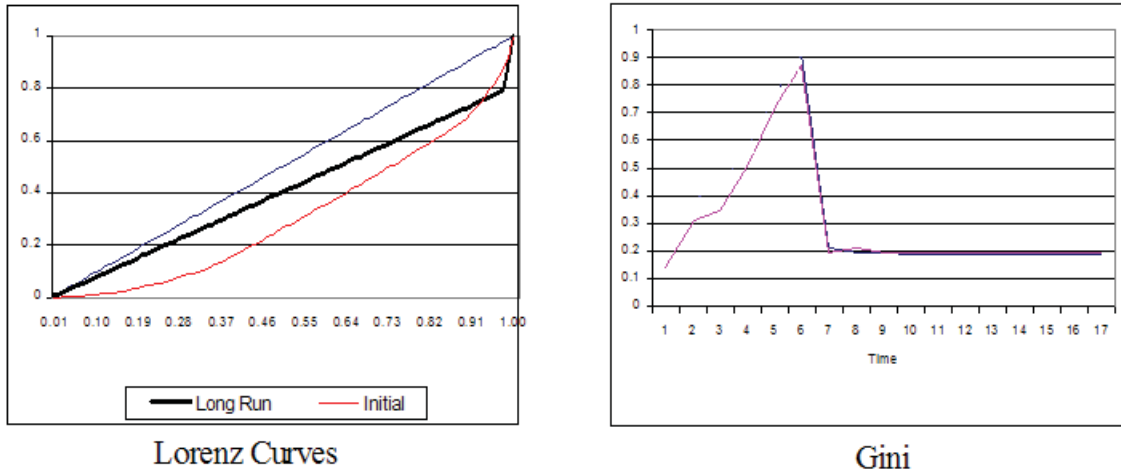


Figure 5: Dynamics of Inequality in Example I

Example II (Unemployment prevails)

A second example is presented using the wealth distribution I, the parameters of table (2) but the value of γ is increased in this example to 3.5. In the initial wealth distribution assumed none of the individuals in the economy have wealth above the threshold where it is more profitable to set up a business in the formal sector (w^{AB}). Nonetheless, there are 60 individuals with wealth above the wealth threshold necessary to set up business in the informal sector w^B , but below the fixed point \hat{w}_e^B . Therefore in the long-run it is expected that even if salaries remain constant, entrepreneurs will decrease their scales of production until they eventually shut down their firms. In this example, in the second period 57 entrepreneurs shut down their businesses and only 3 remain with wealth above the required to set up business in the informal sector (w^B). By the third period the wealth of the 3 remaining entrepreneurs decreases further and none of them can set up business. Therefore everyone in the economy remains unemployed. The total aggregate wealth in the long-run is considerably lower than in the first example presented. In the long-run the wealth distribution reaches perfect equality, given that everyone is unemployed earning the same level of γ . Therefore in the long-run inequality declines to zero.

Example III (Formal and informal markets coexist)

A third example is presented using the initial wealth distribution type II, the parameters of table (2) and $\gamma = 3.5$. In the first period there are 2 individuals that set up business in sector A and 83 individuals that prefer to set up business in sector B. In the second period the 85 entrepreneurs in the economy (in sectors A and B) do not increase their labor demand

considerably and therefore salaries do not rise. In the third period, all entrepreneurs raise their demand for labor, but full employment is reached and entrepreneurs in sector A cannot raise salaries to entice entrepreneurs in sector B into becoming workers. Since entrepreneurs in either of the two sectors have to increase salaries to increase their scales of production, it is not possible for any entrepreneur to continue expanding his scale of production indefinitely and therefore growth stops.

In the long-run 2 entrepreneurs remain in sector A and 83 in sector B . Note that the total wealth decreases over time given that the wealth of the workers declined from 2.8 in the original distribution to 0.40 in the long-run distribution and that no entrepreneur could expand considerably his scale of production.

Inequality increased during the first two periods given that entrepreneurs expanded their scales of production. By the third period when no one is able to expand further inequality remains constant over time. Hence, in the long-run distribution is more unequal than on the original distribution.

Example IV (informal market prevails)

This example uses the same parameters as the previous example. However, the initial distribution type II is modified such that the two richest individuals that were able to set up formal businesses in example III have the same wealth as each of the 83 entrepreneurs in the informal sector. In this example at the initial wealth distribution 85 individuals have wealth above the wealth threshold w^B required to set up business in the informal market. Their wealth is below the wealth threshold w^{AB} and above the fixed point \widehat{w}_e^B . In the initial period, these 85 individuals set up business in the informal sector. By the second period all entrepreneurs increase their wealth, but not enough to reach the wealth threshold w^{AB} . By the third period although the wealth of the 85 entrepreneurs increased, none of the entrepreneurs can afford to increase salaries to the level required for some entrepreneurs to prefer to become workers. Given that there is perfect equality among entrepreneurs, no one can increase his scale of production any further and hence entrepreneurs in the long-run keep operating at the scale of production at which they operated in the third period. Therefore in the long-run all the entrepreneurs in the economy operate in the informal market.

The aggregate wealth level decreases with respect to the initial wealth distribution. This is because each entrepreneur is operating at a very small scale and the wealth of workers decreases from 2.8 in the original distribution to 0.40 in the long-run. Inequality increased

during the first two periods. From the third period onwards when no entrepreneur can expand further inequality remain constant. In the long-run the wealth distribution is more unequal than the initial wealth distribution.

	Initial Wealth Distribution					Long-Run Wealth Distribution							
	Total Wealth	Mean	Min	Max	Gini	n^A	n^B	s^A	Total Wealth	Mean	Min	Max	Gini
Example I	273	2.7	0.2	22.6	0.380	0	2	84.4	1,176	11.8	9.42	123	0.19
Example II	273	2.7	0.2	22.6	0.380	0	0	0.0	40	0.4	0.4	0.4	0.00
Example III	366	3.7	2.8	4.0	0.036	83	2	3.5	330	3.3	0.4	4.0	0.13
Example IV	366	3.7	2.8	3.8	0.035	85	0	3.5	330	3.3	0.4	3.8	0.13

Table 3: Initial and Long Run Wealth Distributions of Examples I-IV

4.1 Discussion of the Model

In the model, due to fixed costs and credit constraints, not all individuals are able to set up business in the formal sector. With a large initial wealth inequality, the fewer entrepreneurs there are, the smaller the demand for labor and the lower the competitive salaries will be. In these conditions, there will be individuals that would prefer to operate in the informal economy in order to receive a higher return than the one they would obtain in the formal sector.

A novelty of this paper is the role that the initial wealth distribution plays in the existence and evolution of the informal sector over time. The model suggests that the more unequal a country is, the larger the informal sector will be in the short run and the more likely that the economy will converge to a low aggregate income and more unequal wealth distribution in the long-run. Although Chong and Gradstein (2007) argue that income inequality is a particularly significant determinant of informality, the model developed in this paper suggests that the relationship between inequality and informality is in fact endogenous if the dynamics of the informal sector are considered over time. Indeed high initial inequality in wealth causes a large informal sector, however, a large informal sector increases inequality as the difference between the rents of those engaged in the formal and informal sectors deepens over time. This is one of the possible reasons why Latin America, the Sub-Saharan Africa and even the United States have both high levels of inequality and large informal sectors.

A distinctive aspect of the paper is that by allowing entrepreneurs to adjust their scale of production (in the formal or informal sector) over time, the model predicts that firms of smaller scale would be set up in the informal sector. By allowing entrepreneurs to vary

their scale of production the dynamics of the model depend not only on who is not credit constrained (as in other similar occupation choice models e.g. Banerjee and Newman, 1993), but on the inequality among those who are credit and non-credit constrained. In particular, the wealth difference among entrepreneurs matters. The larger the difference between entrepreneurs the more likely that low salaries will be paid, since the largest entrepreneurs can drive out small entrepreneurs by simply increasing salaries marginally. In the long-run the model can converge to four distinct types of equilibria depending on the parameters of the economy and the initial wealth distribution: formal sector prevails, or informal sector prevails, or a combination of both formal and informal sector prevails, or only unemployment prevails.

Regarding the labor market, assuming that all potential workers have the same skills, the model finds that in equilibrium the net monetary salary is the same in the informal and formal sectors. Despite this, workers might prefer to be hired in the formal sector as paying taxes gives access to health insurance, pension and other publicly provided goods. Had the model introduced heterogeneity in skills for instance, the equilibrium net salary would have been different in the two sectors. In that case the formal sector would offer a wage premium to hire the more skilled people thereby increasing the productivity gap between the formal and the informal sectors over time.

The numerical examples presented showed that a bigger formal business might drive a small formal business into informality by increasing salaries. Badaui et. al. (2009) develop a theoretical model and find evidence in Ecuador where the formal sector pays a wage premium to reflect the firm size premium and firms in the informal sector pay lower salaries even when there is no heterogeneity amongst either workers or firms ex ante. The model developed here, suggests that indeed firms might choose to pay a higher salary (or a wage premium) when they are in expansion, however over time, this wage premium will disappear once the firms reach their potential or do not find it profitable to expand their scale of production any further. This might explain the varied empirical evidence where in some countries there are no differences in net salary between the formal and informal sectors (such in Argentina found by Pratap and Quintin, 2006) while in other big firms in the formal sector tend to pay a wage premium.

The model also addresses issues that other papers do not, such as the different borrowing cost between the formal and informal credit markets (such as Turknovsky and Basher, 2009),

and allows both entrepreneurs and workers to migrate between both sectors. Other papers in the literature have used a search-matching framework to model the informal labor market. For instance, Albreach et. al. (2009) and Mortensen and Pissarides (1994) develop a matching model to incorporate a self-employed informal sector where there is heterogeneity in workers' productivity in that more productive workers may opt to wait for a formal sector job, while others may select into the informal sector. The model presented here instead argues that the existence of informality is not per se due to the different workers skills but due to the lack of jobs in the formal economy that is unable to employ people that do not have enough wealth to set up business (in either sector). Hence, the policy recommendation that emerges from this model is that unemployment will rise if policies are aimed solely at penalizing those working in the informal economy. The model instead suggests that in order to accelerate wealth creation and the formalization of the economy policies should be aimed at the creation of firms and their expansion over time.

The informal economy provides an alternative source of employment, creates a more flexible environment for entrepreneurs by not charging high registration fees nor requiring lengthy procedures for gaining access to the credit markets and most importantly it also contributes to the production of goods and services that would otherwise not exist in the economy. For these reasons, it is likely that developing countries will often prefer a “soft policy” against the people working in the informal sector. For instance, the risk involved in working as a street vendor in Mexico, one of the most common types of informal activity, is being caught by the patrol-monitoring group in charge of confiscating the merchandise of street vendors. The merchandise is returned after paying a fine equal to 25 cents of a peso, about .02 cents of a dollar, and completing 5 forms. This fine is completely disproportionate to the registration fee of \$1,109.66 that formal entrepreneurs face and the 7 procedures they need to follow.⁷

Why would a country prefer a “strict policy” against those who are willing to set up business formally? One possible explanation is that in a democracy individuals would prefer to vote for governments that impose high taxes and registration fees if they expect that

⁷This fine does not even cover the cost that the government incurs in running the patrol monitoring, which is at least the \$540 dollars a month paid to the three bureaucrats working in the patrol, and .50 cents of a dollar in providing the forms. This fine is specified not by bureaucrats nor by current politicians, but by a law imposed 40 years ago. The number of procedures to start a business in Mexico varies according to the sector and state. For instance, in addition to the state and county level, there are 2,519 procedures at the federal level.

they will not be liable to pay these taxes, while benefiting from an increase in public services. Hence, in the most unequal countries the poor classes will support a government that imposes a high burden on the formal sector, given that they know that with their current wealth the poor will be unable to become entrepreneurs in the formal sector. These high taxes in addition to putting off entrepreneurs from the formal market, may contribute to tax evasion and other corruption practices. The model presented could be extended to include such cases, where entrepreneurs in the formal sector could choose to avoid taxes and hire workers in the informal market at lower salaries. In the long-run the economy is likely to converge to lower salaries in both sectors and lower aggregate wealth level.

5 Concluding Remarks

This paper developed a dynamic model that sheds light on the evolution of the informal sector and its impact on development. According to the model the informal market exists because of fixed costs and credit constraints, as a result of which not all individuals can afford to enter the formal sector. The advantage of entering into the formal sector is access to credit at a lower interest rate than in the informal sector. Hence, the more unequal the country is, the more unequal the opportunities to become an entrepreneur in the formal sector will be, the lower the salaries and the greater the incentives to become an entrepreneur in the informal sector.

The model shows that the informal sector is composed by a mixture of people. On the one hand, workers who cannot find jobs in the formal sector will prefer to enter the informal economy than stay unemployed. There are also informal entrepreneurs (of smaller scale than in the formal sector) that out of choice decide to become informal entrepreneurs as they will earn more than becoming workers in either sector. Given this bag of mixed circumstances the informal economy is an unstable sector. Very small informal entrepreneurs will be forced to close their businesses over time, even if salaries and interest rates remain constant, as their profits will not be high enough to compensate for the capital that depreciates over time. On the other hand, bigger informal entrepreneurs will find it more profitable to formalize their businesses over time to gain access to lower interest rates in the formal sector. The formal sector is in principle stable, if salaries and other parameters of the economy remain constant. Nonetheless, as the expansion of firms (in the informal and formal sectors) increases the

demand for labor, salaries will rise too. This rise in salaries could push formal small firms out of business or to the informal sector if down-sizing the scale of production makes it more profitable to operate in that sector.

In the long-run the shrinkage of the informal sector depends on the parameters of the economy but crucially on the initial distribution of wealth. The more people that can set up businesses, either in the informal or formal sector, the more likely it is that the economy will converge to a formalized economy in the long-run. However, if there is a high level of inequality, for example with a few big formal businesses and many small (informal and formal) businesses, the economy can converge to steady state where both the formal and informal sector coexist. In this context, big businesses can drive small formal businesses out of the market by increasing salaries. If the big businesses cannot hire all the potential workers in the economy, because they would have to raise wages above the level at which it is profitable for them, the economy will converge to an equilibrium where many small firms (both in the formal or informal sector) and a few big formal businesses coexist.

The model highlights that to foster both development and equality the expansion and formalization of businesses should be encouraged. This could be done with a number of policies. Evidently having a more "business-friendly" environment will help. It might sound simple, but the evidence presented here and elsewhere has shown that reducing red tape such as the cost of entry to business will encourage small firms to formalize their activities. The data shown in this paper suggest that entrepreneurs in developing countries face a much higher cost of entry than those in developed countries. A great proportion of these costs are not charged directly by governments but by notaries which charge more than \$1,000 in some Latin American countries. These high start up fees, could either be abolished or substantially diminished for example by allowing them to be paid over time rather than up front. Tax rates proportional to the size of the scale of production will also help small businesses to gradually expand and formalize their businesses. Less straight forward policies are micro-credits and wealth distribution. The model presented suggests that micro-credits or a one-off Pigou-Dalton wealth transfer towards small entrepreneurs (especially in the informal sector) will not have long-term effects. For these policies to be successful they need to enable small enterprises to stay afloat and increase the size of production over time, not just in the short-run.

References

- [1] Albrecht, J., Navarro, L. and Vroman, S., 2009. "The Effects of Labour Market Policies in an Economy with an Informal Sector." *Economic Journal*. in print.
- [2] Antunes, A.R. and de.V. Cavalcanti, T.V., 2007. "Start up costs, limited enforcement, and the hidden economy." *European Economic Review*. 51 (1): 203-224
- [3] Badaui E. E., Strobl, E. and Walsh, F., 2009. "The formal sector wage premium and firm size." *Journal of Development Economics*. in print.
- [4] Banerjee, A.V. and Newman, A.F., 1993. "Occupational choice and the process of development." *Journal of Political Economy* 101 (2): 274-298.
- [5] Cebula, R.J., 1997. "An empirical analysis of the impact of government tax and auditing policies on the size of the underground economy: the case of the United States, 1973–94." *The American Journal of Economics and Sociology* 56 (1) 73-186.
- [6] Chong, A. and Gradstein M., 2007. "Inequality and Informality." *Journal of Public Economics* 91 (1):159-179.
- [7] De Soto, H., 1989. *The Other Path*, Harper and Row.
- [8] Djankov, S., La Porta, R., Lopez-e-Silanes, F. and Shleifer, A., 2002. "The Regulation of Entry." *Quarterly Journal of Economics* 117 (1):1-37.
- [9] Friedman, E., Johnson, S., Kaufman, D. and Zoido-Lobaton, P., 2000. "Dodging the grabbing hand: The determinants of unofficial activity in 69 countries." *Journal of Public Economics* 76 (3):459-493.
- [10] Fugazza, M. and Jacques, J.F., 2004. "Labor market institutions, taxation and the underground economy." *Journal of Public Economics* 88 (2):395-418.
- [11] Giles, D. E. A. and Tedds, L. M., 2002. "Taxes and the Canadian Underground Economy. Canadian Tax Paper." vol. 106. Canadian Tax Foundation, Toronto.
- [12] Ihrig, J. and Moe, K. S. 2004. "Lurking in the shadows: the informal sector and government policy." *Journal of Development Economics* 73 (2):541-57.

- [13] Johnson, S., Kaufmann, D and Zoido-Lobaton, P., 1997. "The Unofficial Economy in Transition." *The Brookings Review* 2:159-239.
- [14] Johnson, S., Kaufmann, D and Zoido-Lobaton, P., 1998. "Regulatory Discretion and the Unofficial Economy." *American Economic Review* 88(2):387-92.
- [15] Loayza, N. V., 1996. "The Economics of the Informal Sector: A Simple Model and Some Empirical Evidence from Latin America." *Carnegie-Rochester Conference Series on Public Policy*, 45, 129-62.
- [16] Maloney, W., 2004. "Informality Revisited." *World Development*, 32(7):1159-1178.
- [17] Montiel, P., Agénor, P. and Ul Haque, N., 1993. *Informal Financial Markets in Developing Countries*, Blackwell.
- [18] Mortensen, D. and Pissarides, C., 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies* 61 (3):397-415.
- [19] Perry, E. G., Maloney, W., Arias, O., Faynzylber, P., Mason, A. and Saavedra-Chandivi, J., 2007. *Informality Exit and Exclusion*. The World Bank, USA.
- [20] Portes, A., Castells, M., and Benton, L. 1989. "World Underneath: The Origins, Dynamics, and the Effects of the Informal Economy." in *The Informal Economy: Studies in Advanced and Less Developed Countries*. Eds. A. Portes, M. Castells, and L, Benton. Johns Hopking, Baltimore.
- [21] Pratap, S. and Quintin, E., 2006. "Are labor markets segmented in developing countries? A semiparametric approach." *European Economic Review* 50 (7):1817-41.
- [22] Schneider, F. and Enste, D. H., 2000. "Shadow Economies: Size, Causes, and Consequences." *Journal of Economic Literature* 38 (1):77-114.
- [23] Strabu, S., 2005. "Informal sector: The credit market channel." *Journal of Development Economics* 78 (2):299-321.
- [24] Turknovsky, S. J. and Basher, Md. A., 2009. "Fiscal policy and the structure of production in a two-sector developing economy." *Journal of Development Economics* 88 (2):205-216.

[25] World Bank. 2003. "World Development Indicators." The World Bank, Washington, DC.

[26] World Bank. 2003. "Doing Business Database." The World Bank, Washington, DC.