Poverty Traps, Economic Inequality and Delinquent Incentives

Andrés Salazar
Departamento Nacional de Planeación

Edgar Villa
Pontificia Universidad Javeriana

July 15, 2010

Abstract

This paper explores the theoretical linkages between poverty traps, economic inequality in a two sector overlapping generations model under perfect competition in which barriers to skilled educational attainment and delinquent incentives interact. We find that the existence of a poverty trap under high economic inequality and costly indivisible human capital investments generate persistent delinquency in the long run. .We study technological shocks that increase skilled wages or reduces land for the unskilled. These temporal shocks produce an outburst of delinquency in the short run that die out later on. If the shock is permanent then delinquency increases permanently in the long run. Given that the optimal social level of deliquency is zero we study the trade off between deterrence policies and education based policies to reduce long run persistent delinquency. We find that for higher levels of delinquency education based policies yield a lower trade off between increasing subsidies to invest in human capital for unskilled workers while decreasing the apprehension probability by hirign less policemen.

JEL code: I30, J31, K42, O11, O17
Key words: Poverty Traps, Inequality, Delinquency, Human Capital

Introduction

Since the pioneering work of Becker (1968) and Ehrlich (1973) the economics of crime literature has argued that in order to lower crime rates it seems necessary to extend and increase law enforcement policies to deter individuals from choosing delinquent activities or incapacitate them. Even so there has been a growing consensus in this literature that delinquency is not likely to be eliminated completely only through deterrence and incapacitation measures from law enforcement authorities since delinquency is an economic choice that individuals turn to in the face of poverty when lacking other economic opportunities. Other scholars have argued that poverty traps could emerge as a bad equilibrium in environments in which illegal activities feed on legal ones (Mehlum et al. 2002, 2006) depending on the economic returns they face from both activities. Furthermore, some argue that individuals tend to be more likely to choose an illegal organization when young when they live in poverty (Blattman-Miguel 2008) or live with high levels of economic inequality which allows high expected gains[1] from illegal activities (Bourguignon (1999), Fajnzylber et al. 2001, 2002). This

[1] Theoretically net gains from criminal activities have been represented in different ways. For example, Bourguignon (1999) understands them as wealth differences between rich and poor, Imrohoroglu, Merlo and Rupert (2000) considers them as income differences among complex heterogeneous agents while other authors as Kelly (2000) consider income inequality as a measurement of the distance between gains from crime and its opportunity costs.

suggests that we should understand the incentives for an individual to choose a delinquent life in environments where there exists both poverty and high economic inequality for given levels of law enforcement.

Another strand of the economics of crime literature finds evidence that human capital accumulation can weaken delinquent incentives (Lochner (2004, 2010), Lochner-Moretti (2001)). In particular, this literature argues that educational attainment is causally related to higher returns in the labor market as well as positive externalities at the social level which suggests that policies that enhance education opportunities for riskier segments of the population have a positive externality that lowers delinquent incentives.

This paper builds an overlapping generations model similar to Galor-Zeira (1993) under perfect competition that abstracts from unemployment which allows to explore the theoretical linkages between poverty traps, economic inequality and educational attainment. It builds on a dual economy in which delinquents are parasites that prey on legal workers that come from less wealth individuals in a poverty trap. It finds that for given levels of law enforcement measures of deterrence and incapacitation delinquency is persistent in the long run if both wealth inequality and wage differentials are large enough relative to costly indivisible human capital investments. We study then both deterrence and incapacitation policies as well as education based policies to reduce persistent delinquency. We find that even though in the long run these policies may not eliminate completely persistent delinquent levels they can attenuate it. We show that contrary to common intuition education based policies that subsidize human capital investments can increase in the short run delinquent incentives, a short run trade off that policy makers should to be aware of.

This paper is organized in five parts. The first part reviews a strand of literature that links both delinquency to economic inequality and poverty while also reviewing another strand of literature that links education attainment and delinquency. The second part builds up the formal model which explores the theoretical linkages between poverty traps, economic inequality and delinquent incentives. The third part explores comparative dynamics with respect to technological shocks tracing out the effect on delinquency. The fourth part examines policies that can decrease delinquent incentives such as law enforcement and education based policies. The fifth part concludes.

1 Literature review

The modern literature on the economics of crime, based on Gary Becker's seminal (1968) article, has focused on the effect of deterrence and incapacitation as incentives on criminal behavior. This tradition understands crime as a result of individual rational choice where benefits of illegal activities outweigh the costs (punishment in case of apprehension and conviction) as well as their current set of opportunities. As a consequence, deterrence theory research has been predominantly concerned with the isolated effects of the severity and certainty of sanctions on illegal behavior, which has been an argument to extend and increase law enforcement policies in order to reduce crime rates. However, economic and social literature argue that delinquency is not likely to be eliminated completely only through deterrence policies since delinquency is an individual choice in the presence of barriers to enter legal sectors that yield economic opportunities (Eide 1997). Specifically, both poverty and high economic inequality are social conditions that induce illegal behaviors due the lack of other legal ways to acquire incomes and assets. Still human capital accumulation can have a role to play in reducing the delinquent incentives. We now turn to these two relations.

1.1 Poverty, inequality and delinquency

There is a recent literature in economics that studies institutional changes in a dynamic evolutionary framework where inequality and poverty traps emerge with endogenous inefficient institutional arrangements (Bowles 2006). Other theoretical contributions argue that parasitic enterprises can feed on productive businesses (Mehlum, Moene and Torvik (2003, 2006)) where the fraction of parasitic enterprises is determined endogenously depending on the institutional arrangements in which an economy operates, namely depending crucially on legal versus the illegal opportunities they face. This literature understands poverty traps as bad equilibria in a multiple equilibrium environments where inefficient or perverse institutions sustain them and could become persistent in time.

According to Kelly (2000) the link between inequality and crime has been studied by the three main theories of crime: economic theory of crime, social disorganization theory, and strain theory. In the economics crime literature, it has been argued that economic differences have been a necessary condition to keep the incentives to commit crimes, hence, property crime may partly be the consequence of excessive economic inequality (Bourguignon (1999), Fender (1999)). Others have considered the effect of inequality on crime, for example, Ehrlich (1973) uses the fraction of the population in an area earning less than half the median income as a proxy for inequality, and shows that the decision to participate in criminal activities involving material gains is positively associated with income inequality. Witte and Tauchen (1994) examine the impact of earnings on criminal participation and Kelly (2000, pag. 537), using FBI Uniform Crime Reports in US, concludes that "the impact of inequality on violent crime is large, even after controlling for the effects of poverty, race, and family composition". Moreover, some other authors have found evidence of a causal link between income inequality and crime rates using cross country data. For example, Krohn (1976), Soares (2004) and Fajnzylber et al (2002) show that countries with more unequal income distribution tend to have higher crime rates than those with more equal patterns of income distribution, for different samples of countries. Another study finds that a one-point rise in a country's Gini coefficient is associated with nearly a one-point increase in its homicide rate (UN Global Report on Crime and Justice (1999) quoted in Buvinic and Morrison (2000)).

The social disorganization theory emphasizes that the existence of several factors such as poverty, family stability, residential mobility and ethnic heterogeneity push some members of communities to commit crimes and weakens the social control of this behavior (Shaw and McKay (1942)). This theory conjectures that income inequality causes crime in an indirect way due to the fact that inequality is related with poverty and this factor induces more likely individuals to commit illegal acts.

Finally, strain theory based on Merton's (1938) work developed the idea of *anomie*, as the lack of social norms or the failure of a social structure to provide mechanisms and pathways necessary for people to achieve their goals, generating deviant behaviors such as crime. In this theory individual alienation can arise from income inequality, and are also related with other measures of deprivation such as poverty and unemployment. This idea is related with the argument that criminality is based on an individual process that consists of an assessment of economic incentives and social norms (Weibull and Villa (2005)).

However, the relationship between income inequality and crime rates and violence is not completely straightforward. Some countries have decreasing income inequality accompanied by an increase in violence (measured in homicide rates) such as Brazil and Venezuela, or a decrease in homicide rates accompanied by an increase in income inequality such as Costa Rica and Mexico (Morrison, Buvinic, and Shifter (2003)). Moreover, for an specific sample in U.S., income inequality has no significant effects on property crimes such as robbery, burglary and vehicle theft (Allen (1996)).[2]

1.2 Education and delinquent incentives

Complementarily, the economics of crime literature also has found evidence that human capital accumulation can discourage illegal activities. For example, Freeman (1996) shows that educational attainment is a preventive policy for crime and finds an inverse relationship between these two variables. Tauchen, et al (1994) studied a sample of men who attended school relative to those who did not and find a negative relationship between the act of studying or working with the probability of committing criminal acts. They argue that studying and working are associated with greater participation in legal activities and therefore decrease the incentives to commit illegal acts. Lochner and Moretti (2001) also show that there is an inverse relationship between school attainment and crime rates. They find that youths that finish high school are more likely no to enter in delinquent activities. Moreover, they argue that education has a positive externality in reducing crime. In consequence, it is argued that education-based policies play an important role in reducing crime rates (Lochner (2004, 2010)).

2 The model

2.1 Legal and Illegal Sectors

Consider a small open economy that produces one homegenous good that can be used for consumption and investment. The good can be produced by two technologies, one which uses skilled labor and capital and another one that uses only unskilled labor. These define a two legal sector economy that demands workers. Nonetheless, in this economy some potential workers could choose to become delinquents and enter an illegal sector with the explicit purpose of acquiring the consumption good by targeting workers of the legal sector. We now describe these technologies.

[2] To see a complete summary of this evidence, see Soares (2004).

Production in the legal skilled labor sector is described by $Y_t^s = F(K_t, L_t^s)$ where Y_t^s is output, K_t is capital and L_t^s is skilled labor, while F is a concave production function with constant returns to scale. It is assumed that investment in human capital and in physical capital is made one period in advance and that there are no adjustment costs to investment and no depreciation of capital. Legal firms can borrow at interest rate $r > 0$ from world markets and as in Galor-Zeira (1993) we assume the absence of adjustment costs to investment, and given the fact that the number of skilled workers is known one period in advance, the amount of capital in the skilled labor sector is adjusted each period so that $F_K(K_t, L_t^s) = r$. Hence there is a constant capital-labor ratio in this sector, which determines the wage of skilled labor w^s which is *constant* as well. This wage w^s depends on r and on technology only. We follow Galor-Zeira (1993) in assuming that all markets are perfectly competitive and expectations are fully rational.

Production in the legal unskilled labor sector is described by $Y_t^n = G(L_t^n, N)$ where Y_t^n, L_t^n and N are output, unskilled labor and land respectively. Let the aggregate amount of land be fixed at $\bar{N}$, so that demand of unskilled labor is

$$G_L(L_t^n, \bar{N}) \equiv Q(L_t^n)$$

where Q is a function that describes the diminishing marginal productivity of unskilled labor. As in Galor-Zeira (1993) we assume that land is traded in a perfectly competitive market, which operates each period after production takes place and given that there is no uncertainty in production of this deterministic economy, land is an asset which is equivalent to lending.

The illegal sector is an abstraction of an organized sector dedicated exclusively to take away income from legal workers. It abstracts from the different types of illegal pecuniary activities that arise in the real world, like robbery in general, burglary, kidnapping, economic extortion etc, but that can be understood as having the same end in sight, namely material incentives by preying on legal workers.[3] The organization of the "firms" that operate in this sector is conceptualized in the following manner: members of the organization acquire the income from illegal activities and then share equally with all the other members. This simplifies away the hierarchy of the organization that would presumably divide in an unequal fashion the income acquired. The acquisition of the income in the illegal sector is described by the following "pseudo production function" which is assumed to be linear in the input labor where delinquents and workers are matched randomly:

$$E\left(Y_t^d\right) = (1-\pi)\rho[\theta_t W_t^n + \eta_t W_t^s]L_t^d. \tag{1}$$

The term $E\left(Y_t^d\right)$ denotes the expected income that is acquired through delinquency, θ_t and η_t are respectively the probabilities of encountering both unskilled and skilled workers in period t, L_t^d is the labor needed in the delinquency sector, $\rho \in [0, 1)$ represents the fraction of the wealth that a delinquent is able to get from his victims in any given encounter, W_t^n and W_t^s are the unskilled and skilled overall wealth respectively. Since the model only has two kind of individuals, namely

[3] This differs for illegal activities like illegal drugs which are goods that are considered to be illegal but are produced in the same way as legal goods.

legal workers and delinquents, then it must be the case that $\theta_t + \eta_t = 1 - \lambda_t$ where λ_t is the probability in period t of encountering a delinquent in any given random match. We assume that encounters among delinquents do not generate any net gain for them. With probability $\pi \in (0, 1)$ the delinquent is apprehended by law enforcement authorities in which case no wealth is maintained by the delinquent[4], while with probability $(1-\pi)$ a delinquent can obtain a net amount $\rho[\theta_t W_t^n + \eta_t W_t^s]$ of expected wealth since under random matching a delinquent "gains" $\theta_t W_t^n + \eta_t W_t^s$ from legal workers while "not gaining" anything from delinquents.

We can define an average expected "implicit wage" acquired by a delinquent in this economy as $w_t^d \equiv E\left(Y_t^d\right)/L_t^d$ given the assumption of income sharing among members of the illegal sector and therefore one can rearrange (1) to represent w_t^d as

$$w_t^d \equiv E\left(Y_t^d\right)/L_t^d = (1-\pi)\rho\left[(1-\lambda_t-\eta_t)W_t^n + \eta_t W_t^s\right] \tag{2}$$

Note that w_t^d is a decreasing function in λ_t for a given value η_t which means that a higher probability of encountering a delinquent lowers the material incentives for all delinquents in this sector in expected terms. Hence, the illegal sector becomes less attractive when more delinquents enter the sector.

2.2 Preferences and Overlapping Generations

Individuals in this economy live two periods as young and old adults each in overlapping generations. In each generation there is a continuum of individuals of size L. Each individual has just one child (there is no population growth), can work as unskilled in the first period of her life or invest in human capital when young and work as skilled worker when old, or choose a delinquency activity when young. For simplicity we shall assume that all individuals consume when old and only work one period. Unskilled workers and delinquents work when young while skilled workers do so when old. Delinquents enjoy their loot when old if they are not apprehended by law enforcement authorities when young. Moreover, we assume explicitly that decisions are irreversible which implies that a delinquent cannot go back to the legal unskilled sector when old.[5] Individuals that choose to educate themselves invest $h > 0$ when young and are able to work in the skilled labor sector when old given that we assume away unemployment in any sector.

All individuals consume when old, work in one period of their life, care in the same way about their children and can derive utility from leading a non delinquent life. This is modelled with a log utility specification in the following way

$$u = \alpha \log c + (1-\alpha)\log b - d \log I,$$

[4] We assume that once an offense occurs with probability π law enforcement authorities are able to apprehend the delinquent and give back the wealth seized to the victim at no cost to the victim.

[5] This assumption of irreversibility is strong but Tauchen, Witte and Griesinger (1994) found evidence of a negative relation between studying and/or working with the probability of engaging in criminal activities. They argue that this behavior comes from keeping individuals linked to legal activities through there contact with either an educational or labor institution and not necessarily due to a higher education attainment that brings higher wages.

where $0 < \alpha < 1$ captures the weight on consumption of an individual, c is consumption in the second period, b is the bequest left to his/her child, I is a psychic cost[6] of committing delinquent acts, $d = \{0, 1\}$ is a binary variable such that $d = 1$ means that an individual chooses to be a delinquent and zero otherwise. All individuals are born with the same potential abilities, same preferences and psychic cost from engaging in illegal activities. They differ only in the amounts they inherit from their parents in terms of wealth x_t where $D_t(x_t)$ is the cumulative distribution function of wealth x_t in period t with support $[0, \infty)$. This distribution satisfies $\int_0^\infty dD_t(x_t) = L$.

As argued above we assume the existence of financial markets that allow individuals to save and earn interest on their savings at interest rate $r > 0$ given exogenously by world markets. The financial markets lend these funds to firms that pay interest rate r. Nonetheless, we assume an imperfection in the credit market for individual borrowers that want to invest in education, namely that no access to credit is allowed to finance investment in human capital.[7] Hence, individuals born in period t that choose to invest in human capital can do so only if they have enough wealth to pay the investment h. This is a working assumption that can be relaxed with less stringent market imperfections in line with Galor-Zeira (1993) without affecting the main results that we find.

2.3 Optimal Bequests

Recall λ_t denotes the probability in period t for a legal worker to encounter a delinquent. When the encounter occurs the delinquent steals fraction ρW_t from a worker with overall wealth W_t, otherwise the encounter does not occur and the worker loses nothing. Therefore an individual born in period t with wealth W_t chooses b_t in order to maximize expected utility

$$\begin{aligned} \max_{b_t} E(U_t) &= \alpha[(1-\lambda_t)\log(W_t - b_t) + \lambda_t \log((1-\rho)W_t - b_t)] \\ &\quad + (1-\alpha)\log b_t - d\log I \end{aligned} \tag{3}$$

We assume that stealing affects directly the consumption of the individual since it is equal to $W_t - b_t$ if the individual is not matched with a delinquent and is $(1-\rho)W_t - b_t$ if matched with one. The first order condition boils down to

$$\frac{\partial E(U_t)}{\partial b_t} = -\frac{\alpha(1-\lambda_t)}{(W_t - b_t)} - \frac{\alpha\lambda_t}{(1-\rho)W_t - b_t} + \frac{1-\alpha}{b_t} = 0 \tag{4}$$

or equivalently

$$\frac{b_t(1-\lambda)}{W_t - b_t} + \frac{b_t\lambda_t}{(1-\rho)W_t - b_t} = \frac{1-\alpha}{\alpha}.$$

It turns out to be a quadratic function in b_t with solution

$$b_t = W_t \left\{ \frac{\alpha}{2} \left[B(\lambda_t) - \sqrt{B(\lambda_t)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2}} \right] \right\} \equiv W_t \Gamma(\lambda_t) \tag{5}$$

[6] This phsychic cost can represent guilt or shame from commtting criminal acts.

[7] This might be rationalized by assuming that individuals that invest a certain amount (h) in their education through acquiring a credit can leave the country at zero cost without paying back the loan.

where $B(\lambda_t) = 1-\rho(1-\lambda_t)+\left(\frac{1-\alpha}{\alpha}\right)(2-\rho) > 0$ since $\rho(1-\lambda_t) < 1$. Importantly the optimal bequest is a linear function of W_t and we take the negative root as the solution of the problem[8] showing in the appendix that $\Gamma'(\lambda) < 0$ and $0 < \Gamma(\lambda) < 1$ for all $\lambda \in [0,1]$ which guarantees that the optimal bequest is always positive. Interestingly the economic interpretation of $\Gamma'(\lambda) < 0$ is quite intuitive since it means that the more likely an individual is robbed the less likely she will be able to inherit to her child and therefore the more likely she will consume out of her wealth. This shows how the likelihood of being a delinquent victim affects negatively inheritances.

Replacing $b_t = W_t\Gamma(\lambda_t)$ in the expected utility function that is maximized in (3) yields the expected life time indirect utility function

$$EU = \log W_t - d\log I + \varepsilon(\lambda_t) \tag{6}$$

where $\varepsilon(\lambda_t) = \alpha(1-\lambda_t)\log(1-\Gamma(\lambda_t))+\lambda_t\log(1-\rho-\Gamma(\lambda_t))+(1-\alpha)\log\Gamma(\lambda_t)$. Note that $\varepsilon(\lambda_t) \leq 0$ and furthermore $\varepsilon'(\lambda_t) < 0$. Function (6) proves important to determine the different choices that individuals make.

2.4 Wealth Distribution and Short-Run Equilibrium

We now turn to describe individual optimal decisions. There are three occupations that individuals can choose: unskilled worker (n), skilled worker (s) and delinquent (d). Overall wealth consists of inherited wealth denoted by x and income earned during the lifetime of an individual. Therefore the overall wealth levels of unskilled and skilled workers are respectively $W_t^n \equiv x_t+w_t^n$ and $W_t^s \equiv x_t+w^s$ for period t. Consider an individual that inherits $x_t \geq h$ who decides to work as skilled ($d=0$) and invest in human capital, her lifetime indirect utility and bequest are respectively

$$\begin{aligned} EU^s(x_t) &= \log\left[(w^s+(x_t-h)(1+r))\right]+\varepsilon(\lambda_t) \\ b^s(x_t) &= [(w^s+(x-h)(1+r)]\Gamma(\lambda_t). \end{aligned}$$

Consider now an individual who inherits an amount $0 < x_t < h$ of wealth in her first period of life and decides to work as unskilled ($d=0$) and not invest in human capital then her lifetime indirect utility and bequest are

$$\begin{aligned} EU^n(x_t) &= \log\left[(x_t+w_t^n)(1+r)\right]+\varepsilon(\lambda_t) \\ b^n(x_t) &= [(x_t+w_t^n)(1+r)]\Gamma(\lambda_t). \end{aligned}$$

Alternatively, an individual who inherits an amount $0 \leq x_t < h$ of wealth in his first period of life and decides to become a delinquent ($d=1$) loses utility $\log I$ and has lifetime utility and bequest

$$\begin{aligned} EU^d(x_t) &= \log\left[\left(x_t+w_t^d\right)(1+r)\right]-\log I+\varepsilon(\lambda_t) \\ b^d(x_t) &= \left[\left(x_t+w_t^d\right)(1+r)\right]\Gamma(\lambda_t). \end{aligned}$$

[8] This is due to the economic intuition of the solution which shall be explained below.

Since occupational choices are irreversible once taken then a delinquent which chooses this when young cannot become a skilled worker that has to invest in education in her first period of life. Hence no educated deliquents can arise in the model.

If the wage differential between skilled and unskilled is sufficiently wide, taking into account the investment cost h, all legal workers would prefer to work as skilled. To see this notice that $EU^s(x_t) \geq EU^n(x_t)$ is true if and only if

$$w^s - h(1+r) \geq w_t^n(1+r) \tag{7}$$

for every t. We assume that (7) holds true for every value of w_t^n.

The possibility of gaining access to education depends then on inherited wealth. Individuals with inherited wealth x_t strictly less than h cannot educate themselves given that it has been assumed away any possibility for financing this investment with future earnings. These individuals prefer to work as legal unskilled workers relative to becoming a delinquent as long as $EU^n(x_t) \geq EU^d(x_t)$, that is as long as[9]

$$(x_t + w_t^n)I \geq x_t + w_t^d. \tag{8}$$

Note from (2) that $w_t^d = (1-\pi)\rho\left[(1-\lambda_t-\eta_t)\, W_t^n + \eta_t W_t^s\right]$ and by construction $W_t^n \equiv x_t + w_t^n$ and $W_t^s \equiv x_t + w^s$. Replacing these in (**??**) yields a threshold wealth level as a function of λ_t and w_t^n expressed as

$$x_t \geq f\left(\lambda_t, w_t^n\right) = \max\left\{0, \frac{(1-\pi)\rho\left[(1-\lambda_t-\eta_t)\, w_t^n + \eta_t w^s\right] - w_t^n I}{I - 1 - (1-\pi)\rho\,(1-\lambda_t)}\right\}. \tag{9}$$

We assume that $I \geq 2$ from now on which implies that the denominator in (9) is positive while if I is large enough under a small wage gap the numerator can be negative which explains the max operator.

The supply of unskilled labor depends on the wealth distribution of the economy since those who have a wealth level between $f\left(\lambda_t, w_t^n\right)$ and h choose to be unskilled workers given that we shall assume through out that $f\left(\lambda_t, w_t^n\right) \leq h$. Hence, since each individual has one unit of labor each period the supply function of unskilled labor in period t is given by

$$S_{n,t} = L_t^n \equiv \int_{f(\lambda_t, w_t^n)}^{h} dD_t(x_t) \tag{10}$$

Competitive markets in the unskilled sector equate aggregate demand and supply of unskilled labor i.e. $S_{n,t} = Q\left(\bar{N}\right)$ to determine the unskilled wage w_t^n in each period. *Figure 1* illustrates both the demand and supply of unskilled labor such that at their intersection the unskilled wage is determined. Given that the aggregate demand is fixed in any given period for a given value of

[9] We have assumed in this calculation that an individual thinks about himself if becoming a delinquent as such that he does not vary the fraction of delinquents in the economy. This implies that the term $\varepsilon\left(\lambda_t\right)$ can be eliminated on both sides of the inequality.

$\bar{N}$ this unskilled equilibrium wage depends negatively on the fraction of unskilled workers in the economy i.e. $w_t^n = w^n\left(\theta_t, \bar{N}\right)$ such that $\frac{\partial w_t^n}{\partial \theta_t} < 0$. Importantly it increases with the level of land in the economy i.e. $\frac{\partial w_t^n}{\partial \bar{N}} > 0$ given that this would shift demand to the right and for the same supply of workers the unskilled wage must increase. For future reference define $\overline{w^n} \equiv w^n\left(0, \bar{N}\right)$ as the highest feasible unskilled wage when no individual would supply unskilled labor and $\underline{w^n} \equiv w^n\left(1-\eta_t, \bar{N}\right)$ as the lowest yet positive unskilled wage when all labor of individuals with less than h is allocated to the legal unskilled sector.

The amount an individual inherits in her first period of life, therefore, *fully* determines her decisions whether to invest in human capital or work as unskilled or become a delinquent, and how much to consume and bequeath. Hence, the distribution D_t determines economic performance in period t: the amount of skilled labor $L_t^s = \int_h^\infty dD_t(x)$, delinquency $L_t^d = \int_0^{f(\lambda_t, w_t^n)} dD_t(x)$ and unskilled labor $L_t^n = \int_{f(\lambda_t, w_t^n)}^h dD_t(x)$.

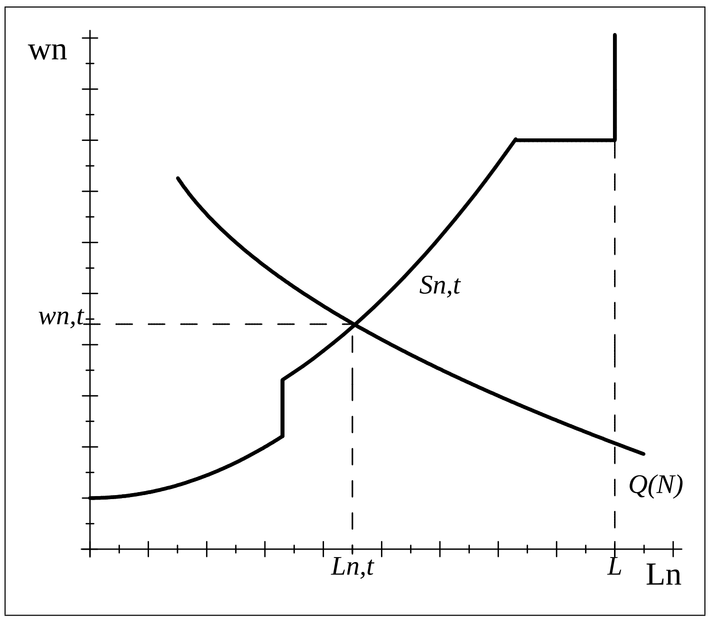

Figure 1

Notice that at $w_t^n = w_s/(l+r) - h$ individuals are indifferent between investing in human capital and working as unskilled, hence the supply curve becomes flat at this wage, then it is upward sloping but can contain horizontal as well as vertical segments. If there is a group of positive measure who inherit the same amount in period t, then there is a horizontal segment in the supply curve. If the distribution D_t is such that there are no inheritances between $f(w_o)$ and $f(w_1)$, then the supply curve is vertical between w_o and w_1. The equilibrium in the unskilled labour market, determines the wage of unskilled, the number of unskilled workers and the number of investors in human capital. It is clear from Figure 1 that this equilibrium depends on the distribution of inheritances D_t.

The amount an individual inherits in her first period of life, therefore, fully determines her decisions whether to invest in human capital or work as unskilled or become a delinquent, and how much to consume and bequeath. Hence, the distribution D_t fully determines economic performance in period t: the amount of skilled labor $L_t^s = \int_h^\infty dD_t(x)$, delinquency $L_t^d = \int_0^f dD_t(x)$ and unskilled labor $L_t^n = \int_f^h dD_t(x) \geq 0$ if $f \leq h$. The distribution of wealth determines the REE in period t since it determines the different actions taken by the individuals.

Rational expectations require consistency of expectations and chosen occupations such that the following are satisfied

$$\begin{aligned} \eta_t &= \frac{\int_h^\infty dD_t(x_t)}{L}; \theta_t = \frac{\int_{f(\lambda_t, w_t^n)}^h dD_t(x_t)}{L}; \\ \lambda_t &= (1-\pi\xi) \frac{\int_0^{f(\lambda_t, w_t^n)} dD_t(x_t)}{L} \end{aligned} \tag{11}$$

where the fraction $\pi\xi \frac{\int_0^{f(\lambda_t, w_t^n)} dD_t(x_t)}{L}$ represents the fraction of delinquents that are apprehended and convicted in period t under random matching[10] and rationalizes that law enforcement authorities can incapacitate at most ξ of the fraction of apprehended delinquents in a given period by putting them in jail.[11] This motivates the following definition.

Definition 1 *A short run rational expectations equilibrium (SREE) of the economy described above consists of a distribution of fractions $\nu_t = [\lambda_t, \theta_t, \eta_t]$ for period t where $\lambda_t + \theta_t + \eta_t = 1$ such that in period t individuals maximize expected utility, firms have zero profits, markets balance and conditions (11) are met.*

Theorem 1 *If the economy described above satisfies (7) for all w_t^n and the equilibrium unskilled wage is low enough $w_t^n \in [\underline{w^n}, M_1)$ for a certain threshold M_1 then it has a unique SREE with $\lambda_t \in (0, 1-\eta_t]$ for any given t. Otherwise if $w_t^n \in [M_1, \overline{w^n}]$ then $\lambda_t = 0$.*

Proof. Firms have zero profits in equilibrium given the assumption of constant returns to scale in both legal sectors. Individuals maximize expected utility and choose optimally bequests and occupations in period t given the theshold values h and $f(\lambda_t, w_t^n)$. Under (7) we get $\eta_t = \frac{\int_h^\infty dD_t(x_t)}{L} > 0$ which is given in period t. Hence to establish the existence of a SREE one has to establish the existence of $\lambda_t \in [0, 1-\eta_t]$ that satisfies (11) recognizing that the cutoff wealth level $f(\lambda_t, w_t^n)$ is a function of λ_t for given η_t from (9). Since $\theta_t = 1 - \lambda_t - \eta_t$ and η_t is given for period t we have that an increase in λ_t is a proportional decrease in θ_t.Given that the equilibrium unskilled wage satisfies $\frac{\partial w_t^n}{\partial \theta_t} < 0$ then it must be the case that $\frac{\partial w_t^n}{\partial \lambda_t} > 0$. Consequently define the following continuous function in λ_t

$$g(\lambda_t) \equiv \lambda_t - (1-\pi\xi) \frac{\int_0^{f(\lambda_t, w_t^n(\lambda_t))} dD_t(x_t)}{L}$$

[10] We denote as the conditional probability of convicting an individual given that he has been apprehended as $P(c|a) = \xi$. Hence the joint probability of apprehending and convicting a delinquent is $P(a,c) = P(a)P(c|a) = \pi\xi$.

[11] Importantly individuals that are put in jail in period t do not circulate in the economy in that period therefore they are modelled here "as if" they disappeared or vanished in the distribution of wealth for (only) period t. They could still have children in jail so the population growth is zero at all times.

in the support $[0, 1-\eta_t]$. Note that

$$g(0) = -(1-\pi\xi)\frac{\int_0^{f(0,\underline{w}^n)} dD_t(x_t)}{L} \leq 0$$

which is zero if $f(0,\underline{w}^n) \leq 0$ or equivalently if $w_t^n \in \left[\frac{(1-\pi)\rho\eta_t w^s}{I-(1-\pi)\rho(1-\eta_t)}, \bar{w}^n\right]$ is satisfied while negative if $f(0,\underline{w}^n) > 0$ or equivalently $w_t^n \in \left[\underline{w}^n, \frac{(1-\pi)\rho\eta_t w^s}{I-(1-\pi)\rho(1-\eta_t)}\right)$ is satisfied and where define $M_1 = \frac{(1-\pi)\rho\eta_t w^s}{I-(1-\pi)\rho(1-\eta_t)} > 0$ given that $I \geq 2$. Furthermore under (7) we have $\bar{w}^n = \frac{w^s}{(1+r)} - h$ such that

$$g(1-\eta_t) = 1-\eta_t-(1-\pi\xi)\frac{\int_0^{f(1-\eta_t,\bar{w}^n)} dD_t(x_t)}{L} > 0$$

which holds since the fraction of skilled workers and delinquents that are not captured by law enforcement authorities cannot exceed one. The continuity of $g(\cdot)$ establishes that there exists a λ_t that satisfies

$$\lambda_t = (1-\pi\xi)\frac{\int_0^{f(\lambda_t,w_t^n(\lambda_t))} dD_t(x)}{L}.$$

Moreover note that by Leibniz rule[12]

$$g\prime(\lambda_t) = 1-(1-\pi\xi)\left[f_1 + f_2\frac{\partial w_t^n}{\partial\lambda_t}\right]\frac{d_t(f(\lambda_t, w_t^n(\lambda_t)))}{L} > 0$$

since $f_1 \leq 0$, $f_2 = (1-\pi)\rho\theta_t - I \leq 0$ and $\frac{\partial w_t^n}{\partial\lambda_t} > 0$ where $d_t(f(\lambda_t, w_t^n(\lambda_t)))$ is the density function of D_t evaluated at $f(\lambda_t, w_t^n(\lambda_t))$ which is always positive. Hence, the SREE is unique for each t. ■

Note that $\lambda_t > 0$ arises in the SREE if assumption $w_t^n \in [\underline{w}^n, M_1)$ is satisfied. This means that the distribution of wealth in period t is sufficiently unequal and there is a sufficient amount of unskilled labor that arises to make the unskilled wage low enough relative to this threshold value M_1. In other words, a positive equilibrium delinquent fraction $\lambda_t > 0$ occurs under a high enough wealth inequality that induces a sufficiently low unskilled wage i.e. sufficient amount of poor unskilled workers. Furthermore, consistent with what has been described above the less wealthy households in period t are the ones self-selected into delinquency which entails a link between poverty, wealth inequality and delinquency in the short run.

2.5 The Dynamics of Wealth Accumulation and the Poverty Trap

The distribution of wealth not only determines equilibrium in period t, but also determines next period distribution of inheritances through the following dynamic system:

$$x_{t+1} = \begin{cases} b^d(x_t;\lambda_t) = \left[\left(x_t + w_t^d\right)(1+r)\right]\Gamma(\lambda_t) & \text{if } 0 \leq x_t < f_t \\ b^n(x_t;\lambda_t) = [(x_t + w_t^n)(1+r)]\Gamma(\lambda_t) & \text{if } f_t \leq x_t < h \\ b^s(x_t;\lambda_t) = [((x_t - h)(1+r) + w^s]\Gamma(\lambda_t) & \text{if } x_t \geq h \end{cases} \tag{12}$$

where for simplicity we denote $f_t \equiv f(\lambda_t, w_t^n(\lambda_t))$. As seen above individuals who have an x less than f_t choose delinquency while individuals who inherit between f_t and h work as unskilled.

[12] Recall Liebniz rule: $\frac{\partial}{\partial z}\int_{a(z)}^{b(z)} f(x,z)\,dx = \int_{a(z)}^{b(z)} \frac{\partial f}{\partial z}dx + f(b(z),z)\frac{\partial b}{\partial z} - f(a(z),z)\frac{\partial a}{\partial z}$.

Using $\overline{x}^n = b^n(\overline{x}^n; \lambda_t)$ in (12), $\Gamma'(\lambda_t) < 0$ and assuming from here onwards a sufficient condition $(1+r)\Gamma(0) < 1$ one has wealth level $\overline{x}_n$ well defined given by

$$\overline{x}^n (\lambda_t) = \frac{(1+r)w_t^n}{\frac{1}{\Gamma(\lambda_t)} - (1+r)} \tag{13}$$

which is positive and where $\lambda_t \in [0, 1-\eta_t]$ is a SREE fraction of non apprehended delinquents. Individuals who inherit more than h invest in human capital hence using $\overline{x}^s = b^s(\overline{x}^s; \lambda_t)$ in (12) and again under $\Gamma'(\lambda_t) < 0$, $(1+r)\Gamma(0) < 1$ one has wealth level $\overline{x}^s$ given by

$$\overline{x}^s (\lambda_t) = \frac{w^s - h(1+r)}{\frac{1}{\Gamma(\lambda_t)} - (1+r)}. \tag{14}$$

Under assumption (7) we have that $\overline{x}^s (\lambda_t) \geq \overline{x}^n (\lambda_t)$ for all $\lambda_t \in [0, 1-\eta_t]$.

Note that wealth $\overline{x}^s (\lambda_t)$ is decreasing in λ_t given that $\Gamma'(\lambda) < 0$ while $\overline{x}^n (\lambda_t)$ is ambiguous since $\frac{\partial w_t^n}{\partial \lambda_t} > 0$ while $\Gamma'(\lambda) < 0$. It turns out that if the elasticity of unskilled wages with respect to λ is large enough such that

$$\varepsilon_{w^n,\lambda} \equiv \frac{\partial w_t^n}{\partial \lambda_t} \frac{\lambda_t}{w_t^n} > \frac{-\lambda_t \Gamma'}{[1-(1+r)\Gamma]\,\Gamma} \tag{15}$$

then we have that $\frac{\partial \overline{x}^n}{\partial \lambda_t} > 0$. The right hand side threshold $\frac{-\lambda_t \Gamma'}{[1-(1+r)\Gamma]\Gamma}$ is non negative under $\Gamma'(\lambda) < 0$, $\lambda_t \in [0, 1-\eta_t]$ and $(1+r)\Gamma(0) < 1$. Assumption (15) simply states that Γ' is not so sensitive to changes in λ relative to the sensitivity in λ of the equilibrium unskilled wage.

Figure 2 illustrates a typical configuration of the *short run* dynamics of wealth accumulation in the economy given by (12). The points in which the curve intersects with the 45 degree line correspond to $\overline{x}^n$ and $\overline{x}^s$ for a SREE value λ_t. Individuals with wealth levels less than h (including unskilled and delinquents) would move in the short run towards $\overline{x}^n$ while those with wealth level greater than h move towards $\overline{x}^s$. Nonetheless, these wealth levels depend explicitly on λ_t and the dynamics of wealth accumulation and should not be considered the long run steady state wealth levels since one would have to determine within the dynamic system the value $\lambda_\infty \equiv \lim_{t\to\infty} \lambda_t$ to which λ_t converges in the long run.

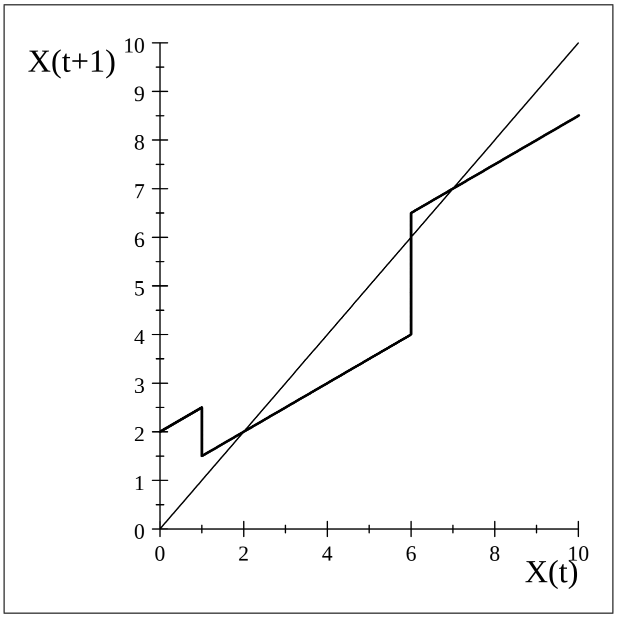

Figure 2

Let us examine the long run behavior of the dynamic system (12). From (9) one can see that the cutoff point f_t and loot w_t^d decrease with λ_t while $\overline{x}^n$ increases with λ_t under assumption (15). Hence in Figure 2 where $f_t < \overline{x}^n$ is satisfied in period t, as the dynamics step in f_{t+1} is higher as a non-negligible fraction of non apprehended delinquents migrate from the illegal sector towards the legal unskilled sector decreasing λ_{t+1}. The wealth level $\overline{x}^s$ increases necessarily as λ_t decreases while $\overline{x}^n$ decreases under assumption (15). Hence, as the economy in Figure 2 moves in time delinquency decreases while the wealth gap $\overline{x}^s - \overline{x}^n$ increases. This motivates two cases to consider: i) a vanishing fraction of delinquents such that $\lambda_\infty = 0$ and ii) persistent delinquency $\lambda_\infty > 0$. If $\lambda_\infty = 0$ then one has a long run behavior as in the Galor-Zeira model abstracting from credit markets for households. Nonetheless, we argue below that in the long run it is possible to have $\lambda_\infty > 0$ under plausible conditions. Regardless of either case this convergence process requires us to consider a steady state in which $\lambda_\infty \equiv \lim_{t\to\infty} \lambda_t$. Consequently a steady state in the long run of the dynamics of wealth accumulation is such that the migration outflow is exactly the migration inflow to the delinquent sector making λ converge to a constant. This motivates the following definition.

Definition 2 *A long run rational expectations equilibrium (LREE) consists of a SREE in which* $\lim_{t\to\infty} \overline{x}^i(\lambda_t) = \overline{x}^i(\lambda_\infty)$ *for* $i = n, s$, *and the long run wealth threshold* f_∞ *satisfies* $f_\infty = \overline{x}^n(\lambda_\infty)$ *if* $\lambda_\infty \in (0, 1 - \eta_t]$ *or* $f_\infty < \overline{x}^n(0)$ *if* $\lambda_\infty = 0$.

To get intuition for this definition consider Figure 2 and let us focus on the dynamics of the bequest functions $b^d(x_t; \lambda_t)$ and $b^n(x_t; \lambda_t)$ as time evolves. Since the process starts off such that $f_t < \overline{x}_n(\lambda_t)$ then some fraction of the offspring of (non apprehended) delinquent households cross f (namely those with wealth level arbitrarily close to f_t) and enter the legal unskilled sector inducing a decrease in λ_{t+1}.[13] This in turn increases the threshold f_{t+1}, loot w^d_{t+1} while decreasing $\overline{x}^n$ under assumption (15). The net effect is that λ should eventually decrease weakly so long as $f_s < \overline{x}_n(\lambda_s)$ for some $s > t$. This process continues up to the point in which equality $f_\infty = \overline{x}_n(\lambda_\infty)$ occurs consistent with persistent delinquency $\lambda_\infty \in (0, 1-\eta_t]$. Nonetheless it could happen that delinquency vanishes before this equality is reached i.e. $f_\infty < \overline{x}_n(0)$ consistent with $\lambda_\infty = 0$. A similar logic occurs for the case in which the dynamic process starts off with $f_t > \overline{x}_n(\lambda_t)$. In this case delinquency increases as more households are induced by the dynamics around $\overline{x}_n$ to enter the delinquent sector increasing $\overline{x}_n(\lambda_t)$ under assumption (15) and eventually decreasing f_t. If $\bar{w}^n = w_s/(1+r) - h$ then no delinquency arises since in that case a delinquent would earn the same as a skilled worker but would would have to suffer a psychic cost of I. Hence, not all unskilled labor could choose a delinquent life and $\lambda_\infty \in (0, 1-\eta_t)$.

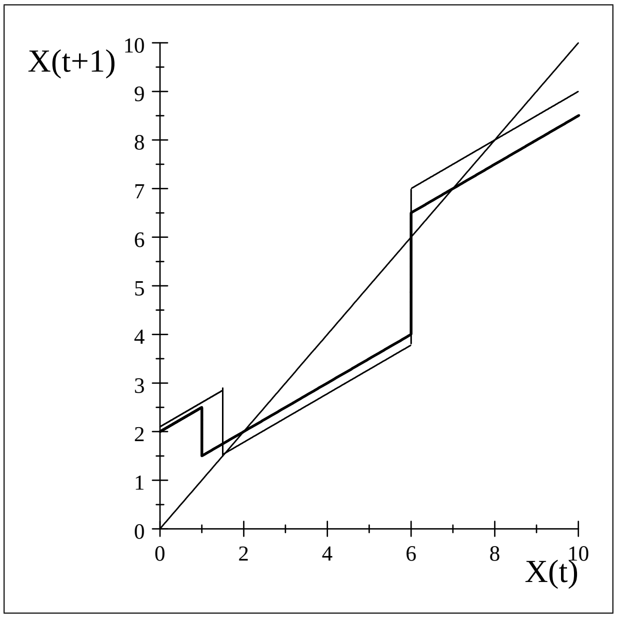

Figure 3

Figure 3 illustrates the limiting behavior of the dynamic system where the *thin* line is consistent with the case $f_\infty = \overline{x}_n(\lambda_\infty)$ for persistent delinquency $\lambda_\infty \in (0, 1-\eta_t)$. Given that $h > \overline{x}_n(\lambda_\infty)$ a *poverty trap* arises which induces persistent inequality and delinquency $\lambda_\infty > 0$ in the long run where there must be a balance between the inflow and outflow of individuals from and to the illegal sector. To see this note that *apprehended* and *convicted* delinquent households that have just $f_\infty = \overline{x}_n(\lambda_\infty)$ (or ε less of wealth) in the long run will not be able to increase their wealth in w^d_∞ forcing them to leave a bequest less than $\overline{x}_n(\lambda_\infty)$ for their offspring, given that they still have to consume when adults. Hence, these offspring would necessarily choose again and again to become delinquents consistent with having persistent delinquency in the economy. On the other hand, *non apprehended*

[13] This is because dynastic delinquent households as they accumulate wealth would cross eventually the threshold f_s for some period $s > t$ given that they are only delayed some finite number of periods by some law enforcement detentions.

delinquents that have just $f_\infty = \overline{x}_n(\lambda_\infty)$ (or ε less of wealth) would be able to secure loot w^d_∞ allowing them to bequest a wealth level greater than $\overline{x}_n(\lambda_\infty)$. Hence, their offspring would choose to become legal workers in the next period. Nonetheless, these households due to the existence of a poverty trap would eventually end up having again $f_\infty = \overline{x}_n(\lambda_\infty)$ in the long run and therefore could end up having some offspring that could choose to become delinquents. It is this outflow and inflow of individuals from and to the illegal sector that would have to be balanced off in the long run consistent with a LREE such that $\lambda_\infty > 0$ remains constant. Figure 4 illustrates a bimodal pdf wealth distribution d_∞ consistent with this dynamic process and corresponds to the thin line of Figure 3.

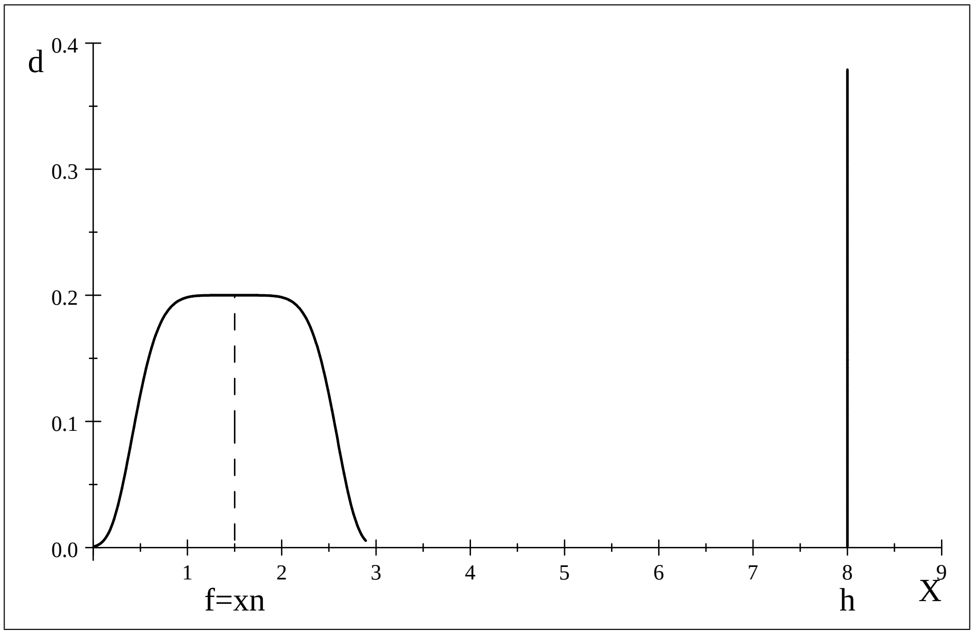

Figure 4

Moreover, to get $\lambda_\infty > 0$ in the long run one requires additionally that $b^d(f_\infty;\lambda_\infty) < h$. To see why consider what would happen if we had $b^d(f_\infty;\lambda_\infty) \geq h$. In this case the offspring of delinquent households with wealth level f_∞ would inherit enough wealth to educate themselves leapfrogging over the poverty trap and entering eventually the skilled sector. Hence, in the long run $\lambda_\infty = 0$. To get persistent delinquency one then requires the necessary condition $b^d(f_\infty;\lambda_\infty) < h$. It remains to show that under certain conditions there exists a LREE with persistent delinquency.

Theorem 2 *If the economy described above satisfies (15), (7) for all w^n_t, $h > I\overline{x}^n(\lambda)$ for all $\lambda \in [0, 1-\eta_t]$ and the long run equilibrium unskilled wage is low enough $w^n_\infty \in [\underline{w^n}, M_2)$ then there exists a unique LREE of the economy described above such that $\lambda_\infty \in (0, 1-\eta_t)$.*

Proof. Consider a SREE and note that assumptions $h > \overline{x}^n$, $\underline{w^n} > 0$ and $(1+r)\Gamma(0) < 1$ implies that $\overline{x}_n > 0$ intersects the 45 degree line and is bounded away from infinity generating a poverty trap since otherwise the b^n function would not intersect the 45 degree line. Assumption (7) guarantees that $\overline{x}^s(\lambda) \geq \overline{x}^n(\lambda)$ for all λ. Define the following function on the domain $[0, 1-\eta_t]$

$$m(\lambda) = f(\lambda, w^n(\lambda)) - \overline{x}^n(\lambda)$$

which is a continuous function of λ given that w_t^n, $\Gamma(\lambda)$ are assumed continuous in λ and the max function is a continuous function. Note that $m'(\lambda) < 0$ since $f'(\lambda) < 0$ and $\frac{\partial \overline{x}^n}{\partial \lambda} > 0$ under assumption (15). Moreover, $f(0, \underline{w^n}) > \overline{x}^n(0)$ arises if $w_\infty^n \in [\underline{w^n}, M_2)$ where $M_2 \equiv \frac{(1-\pi)\rho\eta_t w^s}{I-(1-\pi)\rho(1-\eta_t)+\frac{I-1-(1-\pi)\rho}{\frac{1}{(1+r)\Gamma(0)}-1}} < M_1$. Hence $m(0) > 0$ and $m'(\lambda) < 0$ which generates persistent delinquency since $m(\lambda_\infty) = 0$ must involve $\lambda_\infty \in (0, 1-\eta_t)$. On the other hand if $w_\infty^n \in [M_2, \bar{w}^n]$ one would have $m(0) \leq 0$ and long run steady state is compatible with $\lambda_\infty = 0$. We still need to check that $b^d(f_\infty; \lambda_\infty) < h$ holds. Assumption $h > I\overline{x}^n(\lambda)$ for all $\lambda \in [0, 1-\eta_t]$ implies that $h > I\overline{x}^n(\lambda_\infty)$ which can be rewritten as

$$\begin{aligned}
(1+r)\Gamma(\lambda_\infty)w^n(\lambda_\infty) I &< h\left[1-(1+r)\Gamma(\lambda_\infty)\right] \\
w^n(\lambda_\infty)\left[\frac{1}{1-(1+r)\Gamma(\lambda_\infty)}\right] &< \frac{h}{I(1+r)\Gamma(\lambda_\infty)} \\
\frac{(1+r)w^n(\lambda_\infty)}{\frac{1}{\Gamma(\lambda_\infty)}-(1+r)} &< \frac{h}{I(1+r)\Gamma(\lambda_\infty)} - w^n(\lambda_\infty) \\
f_\infty &= \overline{x}_n(\lambda_\infty) < \frac{h}{I(1+r)\Gamma(\lambda_\infty)} - w^n(\lambda_\infty)
\end{aligned}$$

since in LREE with $\lambda_\infty \in (0, 1-\eta_t)$ we have $f_\infty = \overline{x}_n(\lambda_\infty)$. Moreover

$$\begin{aligned}
(f_\infty + f_\infty(I-1) + w^n(\lambda_\infty) I) &< \frac{h}{(1+r)\Gamma(\lambda_\infty)} \\
f_\infty + w_\infty^d &< \frac{h}{(1+r)\Gamma(\lambda_\infty)}
\end{aligned}$$

since from (2) $w_\infty^d = f_\infty(I-1) + w^n(\lambda_\infty) I$. Note that this last expression rearranged corresponds to $b^d(f_\infty; \lambda_\infty) = \left(f_\infty + w_\infty^d\right)(1+r)\Gamma(\lambda_\infty) < h$. Hence, $b^d(f_\infty; \lambda_\infty) < h$ is satisfied. ■

Some remarks are in order.

i) Assumption $h > I\overline{x}^n(\lambda)$ for all $\lambda \in [0, 1-\eta_t]$ allows the existence of the poverty trap which in turn makes it more likely that persistent delinquency arises in the long run. This is because h sufficiently large relative to $I\overline{x}^n$ implies that the b^n function always intersects the 45 degree line and also generates $b^d(f_\infty; \lambda_\infty) < h$.

ii) Assumption $w_t^n \in [\underline{w^n}, M_2)$ where interestingly $M_2 < M_1$ suggests that persistent delinquency requires that the unskilled equilibrium wage should be lower than the threshold value M_2 and lower even than the level M_1 shown above for having delinquency in the short run. This is intuitive since it says that sufficient poverty that comes out of a high wealth distribution with a sufficient mass of unskilled workers that lowers w^n beyond M_2 allows the poverty trap to generate persistent delinquency in the long run. This implies that poverty coupled with wealth inequality is a necessary condition for delinquency in the short run but no sufficient for persistent deliquency in the long run unless there is a poverty trap.

iii) Consider Figure 3 again and let us focus on the *thin* line that represents the steady state wealth distribution compatible with $\lambda_\infty \in (0, 1-\eta_t)$ in the case $f_\infty = \overline{x}_n(\lambda_\infty)$. The outflow migration from the illegal delinquent sector to the legal unskilled one should be just the same as the inflow migration from the former to the latter. Hence, we have that a continuous flow of households leaving for some periods the illegal sector just to come back eventually to it due to the poverty trap. So it is perfectly possible to have dynastic households that go in and out of delinquency infinitely many times. This circular flow is maintained because of the condition $b^d(f_\infty; \lambda_\infty) < h$ that does not allow delinquent households to leapfrog over the poverty trap. This intuition is illustrated also in Figure 4 where in the long run the wealth distribution around the poverty trap is not a single point but a region.

3 Comparative Dynamics

In this section we study the dynamic behavior of the economy when there are technological shocks to productivity and lower levels of land for the unskilled sector and trace out the effect in the model.

Let us suppose the economy is in its long run SREE such that $f_\infty = \overline{x}_n(\lambda_\infty)$ consistent with $\lambda_\infty \in (0, 1-\eta_t]$. Consider first increasing w^s due to a possible *temporal* one time exogenous technological shock. Let us trace out the effect within the model. In this case initially b^s shifts outward as well as $\overline{x}^s$, while the loot w^d is shifted upward and the threshold f outward in the short run. Hence, f becomes greater than $\overline{x}^n$ making the illegal sector attractive for individuals with lower wealth. This increases subsequently λ in a discrete manner which in turn increases $\overline{x}^n$ while lowering f and $\overline{x}^s$ from there initial upward shift. In the long run it is reestablished that $f_\infty = \overline{x}^n$ at the same wealth level before the shock. Hence, a temporal increase in w^s produces an outburst of delinquency that eventually dies out later. If the shock is *permanent* the logic is the same but there is a permanent increase in λ_∞ since there is a permanent increase in the incentives to enter the illegal sector. In this case the long run wealth level $\overline{x}^s$ does not go back to the initial one while $f_\infty = \overline{x}^n$ increases permanently.

Consider now a *permanent* decrease in $\bar{N}$ which shifts inward the labor demand in the unskilled sector and therefore reduces the equilibrium unskilled wage w^n. Hence, $\overline{x}^n$ decreases in the short run given assumption (15). Since $f > \overline{x}^n$ arises there is an increase in λ which subsequently decreases f with the influx of delinquents. Hence, w^n and $\overline{x}^n$ are increased from the level after the shock to reestablish $f_\infty = \overline{x}^n$ in the long run. Since f has decreased to a lower wealth level we conclude that the overflow of delinquency reduces both $\overline{x}^s$ and $\overline{x}^n$ permanently empoverishing the economy.

4 Budget Balance Equilibrium

Before we consider any policy analysis we construct the budget of a government and consider equilibria that balance this budget. Assume the economy is in the long run steady state where we drop any time subindex for simplicity. To simplify matters let us consider a government that only taxes skilled wages from which the government gets tax revenues[14]. In the above analysis the disposable skilled wage is now $(1-\tau)\,w^s$ and we assume that $\tau \in (0, \tau_{\max})$ where $\tau_{\max} \equiv 1 - \frac{(1+r)w^n - h}{w^s}$ is the highest tax rate that binds (7) so that there are still incentives to invest in human capital. Consider now the total tax revenues

$$\begin{aligned} \text{Total Revenue}_t &= \int_h^{\infty} \tau w^s dD(x) \\ &= \tau w^s \left[L - D(h)\right] \end{aligned} \tag{16}$$

where we assume that skilled workers first pay their taxes and only then delinquents can actually capture part of the corresponding disposable income $(1-\tau)\,w^s$.

Let us assume for simplicity that the fraction of individuals apprehended in steady state is given by the following linear function

$$\pi = \phi \frac{P}{L^d} \tag{17}$$

where P is the mass of policmen that are hired to combat crime, $L^d \equiv \int_0^f dD(x)$ is the mass of delinquents and $\phi \in [0,1)$ represents the technological efficiency of apprehension. We assume that $P < L^d$ in any period. Moreover, the number of apprehended individuals that are *convicted* in any given period is given by the following technology

$$Cv = \min\{\beta J, P\} \tag{18}$$

where J is the number of judges that convict apprehended delinquents in each period and $\beta \geq 1$ is a parameter that measures the technological relation of policemen per judge which we assume constant and where higher levels of β represent more efficient judicial systems. Function (18) is consistent with the idea that there is a bottleneck to process all apprehended delinquents.

Now a government has to choose how many policemen and judges to hire (paying wages w^s) to apprehend and convict delinquents. Therefore the government has to solve the following cost minimization problem taking as given L^d and restricted to a certain conviction rate defined as $\xi \equiv \frac{Cv}{L^d \pi}$

$$\begin{aligned} \min_{J,P} C &= w^s (J + P) \\ s.t. \quad \xi &= \frac{1}{L^d \pi} \min\{\beta J, P\} \end{aligned}$$

[14] It is conceivable that taxes are progressive in the sense that unskilled workers have a tax rate lower than skilled workers. If true the model assumes that the tax rate for unskilled workers is zero while positive for skilled workers. This goes also in line with the interpretation that unskilled and skilled workers are respectively informal and formal workers in a dual economy.

The solution of such a problem is quite simple and yields a cost function given by

$$C = w^s \left(1 + \frac{1}{\beta}\right) L^d \pi \xi \equiv c L^d \pi \xi \tag{19}$$

where in the optimum $\frac{\beta J}{L^d \pi} = \frac{P_t}{L^d \pi} = \xi$ is satisfied and $c \equiv w^s \left(1 + \frac{1}{\beta}\right)$. This function is linear and increasing in the amount of apprehended deliquents $L^d \pi$ and the conviction rate ξ.

Finally consider that a government wants to subsidize education investments for some non negligible amount of households in the economy, namely for households that have wealth levels in $[a, h]$ where a is a cutoff level to be chosen such that $0 \leq a < h$. If so the government would have to spend the following amount

$$E = \int_a^h (h - \bar{x}^n)\, dD(x) = [h - \bar{x}^n]\, [D(h) - D(a)] \tag{20}$$

where in steady state the average wealth level of the unskilled is $\bar{x}^n$. This way of subsidizing is by targeting households that have wealth levels close to h since under the existence of a poverty trap it would be foolish for a government to subsidize educational investments for *anyone* that has a wealth level less than h. It rationally would have to choose households who have the highest possibility of getting out of the poverty trap to maximize the impact in the long run. This is because for sufficiently poor households a subsidy may not allow future generations to choose a legal life.

Let revenues equal expenditures and express the government budget in per capita terms by dividing (16), (19) and (20) through L so that the following equality holds in steady state

$$\tau w^s \eta = c \lambda \pi \xi + [h - \bar{x}^n]\, \sigma \tag{21}$$

where $\sigma = 1 - \eta - \frac{D(a)}{L} \geq 0$ for $0 \leq a < h$ is the fraction of unskilled households to be subsidized by the amount $h - \bar{x}^n$.[15]

The budget balance equilibrium conditions in steady state are equations (21) and conditions

$$\begin{aligned} \eta &= \frac{\int_h^\infty dD(x)}{L}; \qquad \sigma = \frac{\int_a^h dD(x)}{L} \\ \theta &= \frac{\int_{f(\lambda, w^n)}^a dD(x)}{L}; \;\; \lambda = (1 - \pi\xi)\, \frac{\int_0^{f(\lambda, w^n)} dD(x)}{L} \end{aligned} \tag{22}$$

where now

$$f(\lambda, w^n) = \max\left\{0, \frac{(1-\pi)\rho\, [(1 - \lambda - \sigma - \eta)\, w^n + (\eta + \sigma)\,(1-\tau)\, w^s] - w^n I}{I - 1 - (1-\pi)\rho\,(1-\lambda)}\right\}$$

We call tuple $s = (\tau, \pi, \xi, \sigma, \eta, \lambda, \theta)$ a *budget balanced equilibrium state* in space $S = [0,1]^3 \times [0, 1-\eta] \times [0,1] \times [0, 1-\eta] \times [0,1]$ such that $\eta + \sigma + \lambda + \theta = 1$ and equations (21), (22) are satisfied. Call tuple $p = (\tau, \pi, \xi, \sigma) \in P = [0, \tau_{\max}] \times [0,1]^2 \times [0, 1-\eta]$ a *policy*. This establishes an explicit connection between the tax rate τ and the apprehension and conviction rates π, ξ as well as the

[15] In reality government budgets do not always balance. Budget balance is assumed here in order to define a trade-off between taxes and expenditures on delinquency and education subsidies.

fraction of subsidized households σ. For the moment let us assume that $\sigma = 0$ or equivalently $a = h$. It turns out that for any given distribution (η, λ, θ) the equations (21) and (22) determine the existence of a unique policy $(\tau, \pi, \xi, 0) \in P$ such that the tuple $(\tau, \pi, \xi, 0, \nu)$ constitute a balanced equilibrium state. To see this let (η, λ, θ) be given, such that $\eta \in (0, 1)$, and note that the budget equation (21) defines the tax rate τ as a linear increasing function of π or ξ such that $\tau = 0$ when $\pi = 0$.or $\xi = 0$. Second, the last equation of (22) may be rewritten in the following way using (9) to solve for τ

$$\tau = 1 - \frac{1}{\eta w^s} \left[(I - 1 - (1 - \pi) \rho (1 - \lambda)) D^{-1} \left(\frac{L\lambda}{1 - \pi\xi} \right) + w^n (I - (1 - \pi) \rho\theta) \right] \geq 0 \qquad (23)$$

where D^{-1} is the inverse of the c.d.f. D which exists under the mildassumption that D is strictly increasing. This expression is a strictly decreasing function of ξ (respectively of π) for a given level of π (respectively of ξ). Define τ_1 as the minimum tax rate when $\pi = 0$.and/or $\xi = 0$. If $\tau_1 \in [0, \tau_{\max}]$ then it must be the case that the graph of this function intersects the graph of the linear increasing function at exactly one point, proving the existence and uniqueness claims. This result still holds under a sufficient condition when policy $p = (\tau, \pi, \xi, \sigma)$ entails $\sigma > 0$. In this case again τ in (21) is a linear increasing function of σ while τ in (23) is striclty decreasing function of σ. Now let us define τ_0 from (21) as the tax rate when $\pi = 0$.and/or $\xi = 0$, that is

$$\tau_0 \equiv \left(\frac{h - \bar{x}^n}{w^s \eta} \right) \sigma > 0.$$

The existence and uniqueness claims hold as long as the sufficient condition $\tau_1 \geq \tau_0$ holds. This establishes the following theorem.

Theorem 3 *Given $\tau_1 \geq \tau_0$ for every distribution $(\eta, \lambda, \theta, \sigma)$ with $\eta \in (0, 1)$ there exists exactly one policy $p = (\tau, \pi, \xi, \sigma)$ such that $s = (\tau, \pi, \xi, \sigma, \eta, \lambda, \theta)$ constitutes a balanced equilibrium state.*

Thus any distribution $(\eta, \lambda, \theta, \sigma)$ uniquely determines a balanced policy and thereby a balanced equilibrium state. Conversely, for any policy there exists at most one equilibrium distribution such that the distribution $(\eta, \lambda, \theta, \sigma)$ balances the budget. To see this, suppose that $(\eta, \lambda, \theta, \sigma)$ is such that the state $s = (\tau, \pi, \xi, \sigma, \eta, \lambda, \theta)$ is a budget balanced equilibrium, and let $(\eta, \lambda', \theta, \sigma)$ be another distribution such that $s' = (\tau, \pi, \xi, \sigma, \eta, \lambda', \theta)$ is a budget balanced equilibrium then s' does not meet the budget requirement (21) for if $\lambda' > \lambda$ then public spending in $s\prime$ is higher and public revenues are lower, this similarly occurs if $\lambda' < \lambda$.

5 Policy Analysis

Up to now we have not considered the delinquent fraction that maximizes social welfare. Social welfare in general requires usually to do interpersonal utility comparisons which is undesirable to do most of the times. Fortunately, in this context we do not have to invoke any social welfare criteria[16] to choose the optimal level of λ since from the indirect utility function of non-delinquents (with $d = 0$) we have that $EU = \log W + \varepsilon(\lambda)$ which is strictly decreasing in λ given that $\varepsilon(\lambda) \leq 0$ and $\varepsilon'(\lambda) < 0$. Hence, for a legal worker regardless of her wealth heterogeneity, the optimal level of delinquency is $\lambda = 0$, which turns out to be not surprising at all. Naturally, if the economy has a LREE such that $\lambda = 0$ then no policy intervention for delinquency should be considered. Nonetheless, we have shown that a LREE such that $\lambda = 0$ is not necessarily obtained in the long run since persistent delinquency $\lambda \in (0, 1 - \eta)$ can arise mainly due to the existence of a poverty trap. As defined above a policy is a triple $p = (\pi, \xi, \sigma) \in P$ which can be classified as law enforcement policies and education based policies respectively. We are interested in analyzing the trade-off between these two type of policies to generate $\lambda = 0$ in the long run.

5.1 Law Enforcement Policies vs Education Based Policies

The probability of apprehension π is related with police forces that deter and apprehend delinquents while the conditional probability of convicting an apprehended delinquent ξ is related with the judicial system that convicts felons. Both dimensions of law enforcement are needed in order to "punish" delinquents that are found guilty in a court of justice. Naturally both complement themselves in the model since in (11) one sees that it is the product $\pi\xi$ that matters for *incapacitating* delinquents while only π matters to deter potential felons according to (9). On the other hand, education based policies are policies that allow to overcome financial barriers for undertaking human capital investment. We consider policies that subsidize education tuitions such as scholarships or even public schooling (avoiding any concern on quality heterogeneity). As in the last section we consider an education policy that targets households with wealth sufficiently high such that it becomes sustainable for these households to leave a bequest level high enough for their future generations to leave the poverty trap.

The concept of budget balance is the pertinent concept since it allows us to uncover a trade-off between hiring law enforcement and subsidizing human capital investments. Naturally and ideally if tax revenue are high enough they would allow one to increase education subsidies without sacrificing any law enforcement level. Nonetheless, as any economist would argue resources are scarce and one would be interested in assessing the magnitude of the trade off between hiring less judges and policemen in order to subsidize human capital investment of a targeted population. Let us focus then on this trade-off in the LREE steady state which requires that $f = \overline{x}^n(\lambda)$ consistent with $\lambda \in (0, 1 - \eta)$.

The idea is to understand the trade-off between π and σ maintaining all other things equal except θ which would vary one to one negatively with σ. To do so let us consider jointly equations (21), (23) such that $\theta = 1 - \eta - \sigma - \lambda$ which corresponds to a balanced budget equilibrium state $s =$

[16] Actually in this context in which delinquency is a parasite to society a natural social welfare function would be of the Rawlsian type which entails ordinal scale measurability with full comparability across legal workers (See chapter 5 of Boadway-Bruce 1984). The worst-off legal worker would be an unskilled one with wealth $f(\lambda_\infty) = \bar{x}^n$. Hence, the Rawlsian social welfare function would be $SWF = \{\log(f(\lambda_\infty) + w^n) + \varepsilon(\lambda_\infty)\}$ which is strictly decreasing in λ impliying that the social optimum would be $\lambda_\infty = 0$.

$(\tau, \pi, \xi, \sigma, \eta, \lambda, \theta)$. Moreover consider that there exists a neighborhood such that the explicit function $\pi = \pi(\tau, \xi, \sigma)$ exists. To get $\frac{d\pi}{d\sigma}\big|_{\bar{\lambda}}$ write down explicitly from (21) the tax rate as $\tau = \frac{c\lambda\pi\xi + [h - \bar{x}^n]\sigma}{w^s\eta}$ and replace it in (23) to get the following implicit function

$$\frac{c\lambda\pi\xi + [h - \bar{x}^n]\,\sigma}{w^s\eta} - 1 + \frac{1}{(\eta+\sigma)\,w^s}\left[(I - 1 - (1-\pi)\,\rho\,(1-\lambda))\,D^{-1}\left(\frac{L\lambda}{1-\pi\xi}\right) + w^n\,(I - (1-\pi)\,\rho\theta)\right] \equiv 0$$

Invoking the implicit function theorem one can derive with respect to σ keeping other things equal while imposing $\frac{\partial\theta}{\partial\sigma} = -1$. After some manipulations this comes out to be

$$\begin{aligned} & \frac{\partial\pi}{\partial\sigma}\bigg|_{\bar{\lambda}}\left[c\lambda\pi\xi\,(\eta+\sigma) + \frac{HL\lambda\xi}{(1-\pi\xi)^2} + \rho\,(1-\lambda)\,D^{-1}\left(\frac{L\lambda}{1-\pi\xi}\right) + w^n\rho\theta\right] \\ = \;& \frac{\partial w^n}{\partial\theta}\,(I - (1-\pi)\,\rho\theta) - c\lambda\pi\xi - w^n\,(1-\pi)\,\rho - (h - \bar{x}^n)\,(\eta + 2\sigma) + w^s\eta \end{aligned}$$

The sign of the left hand square bracket expression is strictly positive where we have defined

$$H \equiv (I - 1 - (1-\pi)\,\rho\,(1-\lambda))\,D^{-1\prime}\left(\frac{L\lambda}{1-\pi\xi}\right) \geq 0$$

while the sign of the right hand side term is negative if $w^s\eta$ is dominated by the rest of the negative terms. Hence we have $\frac{\partial\pi}{\partial\sigma}\big|_{\bar{\lambda}} < 0$. Note that this derivative becomes $-\infty$ when $\lambda = 0$ since $D^{-1}(0) = \infty$ This suggests that the derivative decreases in absolute value when λ increases away from zero. This gives us the basic conclusion that a society that has a higher delinquency has a better trade off between π and σ if the subsidy decreases the unskilled fraction of workers.

6 Conclusions

Delinquency seems more persistent than one might think in both developed as well as under developed economies. We study an overlapping generations model under perfect competition that abstracts from unemployment similar to Galor-Zeira (1993) which allows us to explore the theoretical linkages between poverty traps, economic inequality and educational attainment. It takes seriously the idea that delinquents choose rationally a criminal life when there is a lack of opportunities to enter a skilled sector that requires previously to attain a certain level of education. It builds on a dual economy in which delinquents are seen as parasites that prey on legal workers. We characterize the optimal bequest of dynastic households in the three occupational activities (delinquency, unskilled and skilled workers) that emerge which govern the wealth accumulation of the economy. We show that a short run delinquency fraction exists and define a steady state of the dynamic system compatible with the possibility of persistent delinquency in the long run. We find that for given levels of law enforcement measures of deterrence and incapacitation delinquency is persistent in the long run if the unskilled wage is low enough due to a high mass of unskilled workers coupled with high inequality that induces delinquent opportunities.

We study technological shocks that increase skilled wages or reduces land for the unskilled. These temporal shocks produce an outburst of delinquency in the short run that die out later on. If the shock is permanent then the delinquency increases permanently in the long run. The optimal social level of deliquency is zero and we study the trade off between deterrence policies and education based policies to reduce long run persistent delinquency. We find that for higher levels of delinquency education based policies yield a better trade off between increasing subsidies to invest in human capital for unskilled workers while decreasing the apprehension probability by hiring less policemen.

Further research would be to allow for unemployment in the skilled sector as trace out the effect on delinquent incentives. Another extension could be to generalize the model to consider illegal activities such as narcotics or gambling that are not necessarily seen as preying on workers but more as activities that sell workers services that are ilicit in the economy.

Appendix

Proposition 1 *Under the assumptions of the model* $\Gamma'(\lambda) < 0$ *and* $0 < \Gamma(\lambda) < 1$ *for all* $\lambda \in [0,1]$.

Proof. First we show that $\Gamma'(\lambda) < 0$ for all $\lambda \in [0,1]$. From (5) differentiating with respect to λ we get

$$\Gamma'(\lambda) = \frac{\alpha}{2}\left[\rho - \frac{1}{2}\left(B(\lambda)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2}\right)^{-\frac{1}{2}}(2B(\lambda)\rho)\right]$$

since $B'(\lambda) = \rho$. It is sufficient to show that

$$1 < \left(B(\lambda)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2}\right)^{-\frac{1}{2}} B(\lambda).$$

which is satisfied since

$$\begin{aligned}
\left(B(\lambda)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2}\right)^{\frac{1}{2}} &< B(\lambda) \\
B(\lambda)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2} &< B(\lambda)^2 \\
\frac{4(1-\alpha)(1-\rho)}{\alpha^2} &> 0.
\end{aligned}$$

We have used the fact that $B(\lambda)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2} > 0$ for all $\lambda \in [0,1]$. To see why this is the case define

$$h(\lambda) \equiv B(\lambda)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2}$$

and note that $h'(\lambda) = 2\rho B(\lambda) > 0$ and $h''(\lambda) = 2\rho^2 > 0$ for all $\lambda \in [0,1]$. Hence the function is strictly convex, increasing and does not attain a minimum in the interval $[0,1]$ since $h'(\lambda) > 0$ because $B(\lambda) > 0$ for all $\lambda \in [0,1]$.

Second we show $0 < \Gamma(\lambda) < 1$ for all $\lambda \in [0,1]$. First let us show that $\Gamma(\lambda) > 0$ for all $\lambda \in [0,1]$. From (5) it is sufficient to show that $B(\lambda) - \sqrt{B(\lambda)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2}}$ is positive for all $\lambda \in [0,1]$. Note

$$\begin{aligned}
B(\lambda) &> \sqrt{B(\lambda)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2}} \\
B(\lambda)^2 &> B(\lambda)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2} \\
\frac{4(1-\alpha)(1-\rho)}{\alpha^2} &> 0.
\end{aligned}$$

Finally to show that $\Gamma(\lambda) < 1$ for all $\lambda \in [0,1]$ it is sufficient to show $\Gamma(0) < 1$ since we have shown $\Gamma'(\lambda) < 0$ for all $\lambda \in [0,1]$ .Notice that for the negative root

$$\Gamma(0) = \frac{\alpha}{2}\left[B(0) - \sqrt{B(0)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2}}\right] < 1$$

if the following holds

$$B(0) < \frac{2}{\alpha} + \sqrt{B(0)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2}}.$$

We know that $\sqrt{B(0)^2 - \frac{4(1-\alpha)(1-\rho)}{\alpha^2}} > 0$ is adding to $\frac{2}{\alpha}$, then we just need to show that $B(0) < \frac{2}{\alpha}$ which comes down to showing that

$$1 - \rho + \left(\frac{1-\alpha}{\alpha}\right)(2-\rho) < \frac{2}{\alpha}$$

which is satisfied since this yields $(1-\alpha)(2-\rho) + \alpha(1-\rho) < 2$ or $-\alpha - \rho < 0$. ■

References

Acemoglu, D. and J. Angrist. 2000. *"How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws"*. National Bureau of Economic Research Macroeconomics Annual, Vol. 15, pp. 9-59.

Boadway, R. and N. Bruce. 1984. Welfare Economics. Basil Blackwell.

Bowles, S. 2006. "Institutional Poverty Traps". In: Poverty Traps. Eds. S. Bowles, S. Durlauf and K. Hoff. Princeton University Press.

Becker, G. 1968. *"Crime and Punishment: An Economic Approach"*. Journal of Political Economy, Vol. 76, No. 2, pp. 169-217.

Blattman, C. and E. Miguel. 2009. *"Civil War"*. National Bureau of Economic Research, Working Paper 14801.

Bourguignon, F. 1999. *"Inequality and Violent Crime"*. Paper prepared for the Annual World Bank Conference on Development Economics, Washington, D.C. April 28-30.

Dilulio, J. 1996. *"Help wanted: Economists, Crime and Public Policy"*. The Journal of Economic Perspectives, No. 10, pp. 3-24.

Ehrlich, I. 1973. *"Participation in Illegitimate Activities: A Theoretical and Empirical Investigation"*. The Journal of Political Economy, Vol. 81, No. 3, pp. 521-265.

Eide, E. 1997. "Economics of Criminal Behavior: Survey and Bibliography". Working paper in law and economics, No. 5, University of Oslo.

Fajnzylber, P., D. Lederman, and N. Loayza. 2001. *"Crimen y victimización: una perspectiva económica"*. In Fajnzylber, Lederman and Loayza (Eds.), Crimen y violencia en América Latina, México D.F., México, Banco Mundial and Alfaomega Grupo Editor.

Fajnzylber, P., D. Lederman, and N. Loayza. 2002. *"Inequality and Violent Crime"*. Journal of Law and Economics, Vol. 45, No. 1, pp. 1-40.

Fender, J. 1999. *A General Equilibrium Modelo of Crime and Punishment.* Journal of Economic Behavior and Organization. Vol. 39, Issue 4, pp. 437-453.

Freeman, R. 1996. *"Why Do So Many Young American Men Commit Crimes and What Might We Do About It?"*. National Bureau of Economic Research, Working Paper 5451.

Galor, O. and J. Zeira. 1993. *"Income Distribution and Macroeconomics"*. The Review of Economic Studies, Vol. 60, No. 1. pp. 35-52.

Harris, J. and M. Todaro. 1970. *"Migration and Employment and Development: A two Sector Analysis"*, American Economic Review, Vol. 40, pp. 126-142.

Imrohoroglu, A., A. Merlo and P. Rupert. 1996. *"On the Political Economy of Income Redistribution and Crime"*. Federal Reserve Bank of Cleveland, Working paper 9609.

Kelly, M. 2000. *"Inequality and Crime"*. The Review of Economics and Statistics, Vol. 82, No. 4, pp. 530-539.

Krohn, M. 1976. *"Inequality, Unemployment and Crime: A Cross-National Analysis"*. The Sociological Quarterly, Vol. 17, No. 3, pp. 303-313.

Lewis, A. 1954. *“Economic Development with Unlimited Supplies of Labor”*. The Manchester School of Economic and Social Studies, Vol. 22, pp. 139-191.

Lochner, L. and E. Moretti. 2001. *“The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports”*. National Bureau of Economic Research, Working Paper 8605.

Lochner, L. 2010. *“Education, Work and Crime: A Human Capital Approach”*. National Bureau of Economic Research, Working Paper 10478. 2004.

Lochner, L. *"Education Policy and Crime"*. National Bureau of Economic Research, Working Paper 15894.

Mehlum, H., K. Moene and R. Torvik. 2002. "Predator or Prey? Parasitic Enterprises in Economic Development". European Economic Review No. 47, pp. 79-98.

Mehlum, H., K. Moene and R. Torvik. 2006. "Parasites". In: Poverty Traps. Eds. S. Bowles, S. Durlauf and K. Hoff. Princeton University Press.

Merton, R. 1938. "Social Structure and Anomie". American Sociological Review No. 3, pp. 672-682.

Soares, R. 2004. *"Development, Crime and Punishment: Accounting for the International Differences in Crime Rates"*. *Journal of Development Economics* No. 73. pp. 155 - 184.

Shaw, C. and H. McKay. 1942. *"Juvenile Delinquency and Urban Areas"*. University of Chicago Press.

Tauchen, H., A. Witte and H. Griesinger. 1994. *“Criminal Deterrence: Revisiting the Issue with Birth Cohort”*. Review of Economics and Statistics, No. 76, pp. 399-412.

Weibull, J. and E. Villa. 2005. *“Crime, Economic Incentives and Social Norms”*, SSE/EFI Working Paper Series in Economics and Finance, No. 610.

Witte, A. and H. Tauchen. 1994. *"Work and Crime: An Exploration Using Panel Data"*. National Bureau of Economic Research, Working Paper 4794.