On the intergenerational nature of criminal behavior

Carlos Bethencourt,* Lars Kunze†

This version: January 2015

Abstract

Empirical evidence suggests that family background and parental criminality are strong predictors of an individuals' criminal behavior. The aim of this paper is to provide a theoretical foundation of the intergenerational nature of criminal behavior. Drawing on the literature of cultural transmission, we model the dynamics of moral norms of good conduct (honest behavior). Individuals' criminal behavior and morality are strategic complementarities that reinforce each other. We establish the existence of multiple steady states and provide conditions on the socialization process under which both types - honest and dishonest - survive in the long run even though parents may commit crime but at the same time agree that honesty is desirable. Our model provides a novel and complementary explanation of why crime is highly concentrated in specific areas and why it tends to be persistent over time.

Keywords crime, cultural transmission

JEL-Classification Z10, K14, K42

*Carlos Bethencourt, Universidad de La Laguna, Departamento de Economía, Campus de Guajara s/n, 38071 Tenerife, Spain, e-mail: *cbethenc@ull.es*, phone: +34 922 317954, fax: +34 922 317204.

†Lars Kunze, TU Dortmund, Department of Economics, 44221 Dortmund, Germany, e-mail: *lars.kunze@tu-dortmund.de*, phone: +49 231 755-3275, fax: +49 231 755-5404.

1 Introduction

...I never wanted this for you [Michael]...I always thought that when it was your time, that you would be the one to hold the strings. Senator Corleone, Governor Corleone, something.

– Vito Corleone

It is a well established fact that crime runs in the family.[1] Still, the familial nature of criminal behavior has only recently attracted the interest of economists, see e.g. Duncan et al. (2005) and Hjalmarsson and Lindquist (2012, 2013).[2] The findings of these studies emphasize the importance of family background and, in particular, parental criminality for predicting an individual's criminal behavior. Those factors have been found to be even more important than income or employment status. According to Hjalmarsson and Lindquist (2012), for example, the probability of having a criminal conviction is 2.06 (2.66) times higher for a son (daughter) with a criminal father as compared to a son (daughter) with a noncriminal father. Moreover, parents' behavior and socialization processes account for a large share of this intergenerational crime relationship. The aim of this paper is to rationalize this stylized fact within a theoretical model based on the cultural transmission of moral values and socialization within the family.

It is by now well accepted that preferences, beliefs and moral norms are acquired through a learning and socialization process or the imitation of role models, which in turn implies that they are transmitted through generations. This idea has recently received considerable attention in the economic literature (Bisin and Verdier, 2000, 2001). In fact, many authors have argued that the transmission of a particular trait (social status, religion, ethnicity, etc.) is the result of a socialization process inside and outside the family (like e.g. role models and peers).[3]

In our model, to account for the familial nature of criminal behavior, we consider the formation of an 'honesty' trait. Individuals' decisions to commit crime are not only affected by economic incentives but are also influenced by rules of good conduct and morality inherited from previous generations.[4] Parents rationally anticipate that their own criminal behavior has

[1] See Thornberry (2009) and Farrington (2011) for recent surveys of the criminology literature.

[2] See also Eriksson et al. (2014) for an attempt to quantify the relative importance of family background and neighborhood effects in determining criminal behavior.

[3] The idea is based on earlier work in sociology and anthropology (see in particular Boyd and Richerson (1985), Cavalli-Sforza and Feldman (1981)). Bisin and Verdier (2011) provide a recent survey of the literature.

[4] See also Tabellini (2008) who studies the cultural transmission of a norm that affects

a negative impact on the transmission process of the 'honesty' trait. More specifically, the transmission process encompasses direct socialization inside the family and indirect socialization via neighborhood effects and social interactions (horizontal and oblique socialization). This creates a strategic complementarity between current behavior and values, which reinforces the effects of changes in exogenous variables and in the external environment (e.g. crime deterrence or education policies)[5]. The more individuals commit crime, the lower is the likelihood of successfully transmitting positive moral values which in turn expands the share of individuals with norms of bad conduct in society. Thus, policies aimed at deterring criminal behavior may not only alter economic incentives but also have long-lasting and amplifying effects through changes in the cultural transmission process.

Our analysis combines two important strands of literature: The economics of crime and cultural transmission. In contrast to most of the cultural transmission literature, however, we assume that all parents, irrespective of their type, agree that one of the traits (honesty) is superior.[6] We further assume that the parents' criminal behavior (rather than effort) has a direct negative impact on the children's probability of adopting the honest trait. While the standard assumption in the literature is that the transmission process requires some costly effort without further specification of its nature or the kind of activities necessary to produce it, we suggest that the observation of the parents' behavior by their kin is the mechanism through which children may assimilate the cultural traits of their parents. This is consistent with recent empirical evidence on the intergenerational nature of criminal behavior (see e.g. Hjalmarsson and Lindquist (2012, 2013) and the evidence presented in section 2) and allows us to establish a clear difference between the distribution of traits and actual observed behavior. Specifically, even individuals with norms of good conduct may commit crime if it is economically profitable (and, similarly, individuals with norms of bad conduct need

individuals' decisions to behave cooperatively in situations like for example, the Prisoner's Dilemma. Similarly, Hauk and Sáez-Marti (2002) analyze the evolution of morality and corruption.

[5] Glaeser et al. (2003) refer to this mechanism as the 'social multiplier'.

[6] Although it seems reasonable that parents try to transmit their own cultural trait when it comes to language or religion, this is different regarding traits and values associated with poor economic outcomes and low socioeconomic status (e.g. working in the informal economy, crime, etc.). So far there are only three studies exploring the theoretical implications when parents with different traits agree on which trait is desirable: Patacchini and Zenou (2011) focus on educational outcomes, Sáez-Marti and Zenou (2012) on work ethics whereas Sáez-Marti and Sjögren (2008) model the merit-guided learning on the part of children. Our analysis complements these studies by exploring the transmission of moral values and their role in determining criminal behavior.

not commit crime if it is not economically profitable).[7] Moreover, and importantly, the interplay between economic incentives and cultural transmission implies that parents may deliberately transmit the bad trait to their children as a by-product of their own behavior even though they agree that it is not the desired one. This case is novel to the literature.[8]

Another important ingredient of our model is the assumption that cultural transmission may be biased in favor of one particular trait or depends on the relative frequency of the trait in the population (see Boyd and Richerson (1985), Bisin and Verdier (2001, sec.2.2.2) and Sáez-Marti and Sjögren (2008)). The main idea behind this kind of oblique transmission process is that children learn from a large group of randomly selected peers.[9] Depending on the children's bias, our model predicts different equilibrium outcomes: (i) With unbiased or positive biased oblique transmission (in favor of the dishonest type), the unique possible steady state is the one in which all individuals acquire the undesired trait even though all parents agree on which trait is desirable. In this case the parents' criminal behavior and the social environment reinforce each other, implying the extinction of the honest type in the long-run. (ii) With frequency-dependent or negative bias, however, both traits can co-exist in equilibrium even if all parents commit crime. Moreover, multiple interior equilibria with either high or low crime rates are shown to exist, so that initial conditions regarding the distribution of cultural traits may determine whether there will be diversity or assimilation of types in equilibrium.

In contrast to the existing literature, our theory can thus explain why criminal behavior persists even though parents agree that norms of good conduct (honesty) are desirable. Furthermore, we show that a public education campaign not only implies the co-existence of honest and dishonest types in society - independent of the children's bias - but also that it is a powerful policy instrument to fight criminal behavior.

Our model provides a novel explanation of why crime is highly concentrated in specific areas (the phenomenon of urban ghettos, see e.g. Grogger and Willis, 2000). In fact, there are many cases of 'twin' cities sharing

[7] Conley and Wang (2006) use a similar approach to model the crime decision taking into account the degree of honesty of individuals. In their model, however, moral values are exogenously given.

[8] The main idea is also captured by Vito Corleone's initial quote from the famous movie *The Godfather*. Even though Vito never wanted his son Michael to be involved in the family's criminal enterprize, and actually hoped he would go into politics, in the end, Michael could not escape the criminal influence of his family.

[9] The importance of peer and neighbor effects in determining criminal behavior has recently been emphasized by Glaeser et al. (1996), Bayer et al. (2009) and Damm and Dustmann (2014).

similar characteristics but still exhibiting very different levels of crime in the United States. For example, the property crime rate is 60% higher in Minneapolis than in St-Paul, 100% higher in Tampa than in St Petersburg, and 46% higher in Oakland than in San Francisco (Marceau and Mongrain, 2011). Existing explanations for such a concentration of crime focus primarily on externalities related to the number of criminals in specific areas (e.g. Freeman et al. (1996), Conley and Wang (2006), Ferrer (2010)) or on some form of social interactions with peers and neighbors (see e.g. Glaeser et al. (1996), Ballester et al. (2010), Acemoglu and Jackson (2014)).[10] Alternative explanations of generating disparities in crime rates within specific areas range from labor market frictions (Burdett et al., 2003), to discrimination and prejudices (Verdier and Zenou, 2004, O'Flaherty and Sethi, 2010). Our model complements these studies by explicitly modeling the role of socialization within the family as a crucial determinant of criminal behavior which has been neglected so far. In fact, our analysis is the first to formally explore the intergenerational nature of criminal behavior.

The remainder is organized as follows. Section 2 provides some further motivation on the intergenerational nature of criminal behavior. Section 3 develops the basic model, establishes the existence of multiple steady states and presents comparative static results. Section 4 introduces a public education campaign into the basic framework and analyzes the consequences for the cultural transmission process and the existence of crime equilibria. Section 5 presents several extensions and robustness checks related to the main assumptions of our theory. Section 6 concludes.

2 The intergenerational nature of crime: Explanations

A substantial body of literature has documented that children with criminal parents are more likely to become criminal themselves.[11] However, the underlying mechanism that generates this intergenerational association in crime remains an open question. Both criminologists and economists have put forward several possible explanations. According to Farrington (2011), for example, the intergenerational transmission of criminal behaviour may be explained by the exposure to multiple risk factors (such as poverty, fa-

[10] See Zenou (2003) for an overview.

[11] Thorough reviews of the criminological literature are provided by Thornberry (2009) and Farrington (2011). From an economic perspective, intergenerational criminal correlations have e.g. been studied by Case and Katz (1991), Duncan et al. (2005), Hjalmarsson and Lindquist (2012) and Hjalmarsson and Lindquist (2013).

milial problems, single or teenage parenting etc.), direct or indirect teaching (i.e. social learning) or genetic mechanisms (as e.g. mental disorders or low cognitive abilities). Similarly, Hjalmarsson and Lindquist (2013) distinguish between pre-birth and post-birth factors. While pre-birth factors comprise genetic influences, prenatal conditions and perinatal factors, post-birth factors refer to social mechanisms, behavioral mechanisms (as e.g. role modeling or socialization) or psychological mechanisms (e.g., childhood traumas and abuse) among others. Using Swedish adoption and police register data to examine parent–son associations in crime, Hjalmarsson and Lindquist (2013) provide evidence that pre-birth and post-birth factors are equally important in determining sons' convictions at the extensive margin, whereas at the intensive margin post-birth factors tend to dominate.

The intergenerational criminal correlations may, however, also stem from correlations in socioeconomic status across generations. For example, better educated parents tend to invest more into their children's education which may in turn reduce crime due to higher earnings in the legal sector and thus higher opportunity costs of committing crime for the children. In this respect, Meghir et al. (2012) exploit a major educational reform in Sweden to study the intergenerational effect of education on crime. They show that the reform lead to a substantial reduction in crime rates of both the targeted generation and their children. Their findings are most likely explained by improved family resources and better parenting through better role models and less parental criminal activities.[12]

The above explanations are not independent of each other and a combination of different mechanisms is likely to explain the intergenerational transmission of criminal behavior. However, for most of the mechanisms, intergenerational transmission should be stronger the more involved parents are in criminal activities or crime prone behavior. According to the social learning mechanism, for example, transmission is explained through direct interactions and influences with others, such as family members and peers. Parents serve as a role model for their children and may either directly or indirectly teach their children beliefs and moral values that in turn favor criminal behavior. The importance of parents as role models for their children in explaining the intergenerational associations in crime has been emphasized by Duncan et al. (2005) and Hjalmarsson and Lindquist (2012).

Moreover, according to Sutherland's Differential Association Theory, criminal behavior depends on norms which are acquired by children if such a be-

[12] An additional explanation of the intergenerational correlations in crime is related to mobility and neighborhood environments, see Ludwig et al. (2001) and Kling et al. (2005) who analyze the Moving to Opportunity program, which aim is to move very poor households to richer areas.

havior is more highly reinforced than noncriminal behavior (Akers and Jennings, 2009). Hence, the more involved parents are in criminal activities, the more opportunities to observe, imitate and learn their parents' delinquent behavior and motivations do children have. This, in turn, implies that there should be more offspring crime.[13] Still, it is important to note that existing studies find co-offending of parents and children to be unusual (West and Farrington (1977); Reiss and Farrington (1991)). Moreover, parents do not tend to encourage their children to become criminal (Rowe and Farrington, 1997, McCord, 2001). Yet, children may still observe and imitate parents behavior and adopt their moral values despite their parents disapproval.

The present paper provides a theoretical foundation of the intergenerational transmission of criminal behavior, and in particular the social learning mechanism. However, as the empirical literature to date has not yet reached a clear consensus about which of the above mechanisms are most relevant in explaining intergenerational associations in crime, the model developed in the preceding section accounts for both pre-birth (through differences in individuals abilities) and post-birth factors (through socialization processes within the family and interactions with peers). Furthermore, it yields several novel and empirically testable predictions with respect to parents' behavior and children's criminality.

3 The Basic Model

We consider a society populated by overlapping generations where the size of each generation is normalized to one. All agents live for two periods. When young (first period), individuals acquire their preferences; when old (second period), they become parents, have one child and decide how to split their time between joining the labor force and engaging in criminal activities.[14] Each old individual may either possess a high productivity θ^h or a low productivity θ^l to earn income in the legal market. We denote the time invariant share of individuals with low productivity in the economy by η.[15] The ability to commit crime is assumed to be the same for all individuals[16] and the net

[13] A similar hypothesis is also put forward by Thornberry (2005). As parents' criminality strongly affects their own development and parenting styles (and hence the quality of their parenting), the children's risk to develop criminal behavior is higher.

[14] Similar assumptions can, e.g., been found in Block and Heineke (1975) and Lochner (2004). Note further that we normalize the individuals' time endowment to one.

[15] Note that we implicitly assume that a parent with high (low) ability has a child with high (low) ability with probability one. This assumption is relaxed in Section 5.3.1.

[16] There is no empirical evidence of a clear positive correlation between the market productivity and the skills to commit a crime, that is, individuals with high productivity are

return from committing crime depends on economic factors (market income, apprehension probability, individuals' ability) and on moral norms of good conduct which are in turn determined by a transmission and socialization process from their parents. Hence, the main focus of our paper is on economic crimes, as e.g. theft, robbery, burglary or fraud. However, our model could as well easily be interpreted in terms of violent crimes, such as family violence, organized crime, etc.

In the following, we first consider how young agents adopt their moral values which in turn affect economic outcomes during old-age.

3.1 The transmission process

There are two different types of cultural traits in society, L and H. These types are referred to as dishonest and honest, respectively. Parents are altruistic and care about the type-dependent utility of their children. Consistent with the existing evidence (see e.g. Rowe and Farrington (1997)), we assume that all parents, independently of their own type, agree that one of the traits (honesty) is preferable.[17] Honest behavior and moral values of good conduct are thus considered to be vertically differentiated characteristics, similar to educational outcomes or good working habits. By contrast, traits like religion or ethnicity are typically modelled as being horizontally differentiated so that each parent prefers to transmit his own trait. Section 5.2 explores the robustness of our findings with respect to this assumption.

Socialization affects the adoption of traits only during childhood so that adult individuals keep the acquired trait throughout their lifetime. As in Bisin and Verdier (2000, 2001), the transmission of traits is modeled as a combination of socialization inside the family (vertical socialization, namely the parents' behavior) and socialization outside the family (oblique socialization, namely the social environment where children live). However, in contrast to the existing theoretical literature, but consistent with the evidence cited in the previous section, we posit that the parents' behavior (the decision to commit crime) rather than effort has a direct positive effect on the probability of the child adopting the bad type through vertical transmis-

not more productive in the crime sector (see for example Lochner (2004) and Lochner and Moretti (2004)).

[17] Indirect evidence in favor of this assumption also comes from survey data. In the response to NORC's General Social Survey's question, 'Which three of the qualities listed would you say are the most desirable for a child to have?', 'honesty' is the most cited quality across the sample (Bisin and Verdier, 2011, p.394).

sion,[18] e.g. through social learning or adopting parental role models.[19] Still, the probability of adopting a specific type also depends on peer group effects and thus on socialization by society through a process of oblique transmission. The importance of peer group effects in determining criminal behavior has been emphasized, e.g., by Damm and Dustmann (2014), Bayer et al. (2009) and Glaeser et al. (1996).

More formally, let μ_t $(1-\mu_t)$ be the proportion of L-type (H-type) adults and $x_t^{L,k} \in [0,1]$ ($x_t^{H,k} \in [0,1]$) the fraction of time a L-type (H-type) parent with ability $k = h,l$ devotes to criminal activities in period t. Then, the total probability that the child of an L-type (H-type) adopts L (resp. H) is given by:

$$P_k^{LL} = x_t^{L,k} + (1-x_t^{L,k})S(\mu_t) \tag{1}$$

$$P_k^{LH} = (1-x_t^{L,k})(1-S(\mu_t)) \tag{2}$$

$$P_k^{HH} = (1-x_t^{H,k})(1-S(\mu_t)) \tag{3}$$

$$P_k^{HL} = x_t^{H,k} + (1-x_t^{H,k})S(\mu_t) \tag{4}$$

where $S(\mu_t)$ captures the process of oblique transmission, namely how children are influenced by society (peers). Before discussing the properties of the transmission function $S(\mu_t)$ in more detail, we first interpret equations (1)-(4). The child of a dishonest parent will also be dishonest with probability equal to the parents' time spent on criminal activities (eq.1). If this direct transmission fails (with probability $1-x_t^{L,k}$), the child acquires the dishonest trait from his/her neighborhood (with probability $S(\mu_t)$). The probability that a child of dishonest parents becomes honest is defined by equation (2). This may only happen if the child does not acquire the bad trait from either his/her parents or his/her peers. For honest parents (equations (3) and (4)), the interpretation is similar. Note further that an increase in the parents' criminal activities unambiguously lowers the probability of the children to adopt the good trait, i.e. $\partial P^{iH}/\partial x_t^{i,k} < 0$, $i = L,H$, $k = h,l$.[20]

[18] **Not only the parents' behavior has a direct impact on the socialization process, but also all related aspects that are involved, for example familial problems due to arrested fathers, etc. –> We should think of more example here!! So our idea is that every aspect of the parents' criminal behavior has an impact on the children's socialization, not only the observed behavior itself!**

[19] Section 5.1 extends the basic model to allow for a socialization process which depends on both effort and the parents' behavior.

[20] Note that, for reasons of simplicity, we assume that children can perfectly observe their parents' criminal activities. The qualitative results would be the same, however, if children could only observe a fraction of their parents' criminal activities, i.e. $\epsilon x_t^{i,k}$, $i = L,H$, $k = h,l$ and $\epsilon \in (0,1)$ in (1)-(4). Furthermore, our qualitative results are also robust against

The oblique transmission function $S : [0,1] \to [0,1]$ is assumed to be twice continuously differentiable and increasing such that $S(0) = 0$ and $S(1) = 1$. The standard assumption in the literature is that children are randomly matched to one role model who provide them the trait to copy. This results in an unbiased oblique transmission function, $S(\mu_t) = \mu_t$, so that children acquire each of the traits with a probability equal to their share in the population. Following Sáez-Marti and Sjögren (2008), Bisin and Verdier (2001, sec.2.2.2) and Boyd and Richerson (1985), however, we assume that oblique transmission may well be biased. Such biases result if the peer group children interact with is of fixed size and consists of randomly-drawn individuals from the whole population, and children evaluate the relative merit of the variants of traits observed in this group. Specifically, three different cases are distinguished in the sociobiological literature:

- Positive bias: the probability that the naive agent adopts type L is always larger than if he had acquired one role model randomly, $S(\mu_t) > \mu_t$ for all $\mu_t \in (0,1)$.
- Negative bias: the probability that the naive agent adopts type L is always smaller than if he had acquired one role model randomly, $S(\mu_t) < \mu_t$ for all $\mu_t \in (0,1)$.
- Conformism or frequency-dependent bias: when the frequency of trait L in the population is smaller (larger) than $\hat{\mu}_t$, the probability that a naive agent adopts L is decreased (increased) relative to the unbiased transmission, $S(\mu_t) \gtreqless \mu_t$ for $\mu_t \gtreqless \hat{\mu}_t$. 'Pure' conformism corresponds to $\hat{\mu}_t = 1/2$.

A graphical representation of the different cases is provided by Figure 1.

[Figure 1 about here]

3.2 The parents' decision

We now turn to the analysis of the parents' decision to engage in criminal activities. This decision depends on economic incentives as well as on type-dependent moral costs, γ^i, $(i = H, L)$. Furthermore, as explained in the previous section, parents are altruistic and care about their children's type-dependent utility.

assuming that the transmission process is affected be the parents' gains from crime instead of the time spent on criminal activities.

Risk neutral parents maximize utility

$$U_t^{i,k} = (1-x_t^{i,k})(1-\lambda(1+\phi))\theta^k + \rho\lambda x_t^{i,k}\tilde{\theta} - (1-\rho)\psi - \gamma^i \frac{(x_t^{i,k})^2}{2} + \beta\left(P_k^{iH}V^H + P_k^{iL}V^L\right) \tag{5}$$

subject to equations (1)-(4) by choosing the fraction of time devoted to criminal activities $x_t^{i,k}$, $(i = H,L,\ k = h,l)$. This utility function has a standard cost/benefit structure established by the seminal contribution of Becker (1968). The benefits from the criminal activity are given by $\rho\lambda x_t^{i,k}\tilde{\theta}$ and are increasing in the average productivity in the legal market $\tilde{\theta} = \eta\theta^L + (1-\eta)\theta^H$, the fraction of income that can be stolen λ and the probability of not being apprehended ρ.[21]

The costs of committing crimes are measured by the probability of being caught $1-\rho$ times the fine ψ (which is a minimum level of income in case a criminal is caught) and the opportunity costs of forgone earnings in the legal sector $(1-\lambda(1+\phi))\theta^k$ $(k = L,H)$. The parameter $\phi \in [0,(1-\lambda)/\lambda[$ captures additional monetary costs incurred when being victimized, e.g. broken windows or doors, time spent reporting the crime to the police, etc. Also, as in Conley and Wang (2006) and Verdier and Zenou (2004), agents have a type-dependent moral cost of committing crime equal to $\gamma^i(x_t^{i,k})^2/2$ $(i = H,L)$, where γ^i captures their degree of honesty. So the higher γ^i and the crime effort, the higher the moral cost.

Finally, the last term of the utility function captures the altruistic concern and β the degree of altruism. Parents care about the well-being of their children in adult life, i.e. V^H, V^L with $V^H \geq V^L$, which is in turn affected by moral concerns regarding criminal behavior and thus by the type adopted by the child. More specifically, an increase in the fraction of time spent on criminal activities enhances the probability of the child adopting the L-type and thus reduces the utility from the altruistic concern.[22]

Solving the parents' maximization problem gives the optimal fraction of time spent on criminal activities:

$$x_t^{i,k} = \frac{1}{\gamma^i}\left(\rho\lambda\tilde{\theta} - (1-\lambda(1+\phi))\theta^k - \beta(1-S(\mu_t))\Delta V\right) \tag{6}$$

with $k = h,l$ and $i = H,L$. Clearly, the optimal share $x_t^{i,k}$ is increasing in

[21] For technical simplicity, we have assumed that each agent is equally sensitive to being the victim of a crime and that each criminal steals the same amount of income per unit of time devoted to criminal activities in the economy, that is, criminals cannot monitor and pick up their victims based on their income (see Imrohoroglu et al. (2004, 2006) and Bethencourt (2014)).

[22] In fact, it is straightforward to prove that $P_k^{iH}V^H + P_k^{iL}V^L$ is decreasing in $x_t^{i,k}$ $(i = H,L,\ k = h,k)$.

the average productivity in the legal market $\tilde{\theta}$, the share of income that can be stolen λ and the probability of not being apprehended ρ whereas it is decreasing in individuals' productivity in market activities θ^k.

We also we obtain some additional and novel empirical testable predictions with respect to the variables affecting the socialization mechanism: The probability with which children acquire the honest trait from their neighborhood, $1-S(\mu_t)$, increases the level of criminal activities. The intuition behind this findings is that a higher share of dishonest types in society reduces the probability of transmitting the desired trait (for a given level of $x_t^{i,k}$) and thus lowers the marginal costs of committing crime. Consequently, horizontal transmission and parents' time spend on criminal activities are complementary. Furthermore, an increase in the parents' relative utility of having a type H child, $\Delta V \equiv V^H - V^L$, decreases criminal activities because the value of transmitting the desired trait increases.

3.3 Dynamics and steady states

Given equation (6), the dynamics of the population of agents with type L are then determined by the following difference equation:

$$\begin{aligned}\mu_{t+1} =& \mu_t \eta P_l^{LL} + \mu_t(1-\eta)P_h^{LL} + (1-\mu_t)\eta P_l^{HL} + (1-\mu_t)(1-\eta)P_h^{HL} \\ =& S(\mu_t) + \left[\tilde{y} - \beta\Delta V(1-S(\mu_t))\right]\left[\frac{\mu_t}{\gamma^L} + \frac{1-\mu_t}{\gamma^H}\right](1-S(\mu_t)) \end{aligned} \tag{7}$$

where $\tilde{y} = \tilde{\theta}(\rho\lambda - (1-\lambda(1+\phi)))$. The change in the fraction of L-types can be obtained from equation (7) as:

$$\triangle \mu_{t+1} = S(\mu_t) - \mu_t + \Gamma(\mu_t) \tag{8}$$

with

$$\Gamma(\mu_t) = \left[\tilde{y} - \beta\Delta V(1-S(\mu_t))\right]\left[\frac{\mu_t}{\gamma^L} + \frac{1-\mu_t}{\gamma^H}\right](1-S(\mu_t)) \tag{9}$$

It is straight forward to see that $\Gamma(\mu_t) \geq 0$ for all $\mu_t \in [0,1[$. In the following, we are looking for conditions under which the different traits coexist in equilibrium even if all parents agree that the honest trait is preferable but, at the same time, parents may devote some (individuum specific) fraction of their time to commit crime.

We denote by $\mu(t,\mu_0)$ the path produced by equation (7) when the initial condition is μ_0, M the set of steady states and $\bar{x} \equiv (\tilde{y} - \beta\Delta V)/\gamma^H$ the average time honest individuals spend on criminal activities when $\mu_t = 0$.

Proposition 1 ***Existence of steady states.***

(i) $\mu^* = 1 \in M$.

(ii) *If oblique transmission is unbiased or positively biased (in favor of trait L), then* $M = \{1\}$ *and* $\mu(t, \mu_0) \to 1 \ \forall \mu_0$.

(iii) *If oblique transmission is negatively biased (against trait L), two cases may arise:*

- *For* $S'(1) > \frac{1}{1-\tilde{y}/\gamma^L}$, *then there exists at least one interior equilibrium* μ^* *with* $\mu^* \in [0,1)$ *and* $\mu(t, \mu_0) \to \mu^* \ \forall \mu_0 \neq 1$.
- *For* $S'(1) < \frac{1}{1-\tilde{y}/\gamma^L}$ *and* $\bar{x}$ *small enough, then there exist at least two interior equilibria,* μ_1^*, μ_2^*, *with* $\mu_1^* < \mu_2^* < 1$, *such that:* $\mu(t, \mu_0) \to \mu_1^* \ \forall \mu_0 < \mu_2^*$ *and* $\mu(t, \mu_0) \to 1 \ \forall \mu_0 > \mu_2^*$. *For* $\bar{x}$ *large enough,* $M = \{1\}$ *and* $\mu(t, \mu_0) \to 1 \ \forall \mu_0$.

(iv) *If oblique transmission is conformist and* $\bar{x}$ *is small enough, then there exist at least two interior equilibria,* μ_1^*, μ_2^*, *with* $\mu_1^* < \mu_2^* < 1$, *such that:* $\mu(t, \mu_0) \to \mu_1^* \ \forall \mu_0 < \mu_2^*$ *and* $\mu(t, \mu_0) \to 1 \ \forall \mu_0 > \mu_2^*$. *For* $\bar{x}$ *large enough,* $M = \{1\}$ *and* $\mu(t, \mu_0) \to 1 \ \forall \mu_0$.

Proof: *See appendix.*

Proposition 1 establishes that both traits can only survive if the desired trait is easy to adopt through oblique transmission. Otherwise, even though parents agree that honesty is desirable, as long as committing crime is economically profitable, the desired trait disappears. Figure 2 illustrates the cultural dynamics for two different cases of proposition 1: (i) when the oblique transmission function is negatively biased and there exists one interior rest point and (ii) when there is conformism and there exist two interior steady states. In all cases we have represented $\triangle\mu$ as a function of μ. We denote all steady states: stable equilibria are marked with green circles and unstable ones with red circles. Clearly, the honest type can survive even in a situation in which all parents commit crime:

First, when the interaction with peers is negatively biased against acquiring trait L, both traits L and H will survive in equilibrium. Second, when the oblique transmission function is characterized by conformism, an equilibrium with diversified culture characterized by a low share of the L-trait exists. However, if the starting share of trait L in the population is sufficiently large, then trait H will disappear. Notice that in both cases (with conformism or a negative bias), multiple interior stable steady states can emerge. Such a result explains why - depending on initial conditions

- the economy may reach an equilibrium of assimilation or an equilibrium with cultural diversity. This has important implications for the analysis of suitability and effectiveness of deterrence policies against crime or policies devoted to modify the values that children assign to different traits.

[Figure 2 about here]

It is now interesting to see how the possible equilibrium multiplicity of types translates into different individuum specific crime rates. The aggregate share of criminal activities in period t is given by

$$\begin{aligned}\tilde{x}_t =& \mu_t \eta x_t^{L,L} + \mu_t(1-\eta)x_t^{L,H} + (1-\mu_t)\eta x_t^{H,L} + (1-\mu_t)(1-\eta)x_t^{H,H} \qquad (10)\\ =& \left(\tilde{y} - \beta\Delta V(1-S(\mu_t))\right)\left(\frac{(1-\mu_t)}{\gamma^H} + \frac{\mu_t}{\gamma^L}\right)\end{aligned}$$

Clearly, this share is monotonically increasing in the share of dishonest individuals μ_t in the economy. At the individual level, however, the contributions of each group to the aggregate share of criminality differ substantially. Figure 3 illustrates the population weighted crime rates (i.e. each of the four summands in (10)) as a function of μ and indicates the resulting steady state levels when there is either a negative bias or conformism. We observe that, despite the overall positive relationship between $\tilde{x}_t$ and μ_t, the contribution of groups with the honest trait is monotonically declining: The positive effect on crime through increases in the share of dishonest individuals is not large enough to offset the reduction in the population weight of these groups.

The multiplicity of interior steady states provides a novel and complementary explanation for understanding the existence of differentials in criminal activities across place and time (see e.g. Glaeser (1998) and Glaeser and Sacerdote (1999)). In fact, there exists considerable variation in US crime rates between cities with similar characteristics. Moreover, center cities showing more criminality than suburbs. For example, the property crime rate in center cities is three to two times larger than that in the suburbs (Grogger and Willis, 2000).

[Figure 3 about here]

3.4 Comparative static analysis

Let us now consider some comparative static results. We are in particular interested in changes in the apprehension probability ρ (e.g. through an

increase in crime deterrence expenditures) and in changes in the relative evaluation of having an honest child ΔV (e.g. by increasing parents' awareness of the importance of honest behavior).

Consider first changes in ρ. Figures 4 and 5 illustrate the cultural dynamics resulting from different levels of ρ and the corresponding average shares of criminal activities both under negative and frequency-dependent transmission. If the apprehension probability is low (solid lines), $\mu = 1$ is the only stable equilibrium and average steady state crime levels are high. Increasing the apprehension probability clearly not only has a direct positive effect on crime reduction by lowering economic incentives (the standard deterrence effect)[23] but also a social multiplier effect resulting from changes in the number and properties of steady state equilibria. This amplifying effect emerges as a decrease in the share of criminal activities increases the probability of individuals adopting the honest trait. The larger is the share of honest individuals in the economy, the lower is the share of criminal activities. This positive feedback process terminates at a new steady state with a lower share of dishonest individuals in the population and a lower share of criminal activities. Graphically, the standard effect implies the downwards shift of the curve $\tilde{x}(\mu)$, while the multiplier effect reflects the transition to a new lower stable steady state determined by the shifted curve. If the increase in ρ is sufficiently large, economic conditions prevent parents from committing crime which in turn implies that $\mu = 0$ is a stable steady state. The reason is that, with a negative bias, children always have a higher probability of adopting the good type whereas it depends on the relative frequency of types with conformism.[24]

[Figures 4 and 5 about here]

Now, consider an increase in the parents' awareness of the relevance of the honest trait, i.e. ΔV. With a negative bias this always increases the share of honest types in the steady state and decreases the average share

[23] In fact, there is a long-established relationship between the probability of apprehension and criminal outcomes dating back to the seminal contribution by Becker (1968). Moreover, this relationship has found strong empirical support in many countries (see e.g. Di Tella and Schargrodsky (2004), Draca et al. (2011) and Levitt (2004)) For the US, Imrohoroglu et al. (2004, 2006) and Engelhardt et al. (2008) use calibration methods and data on property crimes to show that the apprehension probability is one of the most important factors to explain variations over time in thee crime rates.

[24] These comparative static results are similar to the ones in Funk (2005), who studies a model in which moral norms are shaped by social interactions with peers. In contrast to the present paper, however, there is no direct role of socialization mechanisms or the parents' criminal behavior in her model.

of criminal activities. Figure 6 illustrates the cultural dynamics resulting from different levels of the the parents' awareness and the corresponding average crime rate under negative bias. In particular, for $\Delta V = 0$ (solid line), the crime rate is independent of μ as the parents' decision to commit crime has no effect on the transmission process. By contrast, if relative differences of type evaluations are sufficiently large, parents' concerns about their children's well being prevents them from committing crime, which in turn implies that $\mu = 0$ is a stable equilibrium. With conformism (see figure 7), the effects of increasing parents' awareness are qualitatively similar to reductions in ρ. The same conclusion can also be reached for the comparative static effect of decreasing the share of income that can be stolen, λ. The important policy implication is that crime deterrence policies and information/education policies shaping the evaluation of types by parents are substitutes in fighting criminal behavior.[25]

[Figures 6 and 7 about here]

Finally, we consider two exercises related to income growth and inequality. The first one consists of an increase in the average market income through increases in individuals' productivities (θ^H and θ^L) while leaving the level of inequality (the skill premium θ^H/θ^L) unchanged. Such an increase in the average income has a direct positive effect on criminal activities by increasing the net return from committing crime.[26] However, like in previous exercises, there also exists a social multiplier effect derived from the cultural transmission mechanism as a higher share of criminal activities increases the children's probability of adopting the dishonest behavior. As a result, a new steady state with a higher share of dishonest individuals in the population and a higher crime rate arises. Graphically, these changes could be plotted as in figures 6 and 7 but with curves moving into opposite directions. The finding that a higher average income increases crime stands in contrast to the theoretical predictions in Conley and Wang (2006), who show that income growth that affects all members of a society equally has no impact on the equilibrium level of crime. One important difference between the two models, however, is that moral values are endogenously determined in the present paper whereas they exogenously given in Conley and Wang (2006).

The second exercise consists of an increase in individuals' productivities

[25] See also section 3.

[26] The marginal cost of committing a crime increases less than the marginal benefit as the part of the marginal cost determined by the cultural transmission process remains fixed (see equation (6)).

such that the skill premium θ^H/θ^L (the level of inequality) rises while the average productivity remains unchanged. In this case, net benefits of committing crime increase for low ability individuals, thus rising their time share spent on criminal activities, but decrease for high ability individuals, which in turn reduces their criminal activities. Notice that for a sufficiently large increase in inequality, highly productive individuals will quit their criminal activities. Up to this point, large increases in inequality would thus increase the incentives of low ability agents to commit more crimes and so the aggregate crime rate in the economy would increase. Related figures of such changes would be similar to those derived from an increase in the average income. These predictions provide a theoretical explanation for a positive relationship between property crime and income inequality which is generally supported by empirical evidence (see e.g. Freeman (1996, p.33), Demombynes and Oezler (2005) and references therein).[27] Furthermore, our theoretical findings are in line with Chiu and Madden (1998) who show that an increase in inequality unambiguously increases property crime. In their model, in which moral considerations are absent, income inequality affects property crime by rendering legal work less attractive for poor individuals and by increasing the potential proceeds from crime.

4 Public education campaign

This section analyzes the effectiveness of public education campaigns used to emphasize the importance of norms of good conduct.[28] Following Hauk and Sáez-Marti (2002), we assume that children are first exposed to the influence of their parents before undergoing public education. Hence, only children who have not adopted their preferences through direct socialization by their parents can be affected by public education. An education campaign consists of a publicly chosen effort level $\kappa \in [0,1]$ which is assumed to be equal to the probability with which a child adopts honest preferences in school.[29] Public

[27] Note, however, that few studies also report a negative relationship between inequality and crime, see e.g. Brush (2007) and Chintrakarn and Herzer (2012).

[28] See Lochner (2011) for a survey on the relationship between crime and education.

[29] Clearly, we make two simplifying assumptions: first, we assume that the education campaign only affects the cultural transmission process but leaves individuals' productivity unchanged. Allowing for an explicit process of human capital formation is beyond the scope of this paper. Second, we assume that the public education campaign is exogenously given without stating how it is financed. This is not restrictive, however, since we might assume that the required tax revenue is collected by a lump sum tax. Extending the present model to capture general equilibrium effects arising from the public provision of education is an interesting topic left for future research.

education efforts affect the probabilities of honest and dishonest children as follows:[30]

$$P_k^{LL} = x_t^{L,k} + (1 - x_t^{L,k})S(\mu_t)(1-\kappa) \tag{11}$$

$$P_k^{LH} = (1 - x_t^{L,k})((1 - S(\mu_t))(1-\kappa) + \kappa) \tag{12}$$

$$P_k^{HH} = (1 - x_t^{H,k})((1 - S(\mu_t))(1-\kappa) + \kappa) \tag{13}$$

$$P_k^{HL} = x_t^{H,k} + (1 - x_t^{H,k})S(\mu_t)(1-\kappa) \tag{14}$$

As before, the parents' time spent on criminal activities $x_t^{i,k}$, ($i = H,L$, $k = h,l$), determines the probability of children adopting the dishonest trait (eqs. (11) and (14)). With the complementary probability $1 - x_t^{i,k}$ children remain naive and acquire the bad trait through society (with probability $S(\mu_t)$) given that public education fails (with probability $1-\kappa$). By contrast, a child will be honest if it does not acquire the bad trait from either his/her parents or, if public education fails, from his/her peers (eqs. (12) and (13)).

The parents' optimal fraction of time spent on criminal activities is now:

$$x_t^{i,k} = \frac{1}{\gamma^i}\left(\rho\lambda\tilde{\theta} - (1 - \lambda(1+\phi))\theta^k - \beta(1 - S(\mu_t)(1-\kappa))\Delta V\right) \tag{15}$$

with $k = h,l$ and $i = H,L$. The new change in the fraction of L-types with public education is given by:

$$\triangle\,\mu_{t+1} = S(\mu_t)(1-\kappa) - \mu_t + \Gamma(\mu_t) \tag{16}$$

with

$$\Gamma(\mu_t) = \left[\tilde{y} - \beta\Delta V(1 - S(\mu_t)(1-\kappa))\right]\left[\frac{\mu_t}{\gamma^L} + \frac{1-\mu_t}{\gamma^H}\right](1 - S(\mu_t)(1-\kappa)) \tag{17}$$

It is straight forward to see that $\Gamma(\mu_t) > 0$ for all $\mu_t \in [0,1]$ if each individual spends at least some time committing crime. The introduction of public education has two effects: its direct effect is to increase the proportion of honest agents, while its indirect effect is to decrease parents' time spent on criminal activities which in turn reinforces the direct effect. Note further that $\triangle\mu_{t+1} > 0$ if $\mu_t = 0$ and $\triangle\mu_{t+1} < 0$ if $\mu_t = 1$. These observations imply:

[30] Another interesting socialization mechanism, which would yield qualitatively similar results, is $P_k^{HH} = \kappa(1 - x_t^{H,k}) + (1 - \kappa(1 - x_t^{H,k}))(1 - S(\mu_t))$, $P_k^{LL} = (1 - \kappa(1 - x_t^{L,k}))S(\mu_t)$, $P_k^{LH} = \kappa(1 - x_t^{L,k}) + (1 - \kappa(1 - x_t^{L,k}))(1 - S(\mu_t))$ and $P_k^{HL} = (1 - \kappa(1 - x_t^{H,k}))S(\mu_t)$, which corresponds to the case in which children are simultaneously exposed to public education and their parents' behavior, i.e. $\kappa(1 - x_t^{i,k})$, $i = H,L$. If this first direct socialization process fails, children acquire their trait from their neighborhood.

Proposition 2 *Suppose that the government runs a public education campaign, i.e.* $\kappa \in (0,1]$. *Then, there exists at least one interior equilibrium* μ^* *such that* $\mu(t,\mu_0) \to \mu^*$ *for all* μ_0.

The above analysis establishes the existence of at least one interior steady state such that honest and dishonest types, and therefore groups with high and low criminality, co-exist in society. Importantly, with a public education campaign, the existence of interior rest points no longer depends on the children's bias. In fact, even with unbiased horizontal transmission (i.e. $S(\mu_t) = \mu_t$) culture remains diverse. However, as in proposition 1, the functional form of S still affects the number and stability of interior equilibria. Figure 8, for example, illustrates the cultural dynamics and the average share of criminal activities with conformism for different intensities of the education campaign. The solid lines correspond to $\kappa = 0$. Clearly, if society is initially in the high crime steady state, an intensive education campaign with a high enough κ is successful in fighting crime as it affects affects the population dynamics and the proportion of honest individuals increases.

The crime reducing effect of education is well documented by empirical evidence, see Lochner and Moretti (2004), Lochner (2004), Deming (2011), Meghir et al. (2012) and Anderson (2014). Lochner and Moretti (2004), for example, find that if the average years of schooling increase by one year then, both violent and property crime decline by about 11-12 percent, while Deming (2011), using data from public school choice lotteries, shows that peer effects may account for gains in school quality implying a significant decline in criminality. Finally, Anderson (2014) estimates that if the the compulsory schooling age increases from 16 to 17 or 18 years of age then, arrests at these ages reduce by nearly 10 percent, with similar impacts on both property and violent crime.

[Figure 8 about here]

5 Robustness

We consider several robustness checks in order to demonstrate that our findings are robust to assumptions and choices made. More precisely, we introduce an explicit socialization effort into the basic model (Section 5.1) and consider an alternative transmission process in which each parent prefers the transmission of his own trait (Section 5.2). Finally, we discuss the importance of some specific modeling assumptions.

5.1 Socialization effort vs. criminal behavior

In order to reconcile our approach with the standard assumption in the cultural transmission literature, i.e. that socialization requires parental effort, we now consider a transmission process which is influenced by both the parents' behavior and effort to transmit the desired trait. Then, there are two competing channels: As before, the parents' criminal behavior lowers the likelihood that children obtain the desired trait, whereas parental effort has a countervailing effect and increases the same probability. Formally, such a transmission process can be defined as follows:

$$P_k^{LL} = \alpha x_t^{L,k} + [\alpha(1-x_t^{L,k}) + (1-\alpha)(1-\tau_t^{L,k})]S(\mu_t) \tag{18}$$

$$P_k^{LH} = (1-\alpha)\tau_t^{L,k} + [\alpha(1-x_t^{L,k}) + (1-\alpha)(1-\tau_t^{L,k})]((1-S(\mu_t)) \tag{19}$$

$$P_k^{HH} = (1-\alpha)\tau_t^{H,k} + [\alpha(1-x_t^{H,k}) + (1-\alpha)(1-\tau_t^{H,k})]((1-S(\mu_t)) \tag{20}$$

$$P_k^{HL} = \alpha x_t^{H,k} + [\alpha(1-x_t^{H,k}) + (1-\alpha)(1-\tau_t^{H,k})]S(\mu_t) \tag{21}$$

where $\alpha \in (0,1)$ measures the relative importance of the parents' actual behavior and their own socialization effort in determining the horizontal transmission probabilities. The case $\alpha = 1$ corresponds to the basic model of section 3. By contrast, a socialization process with $\alpha = 0$ has e.g. been analyzed by Sáez-Marti and Zenou (2012) among others.

Parents maximize their utility by choosing the time allocated to criminal activities ($x_t^{i,k}$, as before) and to socialize their children ($\tau_t^{i,k}$, with $k = h,l$ and $i = H,L$). We assume that educating ones child has a cost beyond the missed income from working in the legal market given by $c(\tau_t^{i,k}) = c(\tau_t^{i,k})^2/2$. Hence, parents' utility function is given by

$$U_t^{i,k} = (1-x_t^{i,k}-\tau_t^{i,k})(1-\lambda(1+\phi))\theta^k + \rho\lambda x_t^{i,k}\tilde{\theta} - (1-\rho)\psi - \gamma^i\frac{(x_t^{i,k})^2}{2} + \beta\left(P_k^{iH}V^H + P_k^{iL}V^L\right) - c\frac{(\tau_t^{i,k})^2}{2} \tag{22}$$

The optimal fractions of time spent on criminal activities and education are:

$$x_t^{i,k} = \frac{1}{\gamma^i}\left(\rho\lambda\tilde{\theta} - (1-\lambda(1+\phi))\theta^k - \alpha\beta(1-S(\mu_t))\Delta V\right) \tag{23}$$

$$\tau_t^k = \frac{1}{c}\left((1-\alpha)\beta S(\mu_t)\Delta V - (1-\lambda(1+\phi))\theta^k\right) \tag{24}$$

with $k = h,l$ and $i = H,L$.[31] Parents' incentive to invest in education increases with the number of dishonest individuals in society, μ_t, the rela-

[31] In the following we assume $x_t^{i,k} + \tau_t^k < 1$ holds for all individuals. This requires, e.g., some restrictions on γ^i and c.

tive evaluation of having an honest child, ΔV, and the parents degree of altruism, β.[32] By contrast, it is decreasing in forgone earnings in the legal sector (the opportunity costs of educating children), $(1-\lambda(1+\phi))\theta^k$. Finally, if the share of dishonest individuals is sufficiently small, i.e. $\mu_t < S^{-1}((1-\lambda(1+\phi))\theta^k/((1-\alpha)\beta\Delta V))$, parents will free-ride on trait transmission by society and not exert any effort to educate their child.

The change in the fraction of L-types is given by:

$$\begin{aligned}\triangle\mu_{t+1} &= \alpha S(\mu_t) - \mu_t \\ &+ \alpha(1-S(\mu_t))\left(\frac{\mu_t}{\gamma^L} + \frac{1-\mu_t}{\gamma^H}\right)\left(\rho\lambda\tilde{\theta} - (1-\lambda(1+\phi))\tilde{\theta} - \alpha\beta(1-S(\mu_t))\Delta V\right) \\ &+ (1-\alpha)S(\mu_t)\frac{1}{c}\left((1-\alpha)\beta S(\mu_t)\Delta V - (1-\lambda(1+\phi))\tilde{\theta}\right) \qquad (25)\end{aligned}$$

with $\triangle\mu_{t+1} = \bar{x}^{\alpha} > 0$ if $\mu_t = 0$ and $\triangle\mu_{t+1} = (1-\alpha)(\bar{\tau}-1) < 0$ if $\mu_t = 1$.[33] From these observations it follows immediately:

Proposition 3 *Suppose that $\alpha \in (0,1)$. Then, there exists at least one interior equilibrium μ^* such that $\mu(t,\mu_0) \to \mu^*$ for all μ_0.*

Two remarks are in order. First, similar to the situation with a public education campaign, the existence of interior rest points no longer depends on the children's bias. Second, the introduction of parental effort rules out the steady state in which all individuals end up with the dishonest trait in the long-run. The intuition behind this findings is that effort provides a countervailing force to parents' criminal behavior in determining horizontal transmission.

Altogether, our main results turn out to be robust against the introduction of effort into the basic model.

5.2 An alternative transmission process

$$P_k^{LL} = x_t^{L,k} + (1-x_t^{L,k})S(\mu_t) \qquad (26)$$

$$P_k^{LH} = (1-x_t^{L,k})((1-S(\mu_t)) \qquad (27)$$

$$P_k^{HH} = (1-x_t^{H,k}) + x_t^{H,k}((1-S(\mu_t)) \qquad (28)$$

$$P_k^{HL} = x_t^{H,k}S(\mu_t) \qquad (29)$$

[32] Note that the comparative static analysis for the share of time spent on criminal activities is basically the same as in section 3.

[33] Note that $\bar{x}^{\alpha} \equiv \frac{1}{\gamma^H}\left(\rho\lambda\tilde{\theta} - (1-\lambda(1+\phi))\tilde{\theta} - \alpha\beta\Delta V\right)$ and $\bar{\tau} \equiv \frac{1}{c}\left((1-\alpha)\beta\Delta V - (1-\lambda(1+\phi))\tilde{\theta}\right)$.

5.3 Discussion of further assumptions

5.3.1 Endogenous ability

So far we have implicitly assumed that a parent with ability k ($k = h,l$) receives a child with the same ability. In the following we relax this assumption and assume that a parent with ability h has a child with ability h with probability $\bar{p}$ whereas a parent with ability l receives a child with ability l with probability $\tilde{p}$. Then, the dynamics of individuals with low ability are characterized by the following first-order difference equation:

$$\eta_{t+1} = \eta_t \mu_t \bar{p} + \eta_t (1-\mu_t)\bar{p} + (1-\eta_t)\mu_t \tilde{p} + (1-\eta_t)(1-\mu_t)\tilde{p} \tag{30}$$

which has a unique steady state $\eta^* = \tilde{p}/(1-(\bar{p}-\tilde{p}))$. Hence, allowing for a type-specific transmission probability of abilities may slightly alter the dynamics of the basic model but does not affect the existence of steady states, so that our main result carries over to such a framework.[34]

5.3.2 Endogenous apprehension probability

The basic model abstracts from using an endogenous probability of apprehension. Therefore, the probability that a criminal is apprehended is independent of the total number of crimes committed in the economy. In a more realistic framework, however, one may expect that a larger number of crimes, decreases the effectiveness of enforcement (e.g. the ability of the police to handle each case effectively), so that the probability of not being apprehended increases with the overall level of crime in the economy, i.e. $\rho(\tilde{x})$ with $\partial\rho(\tilde{x})/\partial\tilde{x} > 0$. This kind of externality has been emphasized e.g. by Ferrer (2010) and others. However, adding this perhaps more realistic feature to our basic model would not affect any of the main results **Why?–>explanation?**.

5.3.3 Type specific oblique transmission

The basic model assumes that the oblique transmission function $S(\mu_t)$ is the same for all individuals and thus independent of the parents' type. From a practical standpoint, however, there may be good reasons to expect that children with a type L (H) parent are more likely to be exposed to peers with type L (H) (relative to what is represented by the average of the society). In

[34] Another interesting case would be one in which the parents' behavior or effort directly affects the transmission of abilities. We leave a more thorough investigation of this mechanism for future research.

this case, $S(\mu_t)$ would be type-specific and transmission probabilities can be written as follows:

$$P_k^{LL} = x_t^{L,k} + (1 - x_t^{L,k})S^L(\mu_t) \tag{31}$$

$$P_k^{LH} = (1 - x_t^{L,k})((1 - S^L(\mu_t)) \tag{32}$$

$$P_k^{HH} = (1 - x_t^{H,k})((1 - S^H(\mu_t)) \tag{33}$$

$$P_k^{HL} = x_t^{H,k} + (1 - x_t^{H,k})S^H(\mu_t) \tag{34}$$

where $S^L(\mu_t)$ ($S^H(\mu_t)$) represents an oblique transmission function with a positive (negative) bias. Following the same steps as in the previous analysis, it can be shown that the time spent on committing crime is $x_t^{i,k} = \frac{1}{\gamma^i}\left(\rho\lambda\tilde{\theta} - (1 - \lambda(1+\phi))\theta^k - \beta(1 - S^i(\mu_t))\Delta V\right)$ ($k = h,l$ and $i = H,L$) and that the dynamics are very similar to the ones of the basic model so that both traits can survive in the long-run. Hence, our analysis is robust to these changes.

6 Conclusions

This is the first paper to theoretically account for the intergenerational nature of criminal behavior. To do so, we have proposed a dynamic model of cultural transmission of moral norms. Individuals with heterogenous productivities allocate their time endowment to work in the market sector and to commit crime. The decision to commit crime, in turn, has a direct negative impact on the socialization process within the family (the child's probability of adopting norms of good conduct). We show the existence of high and low crime equilibria. Furthermore, we find that both traits, honesty and dishonesty, can survive even if all parents commit crime but at the same time agree that honesty is desirable.

Our model provides a novel and complementary explanation of why crime is highly concentrated in specific areas ('ghetto culture') and why crime rates tend to be persistent over time based on socialization mechanisms within the family. Furthermore, our theory can explain why criminal behavior persists even though parents agree that norms of good conduct (honesty) are desirable. In the present framework, crime deterrence policies not only alter economic incentive to commit crime but may also directly affect the socialization process and thus have long-lasting and amplifying effects ('the social multiplier') depending on initial conditions. Similar effects arise from policies aimed at shaping the evaluation of types by parents. Moreover, we have shown that a public education campaign which is used to emphasize

the importance of norms of good conduct is an effective tool to reduce crime by increasing the share of honest individuals in society and by altering the existence of steady state equilibria.

We consider our paper to be a first step towards a more systematic and formal analysis of the intergenerational nature of criminal behavior. Important issues for future research include, e.g., an investigation of the interaction between socialization processes and laws (Acemoglu and Jackson, 2014), the analysis of normative implications of crime deterrence and education policies within our framework or a systematic investigation of substitutability and complementarity between private and public socialization efforts, norms and criminal behavior. Empirical research on these interactions, and on the role of socialization mechanisms in determining intergenerational crime outcomes are also promising areas of future research.

Acknowledgements

We are grateful to Victor Rios-Rull, Wolfram F. Richter and to participants at the 2014 Annual Meeting of the EEA (Toulouse). Any remaining errors are ours.

Appendix

Proof of proposition 1:

(i) It is straight forward to see that $\triangle\mu = 0$ if $\mu = 1$. Just note that $S(1) = 1$.

(ii) Since $S(\mu) \geq \mu$ and $\Gamma(\mu) > 0$ for all $\mu \in [0,1[$, $\mu = 1$ is the only possible steady state. If we evaluate the derivative of $\triangle\mu$ with respect to μ at $\mu = 1$, we get

$$\frac{d(\triangle\mu)}{d\mu}|_{\mu=1} = S'(1)(1 - \tilde{y}/\gamma^L) - 1 \tag{35}$$

so that $\mu = 1$ is stable whenever $S'(1) < \frac{1}{1-\tilde{y}/\gamma^L}$.

(iii) As has been shown in (ii), the condition $S'(1) < (>)\frac{1}{1-\tilde{y}/\gamma^L}$ ensures that $\mu = 1$ is a locally (un)stable equilibrium. Also, we have $\triangle\mu > 0$ at $\mu = 0$. Consequently, if $\mu = 1$ is unstable, there is at least one stable interior rest point. If $\mu = 1$ is stable, however, there is either an even number or no interior steady states depending on the size of $\triangle\mu$ evaluated at $\mu = 0$, i.e. the size of $\bar{x}$.

(iv) With frequency dependent transmission, $\mu = 1$ is always a locally stable equilibrium as $S'(1) < 1$ (see also (ii)). The existence of interior rest points and their stability properties follow analogous to (iii).

References

Acemoglu, D., Jackson, M. O., 2014. Social norms and the enforcement of laws. NBER Working Paper 20369.

Akers, R. L., Jennings, W. G., 2009. Social learning theory. In J. Miller (Ed.), 21st Century criminology: A reference handbook. Thousand Oaks: SAGE Publications, Inc. 602, 323–332.

Anderson, D., 2014. In school and out of trouble? The minimum dropout age and juvenile crime. Review of Economics and Statistics 96, 318–331.

Ballester, C., Calvo-Armengol, A., Zenou, Y., 2010. Delinquent networks. Journal of the European Economic Association 8, 34–61.

Bayer, P., Hjalmarsson, R., Pozen, D., 2009. Building criminal capital behind bars: Peer effects in juvenile corrections. Quarterly Journal of Economics 124, 105–147.

Becker, G., 1968. Crime and punishment: An economic approach. Journal of Political Economy 76, 169–217.

Bethencourt, C., 2014. Crime and social expenditure: A political economic approach. Mimeo.

Bisin, A., Verdier, T., 2000. Beyond the melting pot: Cultural transmission, marriage and the evolution of ethnic and religous traits. Quarterly Journal of Economics 115, 955–988.

Bisin, A., Verdier, T., 2001. The economics of cultural transmission and the dynamics of preferences. Journal of Economic Theory 97, 298–319.

Bisin, A., Verdier, T., 2011. The economics of cultural transmission and socialization. Handbook of Social Economics 1A, ch.9.

Block, M. K., Heineke, J. M., 1975. A labor theoretic analysis of criminal choice. American Economic Review 65, 314–325.

Boyd, R., Richerson, P. J., 1985. Culture and the Evolutionary Process. University of Chicago Press, Chicago.

Brush, J., 2007. Does income inequality lead to more crime? A comparison of cross-sectional and time-series analyses of United States counties. Economics Letters 96, 264–268.

Burdett, K., Lagos, R., Wright, R., 2003. Crime, inequality, and unemployment. American Economic Review 93, 1764–1777.

Case, A. C., Katz, L., 1991. The company you keep: The effects of family and neighborhood on disadvantaged youth. NBER Working Paper No. 3705.

Cavalli-Sforza, L., Feldman, M., 1981. Cultural Transmission and Evolution. A Quantitative Approach. Princeton University Press, Princeton.

Chintrakarn, P., Herzer, D., 2012. More inequality, more crime? A panel cointegration analysis for the United States. Economics Letters 116, 389–391.

Chiu, W. H., Madden, P., 1998. Burglary and income inequality. Journal of Public Economics 69, 123–141.

Conley, J., Wang, P., 2006. Crime and ethics. Journal of Urban Economics 60, 107–123.

Damm, A. P., Dustmann, C., 2014. Does growing up in a high crime neighborhood affect youth criminal behavior? American Economic Review 104, 1806–1832.

Deming, D., 2011. Better schools, less crime? The Quarterly Journal of Economics 126, 2063–2115.

Demombynes, G., Oezler, B., 2005. Crime and local inequality in South Africa. Journal of Development Economics 76, 265–292.

Di Tella, R., Schargrodsky, E., 2004. Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. American Economic Review 94, 115–133.

Draca, M., Machin, S., Witt, R., 2011. Panic on the streets of london: police, crime and the july 2005 terror attacks. American Economic Review 101, 2157–2181.

Duncan, G., Kalil, A., Mayer, S. E., Tepper, R., Payne, M. R., 2005. The apple does not fall far from the tree. Russel Sage Foundation. Princeton: Princeton University Press.

Engelhardt, B., Rocheteau, G., Rupert, P., 2008. Crime and the labor market: a search model with optimal contracts. Journal of Public Economics 92, 1876–1891.

Eriksson, K. H., Hjalmarsson, R., Lindquist, M., Sandberg, A., 2014. The importance of family background and neighborhood effects as determinants of crime. CEPR Discussion Paper No. 9911.

Farrington, D. P., 2011. Families and crime. In J. Q. Wilson and J. Petersilia (Eds.), Crime and public policy. New York: Oxford University press., 130–157.

Ferrer, R., 2010. Breaking the law when others do: A model of law enforcement with neighborhood externalities. European Economic Review 54, 163–180.

Freeman, R. B., 1996. Why do so many young american men commit crime and what might we do about it? Journal of Economic Perspectives 10, 25–42.

Freeman, S., Grogger, J., Sonstelie, J., 1996. The spatial concentration of crime. Journal of Urban Economics 40, 216–231.

Funk, P., 2005. Governmental action, social norms, and criminal behavior. Journal of Institutional and Theoretical Economics 161, 522–535.

Glaeser, E., Sacerdote, B., Scheinkman, J., 1996. Crime and social interaction. Quarterly Journal of Economics 111, 508–548.

Glaeser, E. L., 1998. Are cities dying? Journal of Economic Perspectives 12, 139–160.

Glaeser, E. L., Sacerdote, B., 1999. Why is there more crime in cities? Journal of Political Economy 107, 225 – 258.

Glaeser, E. L., Sacerdote, B. I., Scheinkman, J. A., 2003. The social multiplier. Journal of European Economic Association 1, 345–353.

Grogger, J., Willis, M., 2000. The emergence of crack cocaine and the rise in urban crime rates. The Review of Economics and Statistics 82, 519 – 529.

Hauk, E., Sáez-Marti, M., 2002. On the cultural transmission of corruption. Journal of Economic Theory 107, 311–335.

Hjalmarsson, R., Lindquist, M. J., 2012. Like godfather, like son: Exploring the intergenerational nature of crime. Journal of Human Resources 550-582, 1–40.

Hjalmarsson, R., Lindquist, M. J., 2013. The origins of intergenerational associations in crime: Lessons from swedish adoption data. Labour Economics 20, 68–81.

Imrohoroglu, A., Merlo, A., Rupert, P., 2004. What accounts for the decline in crime? International Economic Review 45, 707–729.

Imrohoroglu, A., Merlo, A., Rupert, P., 2006. Understanding the determinants of crime. Journal of Economics and Finance 30, 270–283.

Kling, J., Ludwig, J., Katz, L., 2005. Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment. Quarterly Journal of Economics 120, 87–130.

Levitt, S. D., 2004. Understanding why crime fell in the 1990: four factors that explain the decline and six that do not. Journal of Economic Perspectives 18, 163–190.

Lochner, L., 2004. Education, work and crime: A human capital approach. International Economic Review 45, 811–843.

Lochner, L., 2011. Non-production benefits of education: Crime, health, and good citizenship, in E. Hanushek, S. Machin and L. Woessmann, eds. Handbook of the Economics of Education, 183–282.

Lochner, L., Moretti, E., 2004. The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. American Economic Review 94, 155–189.

Ludwig, J., Duncan, G. J., Hirschfeld, P., 2001. Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment. Quarterly Journal of Economics 116, 655–679.

Marceau, N., Mongrain, S., 2011. Competition in law enforcement and capital allocation. Journal of Urban Economics 69, 136–147.

McCord, J., 2001. Forging criminals in the family. In Susan O. White (Ed.), Handbook of Youth and Justice, Ch.11, 223–235.

Meghir, C., Palme, M., Schnabel, M., 2012. The effect of education policy on crime: An intergenerational perspective. NBER Working Paper No.18145.

O'Flaherty, B., Sethi, R., 2010. The racial geography of street vice. Journal of Urban Economics 67, 270–286.

Patacchini, E., Zenou, Y., 2011. Neighborhood effects and parental involvement in the intergenerational transmission of education. The Journal of Regional Science 51, 987–1013.

Reiss, A. J., Farrington, D. P., 1991. Advancing knowledge about cooffending: Results from a prospective longitudinal survey of london males. Journal of Criminal Law and Criminology 82, 360–395.

Rowe, D., Farrington, D. P., 1997. The familial transmission of criminal convictions. Criminology 35, 177 – 202.

Sáez-Marti, M., Sjögren, A., 2008. Peers and culture. Scandinavian Journal of Economics 110, 73–92.

Sáez-Marti, M., Zenou, Y., 2012. Cultural transmission and discrimination. Journal of Urban Economics 72, 137–146.

Tabellini, G., 2008. The scope of cooperation: Values and incentives. The Quarterly Journal of Economics 123, 905–950.

Thornberry, T. P., 2005. Explaining multiple patterns of offending across the life course and across generations. Annals of the American Academy of Political and Social Science 602, 156–195.

Thornberry, T. P., 2009. The apple doesn't fall far from the tree (or does it?): Intergenerational patterns of antisocial behaviour - The American Society of Criminology 2008 Sutherland Address. Criminology 47, 297–325.

Verdier, T., Zenou, Y., 2004. Radical beliefs, location and the causes of crime. International Economic Review 45, 731–760.

West, D. J., Farrington, D. P., 1977. The delinquent way of life. London: Heinemann.

Zenou, Y., 2003. The spatial aspects of crime. Journal of the European Economic Association 1, 459–467.

Tables and Figures

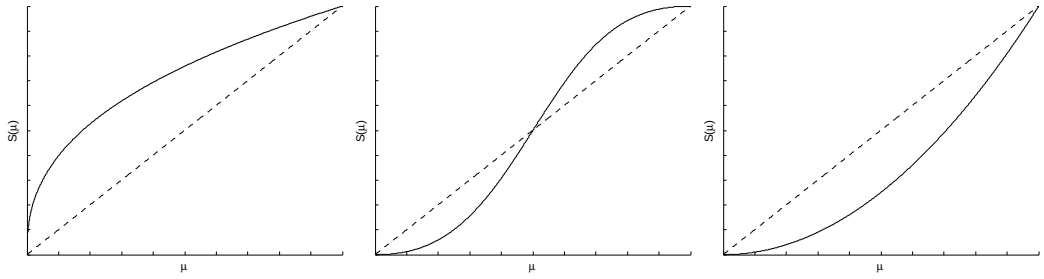

Figure 1: Positive bias (left), frequency-dependent bias (middle) and negative bias (right).

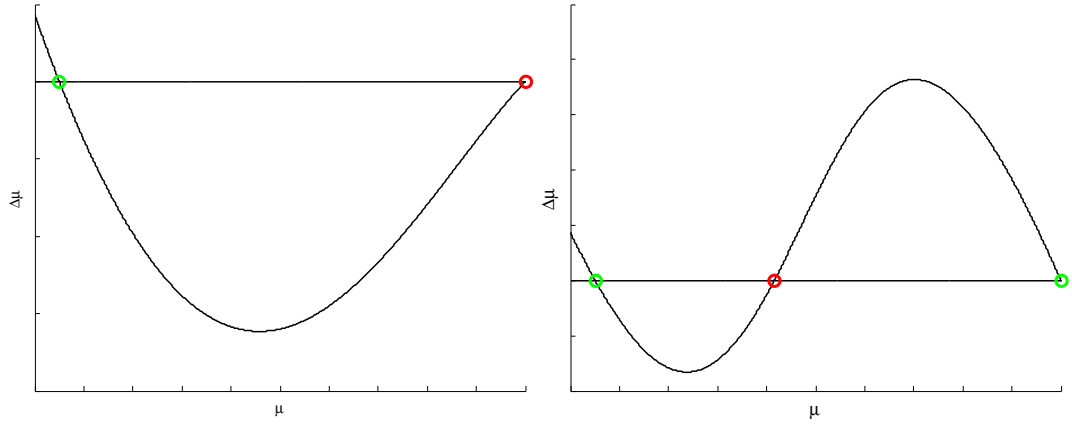

Figure 2: Negative bias with one interior steady state (left panel); Conformism with two interior steady states (right panel).

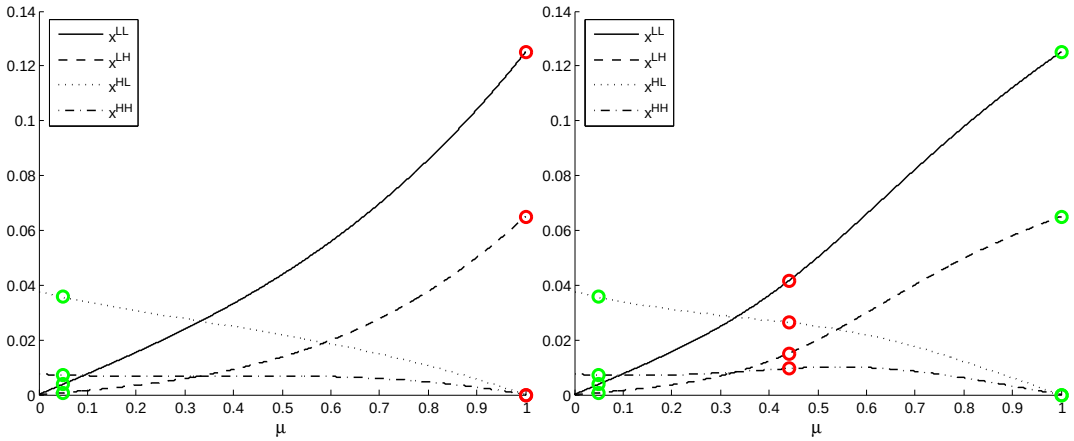

Figure 3: Population weighted crime rates with negative bias and one interior steady state (left panel, $S(\mu) = \mu^2$); Population weighted crime rates with conformism and two interior steady states (right panel, $S(\mu) = \mu^2/(\mu^2 + (1-\mu)^2)$). Parameters: $\theta^L = 1.8$, $\theta^H = 2$, $\gamma^L = 0.5$, $\gamma^H = 1$, $\rho = 0.7$, $\lambda = 0.5$, $\beta = 0.1$, $\phi = 0.4$.

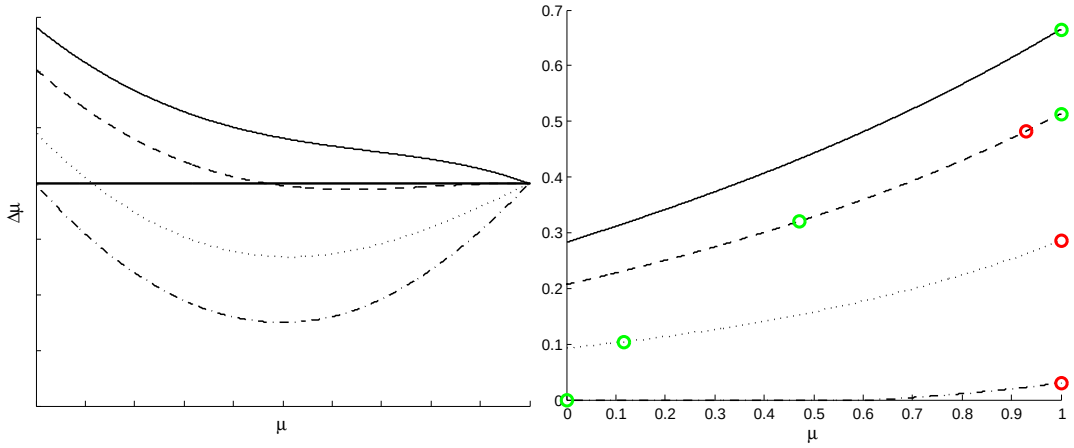

Figure 4: Changes in ρ with negative bias. Dynamics (left) and average share of criminal activity (right).

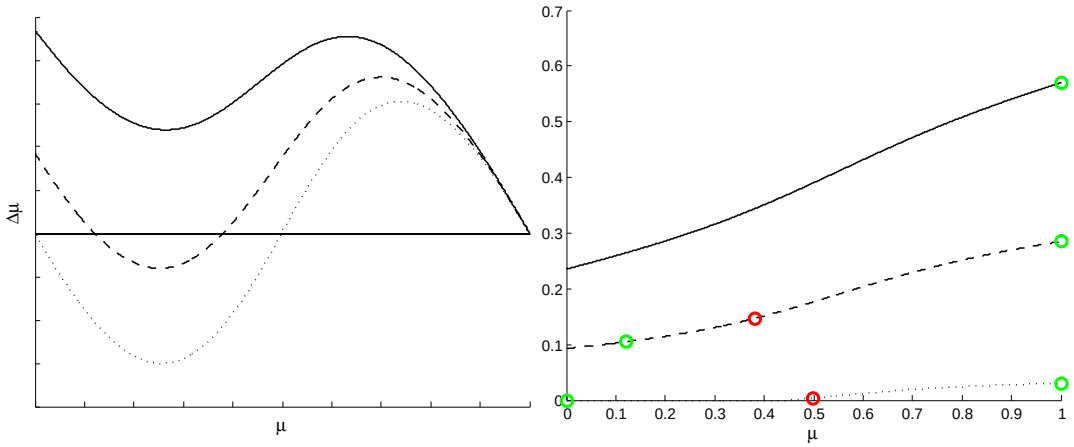

Figure 5: Changes in ρ with conformism. Dynamics (left) and average share of criminal activity (right).

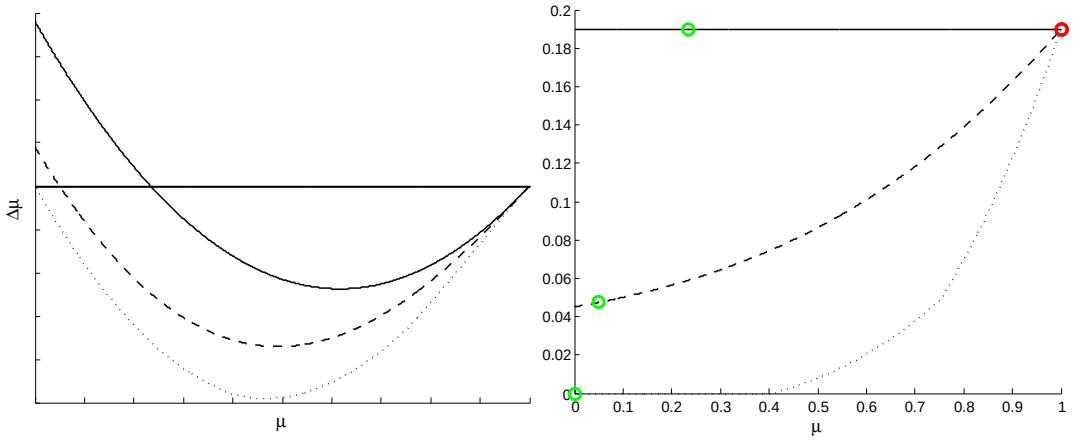

Figure 6: Changes in ΔV with negative bias. Dynamics (left) and average share of criminal activity (right).

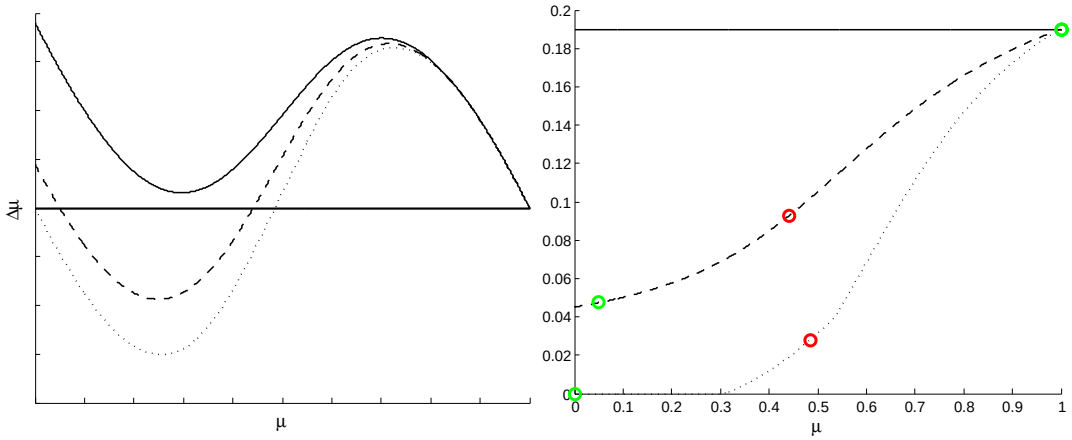

Figure 7: Changes in ΔV with conformism. Dynamics (left) and average share of criminal activity (right).

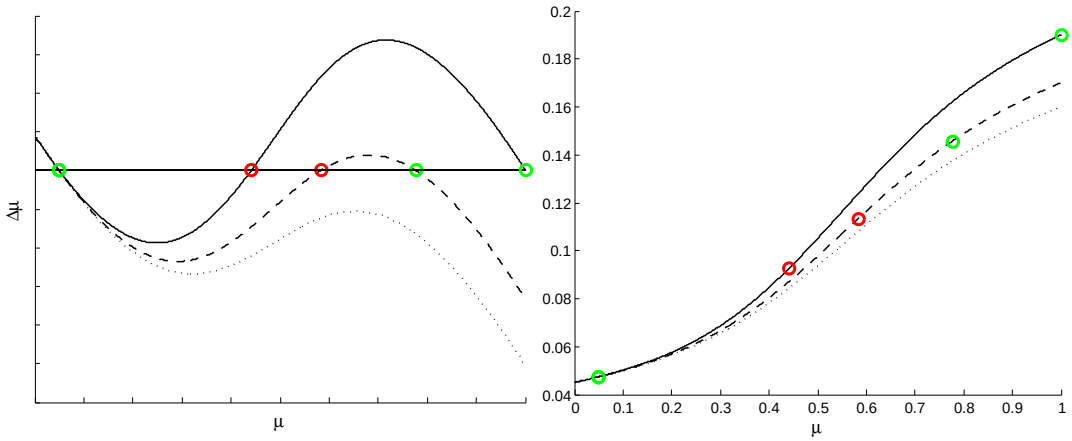

Figure 8: Changes in κ with conformism. Dynamics (left) and average share of criminal activity (right).