The economics of crime and socialization: The role of the family

Carlos Bethencourt[a,*], Lars Kunze[b]

[a] *Universidad de La Laguna, Departamento de Economia, Campus de Guajara s/n, 38071 Tenerife, Spain, phone: +34 922 317954, fax: +34 922 317204.*
[b] *TU Dortmund, Department of Economics, 44221 Dortmund, Germany, phone: +49 231 755-3275, fax: +49 231 755-5404.*

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 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, socialization, cultural transmission, moral norms

JEL-Classification K42, J15, Z13

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. A large criminological literature has documented that children with criminal parents are more likely to exhibit criminal behavior themselves.[1] The familial nature of criminal behavior has recently also attracted the interest of economists, see e.g. Hjalmarsson and Lindquist (2012, 2013), Frimmel et al. (2019) and Bhuller et al. (2021).[2] The findings of these studies emphasize the importance of family background and, in particular, parental criminality for predicting an

*Corresponding author
Email addresses: cbethenc@ull.es (Carlos Bethencourt), lars.kunze@tu-dortmund.de (Lars Kunze)
[1] See Besemer et al. (2017) for a recent survey of this literature.
[2] See section 1.1 for a more detailed discussion.

individual's criminal behavior. Those factors have been found to be even more important than income or employment status. Moreover, parents' behavior and socialization processes account for a large share of this intergenerational crime relationship, consistent with the so-called Differential Association Theory developed by Sutherland and Cressey (1966). 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. 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 (Bisin and Verdier (2011) provide a survey of the literature).

Our model proposes a novel framework to account for the familial nature of criminal behavior. More precisely, we set up an overlapping generations model in which each individual lives for two periods. During the first period of life, corresponding to childhood, individuals do not make any economic decisions but are subject to socialization and acquire their preferences (via direct socialization inside the family and indirect socialization via neighborhood effects and social interactions). As a result of the socialization process, individuals can inherit either a dishonest or an honest cultural trait. The socialization process outside the family may be biased in favor of one particular trait or depend on the relative frequency of the trait in the population (see, e.g., Bisin and Verdier (2001, sec.2.2.2) and Sáez-Marti and Sjögren (2008)).[3] The acquired trait, i.e. rules of good or bad conduct and morality, in turn affects the propensity to commit crime in the next period. During the second period of life, corresponding to adulthood, each individual becomes a parent, has one child and decides how to split the available time between joining the labor force and engaging in criminal activities. The net return from committing crime depends on economic factors and on moral norms of good conduct acquired during childhood. Parents' criminal behavior, in turn, has a key role in determining the outcome of socialization for future generations as it determines the result of the socialization process inside the family. The more children are exposed to their parents' criminal behavior, the higher is the children's probability of acquiring the dishonest trait. Parents, in turn, take into account both the negative influence of their criminal actions on the socialization process and that children are also influenced by their (potential) peers. We use this framework to analyze the dynamics of the formation of the honesty trait (the share of honest individuals) and long run crime rates.

Our analysis combines two important strands of literature: The economics of crime and cultural transmis-

[3]The main idea behind this kind of oblique transmission process is that children learn from a large group of randomly selected peers.

sion. The main contribution of our work is to provide a theoretical framework to study the intergenerational transmission of criminal behavior by integrating the role of socialization of moral norms within the family into an economic model of crime in line with Differential Association Theory. We show that our model predicts different equilibrium outcomes, depending on the children's bias: (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.

Our theory can thus explain why criminal behavior persists even though parents agree that norms of good conduct (honesty) are desirable. 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 not commit crime if it is not economically profitable). 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.[4] This case is novel. In contrast to the existing cultural transmission literature, however, we suggest that the parents' criminal behavior (rather than effort) has a direct negative impact on the children's probability of adopting the honest trait, i.e. the observation of the parents' behavior by their kin is the mechanism through which children may assimilate the cultural traits of their parents. This, in turn, allows us to establish a clear difference between the distribution of traits and actual observed behavior. Supporting evidence for such a mechanism is, e.g., provided by Christopoulou et al. (2013) who analyze the intergenerational transmission of smoking behavior: Even though all parents generally agree on the desired trait (i.e. non-smoking), smoking parents may fail to transmit the desired trait as a result of their own behavior.[5]

A further contribution of our work is to analyze public policies in order to reduce overall crime. As criminal behavior and moral values are complements in our model[6], the effects of changes in exogenous

[4]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.

[5]A related theoretical paper, which considers a similar model to ours, is Bezin et al. (2021). They also analyze a framework in which the decision to commit crime affects the intergenerational transmission of moral norms and in which parents agree that honesty is the desired trait. They show that more intense crime repression (when putting fathers in prison) may increase criminal behavior as it increases the possibility that criminals' sons become criminals themselves. See also section 1.1.

[6]Specifically, the more parents 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 and hence increases criminal activity even more.

variables and in the external environment reinforce each other. 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. In this regard, we show that a public education campaign is a powerful policy instrument to fight criminal behavior as it decreases both the parents' time spent on criminal activities and the share of honest individuals in society (via the socialization mechanism) and thus long-run crime rates. These predictions of the model are consistent with empirical evidence in Meghir et al. (2012) and Chalfin and Deza (2019) who study educational reforms in Sweden and the United States, respectively, and show that these reforms lead to a substantial reduction in crime rates of both the targeted generation and their children.

We also show that the effectiveness of an incarceration policy (i.e. a higher apprehension probability) depends on both direct and intergenerational effects. While the direct effect clearly reduces crime through the deterrence effect, consistent with many empirical studies (see Chalfin and McCrary (2017) for an overview), the effectiveness of the intergenerational effect may depend on the initial share of dishonest individuals in society.[7] If this share is sufficiently large, the intergenerational deterrence effect is not strong enough to offset the inertia of the cultural dynamics, implying the extinction of the honest type in the long-run and, correspondingly, no further decrease in the crime rate. By contrast, if the initial share of dishonest individuals is sufficiently low, both the direct and the intergenerational effect lower long-run crime. Empirical evidence on the intergenerational effect of incarceration on crime turns out to be mixed so far (see Bhuller et al. (2021)).[8] Our model provides a theoretical case for an ambiguous intergenerational effect.

The general notion that parents take into account the impact of their own criminal actions on the socialization process (and thus on the future behavior of their children) and, that they also change their behavior in anticipation of (potential) peer effects, is motivated by two stylized facts. First, it is well documented that anticipation of potential peer effects is a relevant decision variable for parents when it comes to school and neighborhood choice (see, e.g., Barseghyan et al. (2019), Agostinelli et al. (2020), Abdulkadiroglu et al. (2020)). For example, Abdulkadiroglu et al. (2020) study the relationship between parents' preferences for school choice and peer quality. Among other things, they find that parents prefer schools that enroll high-achieving peers. Second, with regard to criminal behavior, parenthood is an important factor determining

[7] As an incarceration policy has both deterrence and punishment effects, it is important to note that the focus of our paper is on the deterrence part of such a policy and not on the effects of changes in the family structure resulting from incarceration. Specifically, our model focusses on ex-ante behavioral choices parents may make rather than on any resulting effect from punishment.

[8] Wildeman and Andersen (2017), e.g., analyze changes in the probability of incarceration for certain crimes resulting from a reform in Denmark. Using a difference-in-difference strategy, they show that incarceration increases the probability of committing a crime for male children but not for females. Other studies use a so called random judge design, which exploits idiosyncratic variation in receiving incarceration. Dobbie et al. (2018), e.g., find that parental incarceration increases criminal behavior of children whereas others do either find no or negative effects (e.g., Bhuller et al. (2018) and Norris et al. (2021)).

desistance from crime (Laub et al., 1998). Empirical support comes from Dustmann and Landerso (2021) who find that young fathers decide to act more responsibly and as a role model when they father a boy. Similarly, Massenkoff and Rose (2020) show that pregnancy implies sharp and lasting declines in crime and that children play a causal role in these findings.

1.1. Related literature

Our work is related to the so-called Differential Association Theory developed by Sutherland and Cressey (1966). According to this theory, criminal behavior depends on norms which are acquired by children if such a behavior is more highly reinforced than noncriminal behavior. Moreover, children are supposed to learn behavior through observation and imitation of role models. As parents are important role models, especially during childhood, 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, even though co-offending of parents and children has been found to be unusual and parents do not encourage their children to become criminal (Reiss and Farrington, 1991; Giordano, 2010). The present paper integrates these elements of Differential Association Theory into an economic model, thereby providing a theoretical foundation of the intergenerational transmission of criminal behavior, and in particular the aforementioned social learning mechanism.

In fact, a large criminological literature has documented that children with criminal parents are more likely to become criminal themselves, see Besemer et al. (2017) for a recent overview. Moreover, there is a growing literature providing causal evidence of substantial intergenerational associations in crime (Hjalmarsson and Lindquist, 2012, 2013; Frimmel et al., 2019). 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. Similarly, Frimmel et al. (2019) provide causal evidence in favor of a strong intergenerational effect in tax evasion behavior related to the transmission of social norms. They find that that paternal non-compliance increases children's non-compliance by about 23 percent and suggest that the most likely transmission channel consists of the children's wishes to conform to the behavior of their fathers. Moreover, a recent and growing literature focusses on estimating the causal intergenerational effect of incarceration, see Bhuller et al. (2021) for an overview. The evidence from this literature, however, turns out to be mixed so far. While some studies find positive effects on child outcomes (Dobbie et al., 2018), others find no or negative effects (Bhuller et al., 2018; Norris et al., 2021).

Another strand of literature highlights the importance of peer effects in determining criminal behavior. In this literature, peers are defined in various ways. While some studies focus on neighbours (see, e.g., Glaeser et al. (1996), Kling et al. (2005) and Damm and Dustmann (2014)), others define peers as friends (see, e.g.,

Patacchini and Zenou (2012) and Lee et al. (2021), co-workers (e.g., Murphy (2019) and Hjalmarsson and Lindquist (2019)), co-offenders (e.g., Lindquist and Zenou (2014) and Bhuller et al. (2018)), people serving time together in prison (e.g., Bayer et al. (2009) and Stevenson (2017)) or family members (e.g., Hjalmarsson and Lindquist (2012); Frimmel et al. (2019)). In our work, peers are also important in determining criminal behaviour, with the main focus being on family members. However, the adoption of the honesty trait and thus the likelihood of committing crime is affected by both parental behaviour and society. To analyze the dynamics of moral norms (honest behaviour) and how they affect criminal behaviour is a new dimension we add to this literature.

Our work is also related to the literature on cultural transmission. Following the seminal papers by Bisin and Verdier (2000, 2001), 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), see Bisin and Verdier (2011) for a survey of the literature.[9] In our paper, all parents are assumed to agree that the honesty trait is superior. Whereas 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 a few 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. The most closely related paper to ours, however, is Bezin et al. (2021). They extend our idea to analyze the role of family structure (the presence of the father within a household) on children's criminal behavior and add an explicit location choice of households which implies the existence of local cultures of crime.

Finally, our work contributes to a large literature which aims at explaining the spatial variation in crime. In fact, it is well documented that crime is highly concentrated in specific areas such as city centers (see, e.g., Bezin et al. (2021) and references therein). Also, there are many cases of 'twin' cities sharing 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 of generating disparities in crime rates within specific areas range from externalities related to

[9]See, e. g., also Michaeli and Wu (2022) for a recent application of the basic modeling approach. They set up a dynamic model of inter-generational cultural transmission to explain recent trends in polarization.

the number of criminals , the spatial variation of police forces, labour market frictions, discrimination and prejudices or land prices (Freeman et al., 1996; Burdett et al., 2003; Verdier and Zenou, 2004; O'Flaherty and Sethi, 2015; Gaigné and Zenou, 2015; Galiani et al., 2018). 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, thereby providing a new mechanism to explain the spatial variation of crime.[10]

The remainder is organized as follows. Section 2 develops the basic model, establishes the existence of multiple steady states and presents comparative static results. Section 3 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 4 presents several extensions and robustness checks related to the main assumptions of our theory. Section 5 concludes.

2. 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.[11] The net return from committing crime depends on economic factors (market income, apprehension probability) 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 other crimes in which the children' observation of parents' criminal behavior would affect children's traits.[12]

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

[10] In a broader sense, our work also contributes to a large literature studying the transmission of socioeconomic outcomes from parents to children. The main focus of this literature has traditionally been on educational attainment and earnings, see e.g. Oreopoulos et al. (2006) and Chetty et al. (2014). More recently, however, further dimensions such as health behavior (Thompson, 2014), consumption (Bruze, 2018) or the existence of family welfare cultures (Hartley et al., 2022) have also been analyzed to gain a better understanding of the intergenerational link. As criminal behavior, which is the main focus of our work, directly affects available income and potentially social status, it is an important channel that affects the intergenerational transmission of socioeconomic status.

[11] Note that we normalize the individuals' time endowment to one.

[12] Examples of these kinds of crimes are: aggression to people, family violence, antisocial behaviour, etc. For instance, Cummings and Davies (2002) show that children may learn that aggression is a normal part of family relationships by observing parental fights, which in turn increases the likelihood of imitating aggressive behaviour modelled by their parents. Also, Dogan et al. (2007) find that children who grow up with antisocial parents are more likely to be antisocial themselves.

2.1. The transmission process

There are two different types of cultural traits in society, D and H. These types are referred to as dishonest and honest, respectively. Parents care about the type-dependent utility of their children. Consistent with the existing evidence (see e.g. Rowe and Farrington (1997) and Giordano (2010)), we assume that all parents, independently of their own type, agree that one of the traits (honesty) is preferable.[13] 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.

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 Differential Association Theory and 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 transmission, e.g., through social learning or adopting parental role models.[14] 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.[15]

More formally, let μ_t $(1-\mu_t)$ be the proportion of D-type (H-type) adults and $x_t^i \in [0,1]$ the fraction of time a type-i parent ($i \in \{H, D\}$) devotes to criminal activities in period t. Then, the total probability that the child of an type-i parent adopts trait $i \in \{H, D\}$ is given by:[16]

$$P^{iD} = x_t^i + (1 - x_t^i)S(\mu_t) \tag{1}$$

$$P^{iH} = 1 - P^{iD} \tag{2}$$

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

[13] Indirect evidence in favor of this assumption 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). Moreover, the assumption is supported by experimental evidence. Houser et al. (2016), e.g., show that parents cheat less when they are observed by their child.

[14] Section 4.1 extends the basic model to allow for a socialization process which depends on both effort and the parents' behavior.

[15] See section 1.1 for evidence on the importance of peer effects in determining criminal behavior.

[16] For reasons of notational simplicity, we drop the time indices for the transmission probabilities throughout the paper. Note further that, by the Law of Large Numbers, P^{ij} also denotes the fraction of children having a type-i parent and acquiring the trait j since there is a continuum of agents.

equations (1)-(2). The child of a dishonest parent ($i = D$) will also be dishonest with probability equal to the parents' time spent on criminal activities (eq.1), i.e., x_t^D. If this direct transmission fails (with probability $1 - x_t^D$), the child can still acquire the dishonest trait from his/her neighborhood through a process of oblique transmission. Given the average agent approach (Bisin and Verdier (2000, 2001) and Sáez-Marti and Sjögren (2008)), the function $S(\mu_t)$ captures the process by which the "naive" child is influenced by peers. Thus, the child acquire the dishonest trait through the oblique transmission with probability $(1 - x_t^D)S(\mu_t)$. 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 ($i = H$), 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. $\partial P^{iH}/\partial x_t^i < 0$, $i \in \{H, D\}$.[17]

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 $S(1) = 1$. A common 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. In this paper, however, we assume that oblique transmission may well be biased. This is consistent with evidence from the cultural anthropological literature, see Boyd and Richerson (1985) and the discussion in Sáez-Marti and Sjögren (2008).[18] 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.[19] Specifically, the existing literature distinguishes three different cases:

- Positive bias: The probability that the naive agent adopts type D 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 D is always smaller than if he had

[17] 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. ϵx_t^i, $i \in \{H, D\}$ and $\epsilon \in (0, 1)$ in (1)-(2). Note also that our qualitative results are robust against assuming that the transmission process is affected by the parents' gains from crime rather than the time spent on criminal activities. Furthermore, our specification implies that the marginal impact of criminal behavior on the probability of acquiring the dishonest trait decreases with $S(\mu_t)$.

[18] Applications of this modelling approach include, e.g., Bisin and Verdier (2001, sec.2.2.2), Sáez-Marti and Zenou (2012), Bezin and Moizeau (2017) and Varvarigos (2020). Moreover, the economic literature has studied conformism in various applications. For example, conformism may result if rational individuals infer the quality of goods or the virtue of a trait from the behavior and the decisions from other individuals (Bikhchandani et al., 1992) or due to the valuation of status in groups (Becker and Murphy, 2003).

[19] Note that an alternative approach would be to assume that the oblique transmission function depends on the aggregate level of crime in the economy. Such a transmission process should yield qualitatively similar results, since the aggregate level of crime, in turn, will depend on the share of dishonest individuals in society. In this paper, however, we follow the cultural transmission literature and leave a more thorough investigation of such an alternative model for future research.

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 D in the population is smaller (larger) than $\hat{\mu}_t$, the probability that a naive agent adopts D 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.

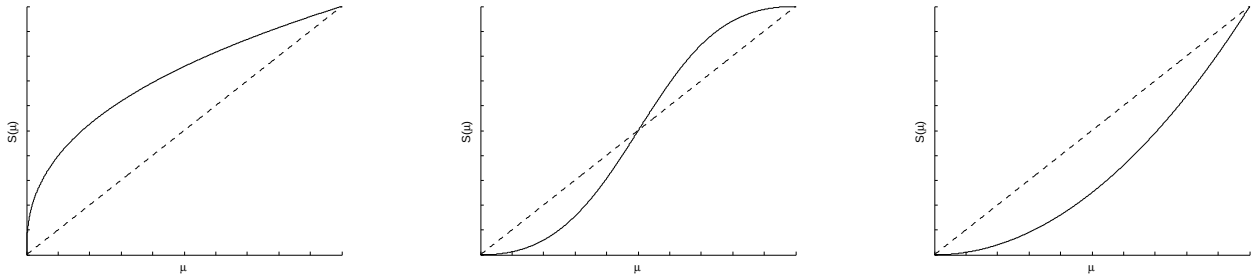

Figure 1: Positive bias (left), frequency-dependent bias (middle) and negative bias (right).

2.2. 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, $(i \in \{H, D\})$ with $\gamma^H > \gamma^D$. Furthermore, as explained in the previous section, parents are altruistic and care about their children's type-dependent utility.

Risk neutral parents maximize utility

$$U_t^i = (1 - x_t^i)w + gx_t^i - pfx_t^i - \gamma^i \frac{(x_t^i)^2}{2} + \beta \left(P^{iH} V^H + P^{iD} V^D \right) \tag{3}$$

subject to equations (1)-(2) by choosing the fraction of time devoted to criminal activities x_t^i, $(i \in \{H, D\})$. This utility function has a standard cost/benefit structure as established by the seminal contribution of Becker (1968). The benefits from the criminal activity are given by g,[20] and the costs of committing crimes are measured by the probability of being caught, p, times the fine, f, and the opportunity costs of forgone earnings in the legal sector, w. Also, as e.g. in Verdier and Zenou (2004), agents have a type-dependent moral cost of committing crime equal to $\gamma^i (x_t^i)^2/2$ $(i \in \{H, D\})$, where γ^i captures their degree of honesty. So the higher are γ^i and the crime effort, the higher is the moral cost.

[20] For technical simplicity, we assume that the apprehension probability (and therefore individuals' utility) is independent of the average level of crime in society. In the appendix, however, we show that our findings are robust against this assumption.

Finally, the last term of the utility function captures the fact that parents care about the child's well-being in adulthood.[21] More specifically, V^i denotes the parents' utility of having a child of type i ($i \in \{H, D\}$). Note that this utility does not depend on the parents' type, since we assume that all parents prefer having honest children. We also assume here that the type adopted by the child is expected to affect its well-being in adult life and that parents care about this. In our model, V^i can be interpreted as an expression of paternalistic preferences for honesty of type i, with $V^H \geq V^D$. The actual type adopted by the child will, in turn, be affected by parents' criminal behavior. More precisely, an increase in the fraction of time spent on criminal activities enhances the probability of the child adopting the D-type and thus reduces the utility from the child's expected well-being in the future.[22]

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

$$x_t^i = \max\left\{0, \frac{1}{\gamma^i}\left(g - w - pf - \beta(1 - S(\mu_t))\Delta V\right)\right\} \tag{4}$$

with $i \in \{H, D\}$. Clearly, the optimal share x_t^i depends on the individual payoff, $g - w - pf$, and the altruistic payoff, $-\beta(1 - S(\mu_t))\Delta V$. Regarding the individual payoff, we observe that the optimal share x_t^i is increasing in the gain from crime g and decreasing in the apprehension probability p and the fine f due to the standard deterrence effect.[23] Regarding the altruistic payoff, we obtain some additional and novel empirically testable predictions with respect to the variables affecting the socialization mechanism and the optimal share x_t^i: The probability with which children acquire the honest trait from their neighborhood, $1 - S(\mu_t)$, decreases the level of criminal activities. The intuition behind this finding is that a higher share of dishonest types in society, μ_t, reduces the probability of transmitting the desired trait (for a given level of x_t^i) and thus decreases the marginal cost of committing crime, implying a positive impact in the crime share. 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, $\Delta V \equiv V^H - V^D$, decreases criminal activities because the value of transmitting the desired trait increases. These behavioral reactions are broadly consistent with the empirical findings that the anticipation of potential peer effects is a relevant decision variable for parents when it comes to school and neighborhood choice (see, e.g., Barseghyan et al.

[21] This is the standard assumption in the cultural transmission literature, in which parents try to socialize their children towards a specific trait if they believe that this trait will enhance the children's welfare, see, e.g., Bisin and Verdier (2000), Bisin and Verdier (2001) and Bezin et al. (2021). Empirical support for this modeling is provided by the empirical analysis of inter vivos transfers, see e.g., Altonji et al. (1997) and Laferrere and Wolff (2006).

[22] In fact, it is straightforward to prove that $P^{iH}V^H + P^{iD}V^D$ is decreasing in x_t^i ($i \in \{H, D\}$).

[23] Note that the effects from changes in the parameters p and f are identical in our model. Furthermore, as in Bezin et al. (2021), we assume that individuals who are caught will also be incarcerated. Therefore, we use the terms 'apprehension probability' and 'incarceration rate' interchangeably in the following. In contrast to Bezin et al. (2021), however, the focus of our paper is on the deterrence part of the policy rather than on the effect of changes in the family structure resulting from incarceration.

(2019), Agostinelli et al. (2020), Abdulkadiroglu et al. (2020)) and that parenthood is an important factor determining desistance from crime (see, e.g. Dustmann and Landerso (2021) and Massenkoff and Rose (2020)).

In order to have $x_t^i < 1$ we need to assume that $\gamma^D > g - w - pf$. Depending on the parameters, there are three cases: (i) if $g < w + pf$, then crime is not economically profitable. (ii) If $g > w + pf$ and $g < w + pf + \beta\Delta V$, then both the crime decision and its extend depend on μ_t. (iii) If $g > w + pf + \beta\Delta V$, then crime is always profitable for both types but $x_t^H < x_t^D$ and, the extend of criminal behavior will depend on μ_t. Throughout the remaining analysis we focus on the third case for ease of exposition.[24]

2.3. Dynamics and steady states

Given equation (4), the dynamics of the population of agents with type D are determined by the following difference equation:

$$\begin{aligned}\mu_{t+1} =& \mu_t P^{DD} + (1-\mu_t)P^{HD} \\ =& S(\mu_t) + \left[g - w - pf - \beta\Delta V(1 - S(\mu_t))\right]\left[\frac{\mu_t}{\gamma^D} + \frac{1-\mu_t}{\gamma^H}\right](1 - S(\mu_t))\end{aligned} \tag{5}$$

The change in the fraction of D-types can be obtained from equation (5) as:

$$\triangle\mu_{t+1} = S(\mu_t) - \mu_t + \Gamma(\mu_t) \tag{6}$$

with

$$\Gamma(\mu_t) = \left[g - w - pf - \beta\Delta V(1 - S(\mu_t))\right]\left[\frac{\mu_t}{\gamma^D} + \frac{1-\mu_t}{\gamma^H}\right](1 - S(\mu_t)) \tag{7}$$

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 (individual-specific) fraction of their time to commit crime.

We denote by $\mu(t, \mu_0)$ the path resulting from equation (5) when the initial condition is μ_0, M the set of steady states[25] and $\bar{x} \equiv (g - w - pf - \beta\Delta V)/\gamma^H$ the average time honest individuals spend on criminal

[24] Note that the focus on case (iii) does not limit our analysis to particular cases. When crime is not economically profitable (case i), however, then $x_t^i = 0\ \forall i$ and the long run equilibria would be characterized by an unique trait in the population. The surviving trait will depend on the bias of the oblique transmission function, e.g. the honesty trait for positive bias or the dishonest trait for negative bias. Moreover, case (ii) is a hybrid of cases (i) and (iii): For a given set of parameters g, w, p, f, β, ΔV, we can define a $\hat{\mu}$ such that: if $\mu > \hat{\mu}$, then $x_t^i(\mu) > 0$ and the dynamics are equivalent to case (iii) and, if $\mu \leq \hat{\mu}$, then $x_t^i = 0$ and the dynamics are equivalent to case (i). The larger $\hat{\mu}$, the higher is the probability of not having interior equilibria as in case (i), whereas for $\hat{\mu}$ small enough, stable interior equilibria as in case (iii) are possible. Note further that we allow for cases (i) and (ii) in the numerical exercises in section 2.4.

[25] According to the fixed point theorem, a steady state satisfies the condition $\mu_{t+1} = \mu_t = \mu^\star$. This implies that $\triangle\mu_{t+1} = 0$

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 D), then $M = \{1\}$ *and* $\mu(t, \mu_0) \to 1 \ \forall \mu_0$.

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

- *For* $S'(1) > \frac{1}{1-(g-w-pf)/\gamma^D}$, *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* $S'(1) < \frac{1}{1-(g-w-pf)/\gamma^D}$ *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 (honest) trait is easy to adopt through oblique transmission, that is, when the oblique transmission is negatively biased (against dishonest trait) or conformist. Otherwise, even though parents agree that honesty is desirable, as long as committing crime is economically profitable and the probability that the naive agent adopts the dishonest trait is always larger than if he had acquired one role model randomly, the desired trait disappears (i.e., $\mu^* = 1$ becomes the unique stable equilibrium). 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. Hence, only with frequency-dependent or negative bias both traits can co-exist in equilibrium even if all parents commit crime. Figure 2 illustrates this. It shows the cultural dynamics for two different cases: (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 circles and unstable ones with asterisks. Clearly, the honest type can survive even in a situation in which all parents commit crime:

---

holds in steady state. There can be more than one steady state. A steady state is stable if the first derivative of the function $\triangle\mu_{t+1} = 0$ evaluated at the steady state is negative, i.e. $\frac{\partial \triangle\mu}{\partial \mu} < 0$. In this case the difference between μ_t and μ_{t+1} becomes negative when μ_t deviates from the steady state so that the dynamic system returns to the steady state. Conversely, a steady state is unstable, if the difference between μ_t and μ_{t+1} becomes larger, i.e., if $\frac{\partial \triangle\mu}{\partial \mu} > 0$ holds in steady state.

First, when the interaction with peers is negatively biased against acquiring trait D, both traits D 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 D-trait exists. However, if the starting share of trait D in the population is sufficiently large, then trait H will disappear, even though it is the desired trait for all individuals. Note 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 policies designed to combat crime or policies devoted to modify the values that children assign to different traits.

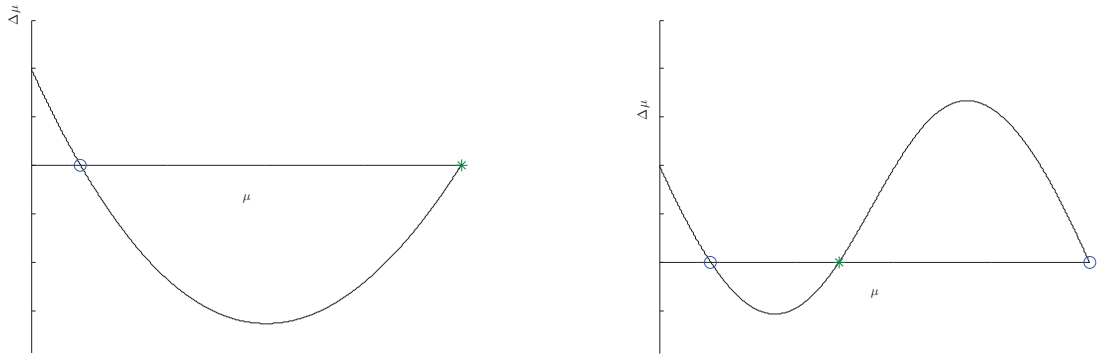

Figure 2: Negative bias with one interior steady state (left panel); Conformism with two interior steady states (right panel). Stable equilibria are marked with circles, unstable ones with asterisks.

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

$$\begin{aligned}\tilde{x}_t =& \mu_t x_t^D + (1-\mu_t) x_t^H \\ =& \left(g - w - pf - \beta \Delta V (1 - S(\mu_t))\right) \left(\frac{1-\mu_t}{\gamma^H} + \frac{\mu_t}{\gamma^D}\right)\end{aligned} \tag{8}$$

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. Figure 3 illustrates the population weighted crime rates (i.e. each of the two summands in (8)) 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 under-

standing the spatial variation in crime (see e.g. Glaeser and Sacerdote (1999) and Bezin et al. (2021)) based on socialization forces within the family. In particular, considering case (iv) of proposition 1, high crime equilibria emerge and tend to persist when, initially, the share of honest individuals is low and many adults become criminals as a result of the socialization process within the family. Together with the criminal behavior of other children, actual criminal behavior and socialization thus reinforce each other. This, in turn, explains why in some neighborhoods or cities crime rates are high and communities collapse whereas in others the opposite is true.

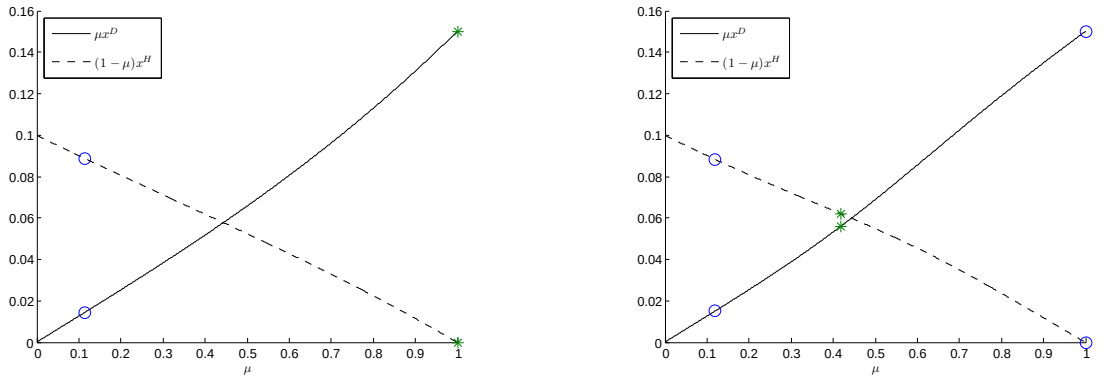

Figure 3: Population weighted crime rates with negative bias and one interior steady state (left panel, $S(\mu) = \mu^2$); Population weighted crime rates with conformism and two interior steady states (right panel, $S(\mu) = \mu^2/(\mu^2 + (1-\mu)^2)$). Parameters: $\gamma^D = 2$, $\gamma^H = 2.5$, $\Delta V = 0.5$, $p = 0.8$, $g = 1.3$, $f = 1$, $\beta = 0.1$, $w = 0.2$.

2.4. Comparative static analysis

Let us now consider some comparative static results. We are in particular interested in changes in the apprehension probability p (i.e., an increase in the incarceration rate) 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 p. Figures 4 and 5 illustrate the cultural dynamics resulting from different levels of p and the corresponding average shares of criminal activities both under negative and frequency-dependent transmission. The solid line corresponds to a high level of p, the dashed line to a medium and the dotted line to a low level. If the apprehension probability is low (dotted line), $\mu = 1$ is the only stable equilibrium and average steady state crime levels are high. Increasing the apprehension probability p, i.e. moving from the dotted to the dashed or the solid line, clearly not only has a direct positive effect on crime reduction by lowering economic incentives (the standard deterrence effect)[26] but also an intergenerational effect resulting from changes in the number and properties of steady state equilibria. This amplifying effect emerges as a

[26]See Chalfin and McCrary (2017) for a recent overview of the literature.

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. With conformism, if the initial share of honest individuals is sufficiently small, however, the intergenerational effect is not strong enough to offset the inertia of the cultural dynamics, implying the extinction of the honest type in the long-run and, correspondingly, no further decrease in the crime rate (see figure 5). These predictions provide a theoretical case for an ambiguous intergenerational effect of incarceration on crime, broadly in line with the inconclusive empirical evidence (see Bhuller et al. (2021)). Graphically, the standard effect implies the downwards shift of the curve $\tilde{x}(\mu)$, while the intergenerational effect reflects the transition to a new stable steady state determined by the shifted curve. In the limit, if the increase in p is sufficiently large, economic conditions prevent parents from committing crime, which in turn implies that $\mu = 0$ is a stable steady state.[27] 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.

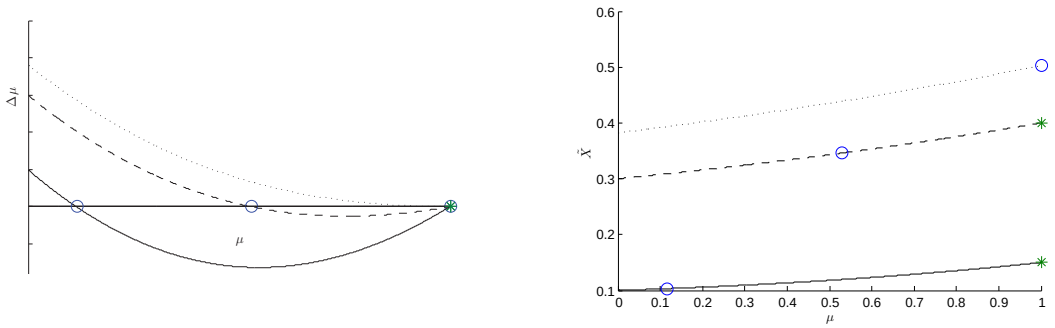

Figure 4: Changes in p with negative bias. Dynamics (left) and average share of criminal activity (right). Stable equilibria are marked with circles, unstable ones with asterisks. We consider three different values: $p = 0.8$ (solid line), $p = 0.3$ (dashed line) and $p = 0.01$ (dotted line) and use the same values as in figure 3 for the remaining parameters.

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 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. The solid line corresponds to a low value of ΔV, the dashed line to a medium value and the dotted line to a high value of ΔV. Note that in the case of a low ΔV parents do not care too much about the transmission process. In particular,

[27] Note that we also allow for corner solutions, i.e. $x_t^i = 0$ in the numerical exercises (cf. equation (4) and the related discussion).

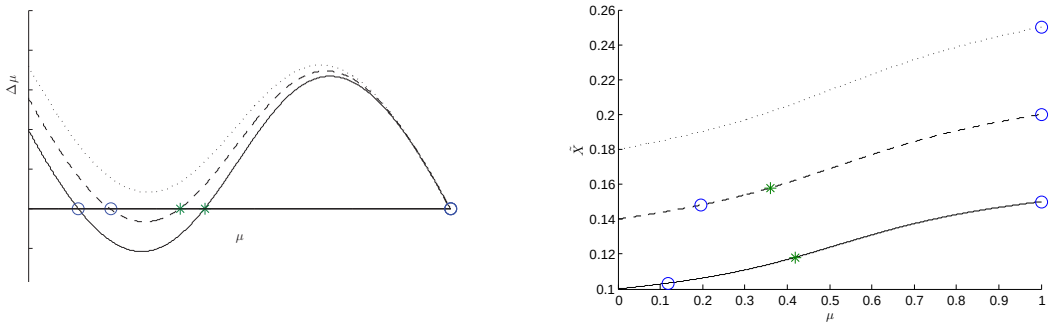

Figure 5: Changes in p with conformism. Dynamics (left) and average share of criminal activity (right). Stable equilibria are marked with circles, unstable ones with asterisks. We consider three different values: $p = 0.8$ (solid line), $p = 0.7$ (dashed line) and $p = 0.6$ (dotted line) and use the same values as in figure 3 for the remaining parameters.

for $\Delta V = 0$, the crime rate is independent of μ as the parents' decision to commit crime does not take into account the 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. For intermediate values, we observe that increases in ΔV leads to a stable interior steady state with a lower share of dishonest individuals, μ, and a lower share of criminal activities. Graphically, this corresponds to downward shifts of the curves $\triangle\mu$ and $\tilde{x}(\mu)$. Thus, a policy intervention that increases the value parents attach to the honest trait, e.g. an information campaign, would reduce the share of dishonest individuals in the population in the long run. Figure 7 illustrates the same experiment with frequency-dependent cultural transmission. Results are similar to the negative bias: higher values of ΔV lead to increases in the share of honest individuals. Consequently, parenting education programs that encourage parents to praise good and honest behavior might have a clear positive impact on reducing criminality. Farrington and Welsh (2005) document the existence of many programs, which have been successfully lowered children's antisocial behavior.[28]

We also observe that (see figure 7), the effects of increasing parents' awareness are qualitatively similar to increases in p. The important policy implication is that policies to deter crime and information/education policies shaping the evaluation of types by parents might be substitutes in fighting criminal behavior. However, as these policies operate through different channels (changes of economic incentives vs. changes of the socialization process), quantitative effects should be different and the implementation of a particular policy will depend on the related cost-benefit analysis.[29]

[28] Olds et al. (1998), for example, investigate the effects of a home visiting program on pregnant women giving them advice about child rearing and the need to avoid some bad habits. A 15-year follow-up showed that the children of visited mothers were significantly lower arrested than the children of non-visited mothers.

[29] Tentative simulation results using the same percentage changes for both parameters p and ΔV suggest that increases in p imply larger reductions in the crime rate as well as in the share of honest individuals. This is not surprising, however, as p

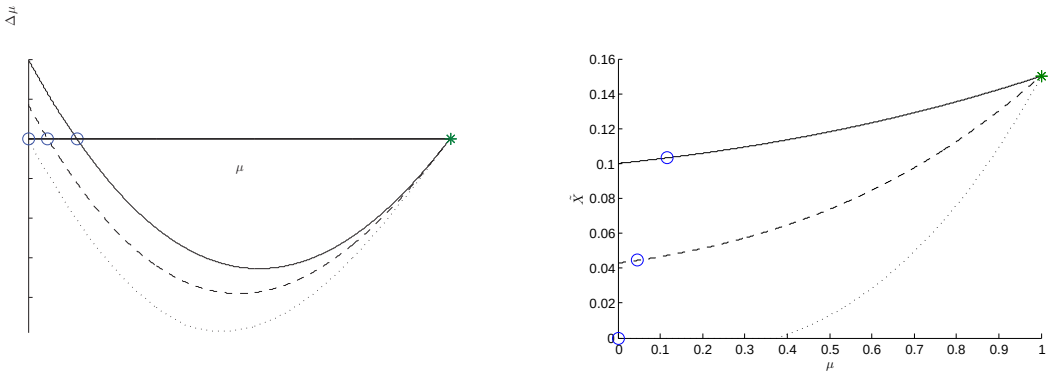

Figure 6: Changes in ΔV with negative bias. Dynamics (left) and average share of criminal activity (right). Stable equilibria are marked with circles, unstable ones with asterisks. We consider three different values: $\Delta V = 0.5$ (solid line), $\Delta V = 1.5$ (dashed line) and $\Delta V = 2.5$ (dotted line) and use the same values as in figure 3 for the remaining parameters.

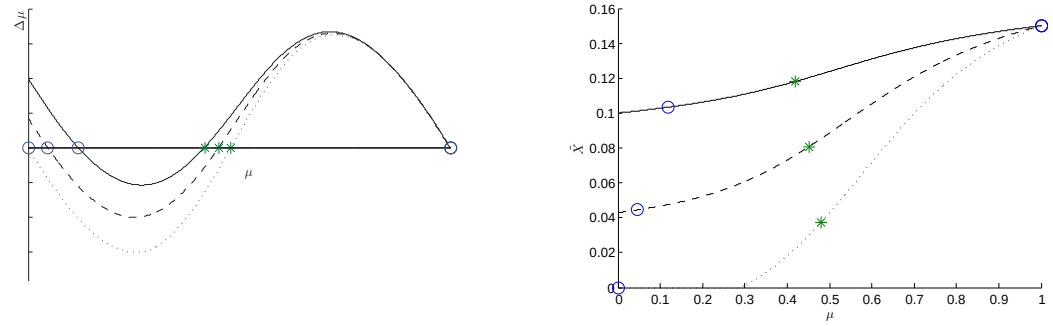

Figure 7: Changes in ΔV with conformism. Dynamics (left) and average share of criminal activity (right). Stable equilibria are marked with circles, unstable ones with asterisks. We consider three different values: $\Delta V = 0.5$ (solid line), $\Delta V = 1.5$ (dashed line) and $\Delta V = 2.5$ (dotted line) and use the same values as in figure 3 for the remaining parameters.

3. Public education campaign

This section analyzes the effectiveness of public education campaigns used to emphasize the importance of norms of good conduct.[30] 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.[31] Public education efforts affect the

affects crime both directly and through changes in the socialization process. A more rigorous simulation of the model would be an interesting topic for future research.

[30] See Lochner (2011) and Bell et al. (2022) for an overview on the relationship between crime and education.

[31] 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

probabilities of honest and dishonest children as follows:[32]

$$P^{iD} = x_t^i + (1 - x_t^i)S(\mu_t)(1 - \kappa) \tag{9}$$

$$P^{iH} = (1 - x_t^i)((1 - S(\mu_t))(1 - \kappa) + \kappa) \tag{10}$$

As before, the parents' time spent on criminal activities x_t^i, $(i \in \{H, D\})$, determines the probability of children adopting the dishonest trait (eq. 9). With the complementary probability $1 - x_t^i$ 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 (eq. 10).

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

$$x_t^i = \max\left\{0, \frac{1}{\gamma^i}\left(g - w - pf - \beta(1 - S(\mu_t)(1 - \kappa))\Delta V\right)\right\} \tag{11}$$

with $i \in \{H, D\}$. The new change in the fraction of D-types with public education is given by:

$$\triangle\mu_{t+1} = S(\mu_t)(1 - \kappa) - \mu_t + \Gamma(\mu_t) \tag{12}$$

with

$$\Gamma(\mu_t) = \left[g - w - pf - \beta\Delta V(1 - S(\mu_t)(1 - \kappa))\right]\left[\frac{\mu_t}{\gamma^D} + \frac{1 - \mu_t}{\gamma^H}\right](1 - S(\mu_t)(1 - \kappa)) \tag{13}$$

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.[33] 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:

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.

---

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.

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

[33] Note that in this case if (i) $g < w + pf + \beta\kappa\Delta V$, then crime is not economically profitable, if (ii) $g > w + pf + \beta\kappa\Delta V$ and $g < w + pf + \beta\Delta V$, then both the crime decision and its extend depend on μ_t and if (iii) $g > w + pf + \beta\Delta V$, then crime is always profitable for both types but $x_t^H < x_t^D$. In order to have $x_t^i < 1$ we assume $\gamma^D > g - w - pf$. As in the baseline model we focus on the third case for ease of exposition.

Proof: *Straightforward from proposition 1.*

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.[34] 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 children's bias 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 the population dynamics and the proportion of honest individuals increases. Specifically, successive increases in κ lead to new stable interior steady states with lower shares of dishonest agents and criminal activities (dashed and dotted lines). Graphically, we observe downward shifts of the curves $\triangle\mu$ and $\tilde{x}(\mu)$ and the disappearance of the high crime steady state if κ is sufficiently large (dotted line).

While the direct effect of education on crime is empirically well documented, see e.g. Bell et al. (2022) and references therein, the importance of intergenerational effects of education on crime has only recently received some attention. Meghir et al. (2012) exploit a major educational reform in Sweden and 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. Similarly, Chalfin and Deza (2019) provide evidence of an intergenerational effect of education on crime by analyzing changes in compulsory schooling laws in the United States. They conclude that previous analyses of compulsory schooling laws - and investments in education more generally - appreciably underestimate the full benefits of investments in education. In a related paper, Chalfin and Deza (2018) also document the existence of a positive intergenerational effect of education on alcohol abuse in the United States. The predictions of our model are consistent with the aforementioned evidence.

[34] As in the baseline model, all individuals devote some type-specific amount of time to criminal activities. However, the model could be modified to consider a case in which some agents never commit crime, for instance, by introducing heterogeneity in wages. We have analyzed this case in a previous version of the paper and both qualitative and quantitative results turn out to be very similar.

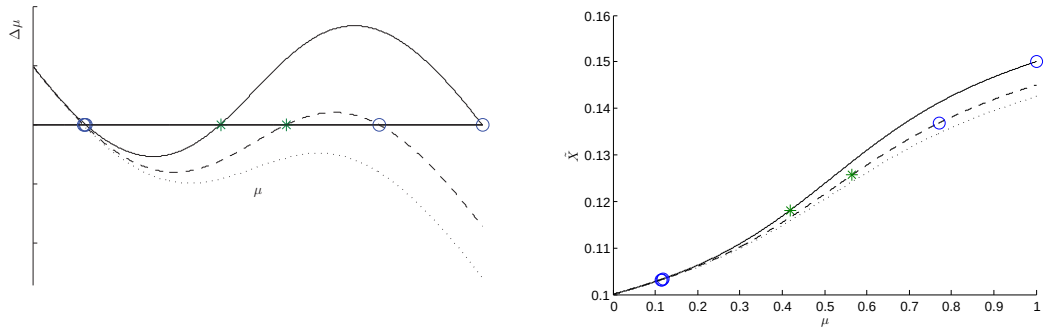

Figure 8: Changes in κ with conformism. Dynamics (left) and average share of criminal activity (right). Stable equilibria are marked with circles, unstable ones with asterisks. We consider three different values: $\kappa = 0$ or no education campaign (solid line), $\kappa = 0.2$ (dashed line) and $\kappa = 0.3$ (dotted line) and use the same values as in figure 3 for the remaining parameters.

4. Robustness

The transmission process in the basic model, i.e. equations (1) and (2), adopts the standard functional forms assumed in the literature of cultural transmission proposed by Bisin and Verdier (2000) and Bisin and Verdier (2001)[35] and, more specifically, in the literature that models the cultural transmission of one desired trait for all types of parents, see e.g. Sáez-Marti and Sjögren (2008), Sáez-Marti and Zenou (2012) and Bezin et al. (2021). In line with this latter strand of literature (see also Boyd and Richerson (1985)), we assume that the oblique transmission process may be biased. The corresponding functional form assumptions regarding the transmission functions are very common in the economic literature, see e.g. Sáez-Marti and Zenou (2012), Bezin and Moizeau (2017) and Varvarigos (2020).[36]

The aim of this section is to consider several sensitivity checks in order to demonstrate that our findings are robust against assumptions and choices made. More precisely, we introduce an explicit socialization effort into the basic model in order to reconcile our approach with the standard assumption in the cultural transmission literature, i.e. that socialization requires parental effort (section 4.1). We also consider two alternative transmission processes. The first one assumes that the honest trait can only be transmitted if it is acquired through oblique transmission by peers, i.e., if vertical and oblique socialization are complementary (section 4.2). Such a transmission process is studied by Bezin et al. (2021). The second one assumes that the oblique transmission is type-specific (section 4.3), e.g. due to parents' school or neighborhood choice.

[35] See, for example, also Hauk and Sáez-Marti (2002) and Bisin et al. (2011) for further motivation of these functional forms.
[36] See also Bikhchandani et al. (1992) or Becker and Murphy (2003) for alternative applications of conformist behavior.

4.1. Socialization effort vs. criminal behavior

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:

$$P^{iD} = \alpha x_t^i + [\alpha(1 - x_t^i) + (1 - \alpha)(1 - \tau_t^i)]S(\mu_t) \tag{14}$$

$$P^{iH} = (1 - \alpha)\tau_t^i + [\alpha(1 - x_t^i) + (1 - \alpha)(1 - \tau_t^i)](1 - S(\mu_t)) \tag{15}$$

where $i \in \{D, H\}$ and $\alpha \in (0, 1)$ measures the relative importance of the parents' actual criminal behavior and their own socialization effort in determining the horizontal transmission probabilities. The case $\alpha = 1$ corresponds to the basic model of section 2. 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, as before) and to socialize their children (τ_t^i, with $i \in \{H, D\}$). We assume that educating ones child has a cost beyond the foregone income from working in the legal market given by $c(\tau_t^i) = c(\tau_t^i)^2/2$. Hence, parents' utility function is given by

$$U_t^i = (1 - x_t^i - \tau_t^i)w + gx_t^i - pfx_t^i - \gamma^i \frac{(x_t^i)^2}{2} + \beta \left(P^{iH} V^H + P^{iD} V^D \right) - c\frac{(\tau_t^i)^2}{2} \tag{16}$$

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

$$x_t^i = \max \left\{ 0, \frac{1}{\gamma^i} \left(g - w - pf - \alpha\beta(1 - S(\mu_t))\Delta V \right) \right\} \tag{17}$$

$$\tau_t^i = \max \left\{ 0, \frac{1}{c} \left((1 - \alpha)\beta S(\mu_t)\Delta V - w \right) \right\} \tag{18}$$

with $i \in \{H, D\}$.[37] Parents' incentive to invest in education increases with the number of dishonest individuals in society, μ_t, the relative evaluation of having an honest child, ΔV, and the parents level of altruism, β.[38] By contrast, it is decreasing in forgone earnings in the legal sector (the opportunity costs of educating children), w. Finally, if the share of dishonest individuals is sufficiently small, i.e. $\mu_t < S^{-1}(w/((1 - \alpha)\beta\Delta V))$, parents will free-ride on trait transmission by society and not exert any effort to educate their child.

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

[38] Note that the comparative static analysis for the share of time spent on criminal activities is basically the same as in section 2.

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

$$\begin{aligned}\triangle\mu_{t+1} = \alpha S(\mu_t) - \mu_t + \alpha(1 - S(\mu_t)) \left(\frac{\mu_t}{\gamma^D} + \frac{1-\mu_t}{\gamma^H} \right) \left(g - w - pf - \alpha\beta(1 - S(\mu_t))\Delta V \right) \\ + (1-\alpha)S(\mu_t)\frac{1}{c}\left((1-\alpha)\beta S(\mu_t)\Delta V - w \right) \end{aligned} \tag{19}$$

with $\triangle\mu_{t+1} = \alpha\bar{x}^{\alpha} > 0$ if $\mu_t = 0$ and $\triangle\mu_{t+1} = (1-\alpha)(\bar{\tau} - 1) < 0$ if $\mu_t = 1$.[39] 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.*

Proof: *Straightforward from proposition 1.*

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 thus turn out to be robust against the introduction of effort into the basic model.

4.2. Complementarities between vertical and oblique socialization

In this section we assume that vertical and oblique socialization are complementary. This means that a child only becomes honest if parents succeed in transmitting the honesty trait and the role model which is randomly met by the child is also honest. Similarly, children acquire the dishonest trait if both parents fail to transmit the honesty trait and the peer group has a negative influence. A problem emerges when there is a conflict between socialization by parents and by society. In this case we assume that children are randomly matched with a role model from society a second time and so, children acquire the trait of this role model. Transmission probabilities can then be written as follows ($i \in \{H, D\}$):

$$P^{iD} = x_t^i S(\mu_t) + x_t^i(1 - S(\mu_t))S(\mu_t) + (1 - x_t^i)S(\mu_t)S(\mu_t) \tag{20}$$

$$P^{iH} = (1 - x_t^i)(1 - S(\mu_t)) + (1 - x_t^i)S(\mu_t)(1 - S(\mu_t))S(\mu_t) + x_t^i(1 - S(\mu_t))(1 - S(\mu_t)) \tag{21}$$

This socialization process implies that a child of dishonest parents ($i = D$) will also be dishonest (eq. 20) if parents transmit their trait through their criminal activity (with probability x_t^D) and if the dishonest

[39] Note that $\bar{x}^{\alpha} \equiv \frac{1}{\gamma^H}\left(g - w - pf - \alpha\beta\Delta V\right)$ and $\bar{\tau} \equiv \frac{1}{c}\left((1-\alpha)\beta\Delta V - w\right)$.

trait is also transmitted by peers (with probability $S(\mu_t)$). However, if children meet honest peers (with probability $1-S(\mu_t)$) conflicting with their parents activities, it is assumed that children have an opportunity of acquiring the dishonest trait when they randomly meet dishonest peers for a second time (with probability $S(\mu_t)$). Finally, if parents do not succeed in transmitting their trait (with probability $1-x_t^D$) but children meet dishonest peers (with probability $S(\mu_t)$), then children acquire the dishonest trait if, in a second random match, they meet dishonest peers again (with probability $S(\mu_t)$). Otherwise, children of dishonest parents with get the honest trait when both parents and peers transmit the honest trait (with probability $(1-x_t^D)(1-S(\mu_t))$). Moreover, if transmission by parents and peers is conflicting, children may also acquire the honest trait if in a second round children randomly meet honest peers (eq. 20). The socialization process of children with honest parents is analogue.

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

$$x_t^i = \max\left\{0, \frac{1}{\gamma^i}\left(g - w - pf - 2\beta(1-S(\mu_t))S(\mu_t)\Delta V\right)\right\} \tag{22}$$

with $i \in \{H, D\}$. The new change in the fraction of D-types is given by:

$$\triangle\mu_{t+1} = S^2(\mu_t) - \mu_t + \Gamma(\mu_t) \tag{23}$$

with

$$\Gamma(\mu_t) = 2\left[g - w - pf - 2\beta\Delta V(1 - S(\mu_t)S(\mu_t))\right]\left[\frac{\mu_t}{\gamma^D} + \frac{1-\mu_t}{\gamma^H}\right](1-S(\mu_t))S(\mu_t) \tag{24}$$

From these equations it is straightforward to see that $\mu = 0$ and $\mu = 1$ are always possible steady states. Moreover, further analysis of equation (23) gives rise to the following proposition:

Proposition 4. *Suppose that* $S'(1) > \frac{1-x^D(1)}{x^H(1)}$ *and* $S'(0) > \frac{x^H(0)}{1-x^D(0)}$. *Then, there exists at least one interior equilibrium* μ^* *such that* $\mu(t, \mu_0) \to \mu^*$ *for all* μ_0.

Proof: *See appendix.*

Similar to the previous extensions, we show that both honest and dishonest types can survive in the long-run and that multiple equilibria with either high or low crime rates may exist. Hence, the main predictions of our model turn out to be robust against alternative modeling approaches to the socialization process. Moreover, as in the previous case, the existence of an interior steady state does not depend on the children's bias.

4.3. 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 D (H) parent are more likely to be exposed to peers with type D (H) (relative to what is represented by the average of the society), which may be the result of parents' school or neighborhood choice. In this case, $S(\mu_t)$ would be type-specific. Thus, the total probability that the child of an type-i parent adopts trait $i \in \{H, D\}$ is given by:

$$P^{iD} = x_t^i + (1 - x_t^i)S^i(\mu_t) \tag{25}$$

$$P^{iH} = 1 - P^{iD} \tag{26}$$

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 $x_t^i = \max\left\{0, \frac{1}{\gamma^i}\left(g - w - pf - \beta(1 - S^i(\mu_t))\Delta V\right)\right\}$ ($i \in \{H, D\}$) and that the dynamics are qualitatively very similar to the ones of the basic model so that both traits can survive in the long-run.

5. 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 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) consistent with Differential Association Theory. 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 thus provides a novel and complementary explanation of why crime rates tend to be persistent over time and why it is highly concentrated in specific areas 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 (if children are negatively biased in favor of it, or if they are conformist and their environment is dominated by the dishonest trait).

In the present framework, policies aimed at deterring crime not only alter economic incentive to commit crime but may also directly affect the socialization process and thus have amplifying intergenerational effects 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.

With regard to policy implications, our theory justifies policy intervention in defense of the honesty trait if such a trait is valued in society. Moreover, policies aimed at fighting criminal behavior and at promoting children's adoption of the honesty trait will be aligned with parental preferences. With respect to the spatial variation in crime, polices that foster the movement of individuals along neighborhoods or cities (reallocation policy or subsidy policy) might be used to reduce aggregate crime. These policies should be aimed at increasing the possibility of having equilibria with a high share of honest individuals (see also Bezin et al. (2021)).

Finally, our model predicts that parents should reduce criminal behavior when societal (potential peer) criminal behavior decreases. While there is broad evidence supporting the general notion that parents take into account the impact of their own criminal actions on the socialization process (and thus on the future behavior of their children) and, that they also change their behavior in anticipation of (potential) peer effects[40], the specific predictions of our model could explicitly be tested by future research. For example, experiments or empirical analysis could address the effect of a schooling reform on parents of children in school to see whether these parents reduce their criminal behavior. Such an analysis would be related to empirical papers using non-school age people as a control group to measure the causal effect of policies, see e.g. Duflo (2001) and Duflo et al. (2021) who evaluate the causal impact of schooling on various outcome measures such as educational attainment, wages or health behavior.

We consider our paper to be a first step towards a more systematic and formal analysis of the intergenerational nature of criminal behavior. While the mechanism considered in this paper helps to improve the understanding of the origins of criminal behavior, the limitations of our analysis, at the same time, provide guidance for future research. Our model could be extended, for example, to include specific features of criminal behavior that differ from culture, such as learning of criminal skills or crime spillovers. Adding these features would generate additional channels for crime to be geographically concentrated. Other important issues for future research include, e.g., an investigation of the interaction between socialization processes and laws (Acemoglu and Jackson, 2017), 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.

[40] Fore example, it is well documented that anticipation of potential peer effects is a relevant decision variable for parents when it comes to school and neighborhood choice (see, e.g., Barseghyan et al. (2019), Agostinelli et al. (2020), Abdulkadiroglu et al. (2020)). Moreover, parenthood is an important factor determining desistance from crime (see, e.g., Dustmann and Landerso (2021) and Massenkoff and Rose (2020)).

Acknowledgements

We are grateful to José-Víctor Ríos-Rull, Yves Zenou, Wolfram F. Richter and to participants at the Annual Meeting of the EEA (Toulouse) and the Annual Conference of the Royal Economic Society (Manchester) for helpful comments and suggestions. Any remaining errors are ours. Finally, Carlos Bethencourt thanks the Spanish Ministry of Science and Technology (R&D Excellence Program) for Grant PID2019-107161GB-C33 for financial support.

Appendix

Proof of proposition 1:

First, note $\triangle\mu = \bar{x} > 0$ if $\mu = 0$ where $\bar{x} \equiv (g - w - pf - \beta\Delta V)/\gamma^H$.

(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

$$\frac{d(\triangle\mu)}{d\mu}|_{\mu=1} = S'(1)(1 - (g - w - pf)/\gamma^D) - 1 \tag{A.1}$$

so that $\mu = 1$ is stable whenever $S'(1) < 1 < \frac{1}{1-(g-w-pf)/\gamma^D}$.

(iii) As has been shown in (ii), the condition $S'(1) < (>)\frac{1}{1-(g-w-pf)/\gamma^D}$ 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 then follow analogous to (iii).

Proof of proposition 4:

If we evaluate the derivative of $\triangle\mu$ with respect to μ at $\mu = 1$ and $\mu = 0$, we get

$$\frac{d(\triangle\mu)}{d\mu}|_{\mu=1} = S'(1)(1 - x^D(1)) - x^H(1) \tag{A.2}$$

$$\frac{d(\triangle\mu)}{d\mu}|_{\mu=0} = S'(0)x^H(0) + x^D(0) - 1 \tag{A.3}$$

so that $\mu = 1$ is unstable whenever $S'(1) > \frac{1-x^D(1)}{x^H(1)}$ and, similarly, $\mu = 0$ is unstable if $S'(0) > \frac{x^H(0)}{1-x^D(0)}$.

Proof of proposition 5:

If we evaluate the derivative of $\triangle\mu$ with respect to μ at $\mu = 1$ and $\mu = 0$, we get

$$\frac{d(\triangle\mu)}{d\mu}|_{\mu=1} = 2S'(1) - 2S'(1)\frac{(g-w-pf)}{\gamma^D} \tag{A.4}$$

$$\frac{d(\triangle\mu)}{d\mu}|_{\mu=0} = 2S'(0)\frac{(g-w-pf)}{\gamma^H} - 1 \tag{A.5}$$

so that $\mu = 1$ is unstable whenever $S'(1) > \frac{1}{2\left(1-\frac{(g-w-pf)}{\gamma^D}\right)}$ and, similarly, $\mu = 0$ is unstable if $S'(0) > \frac{\gamma^H}{2(g-w-pf)}$.

Endogenous apprehension probability:

The basic model abstracts from 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 elaborated framework, however, one may expect that a larger number of crimes reduces the effectiveness of enforcement (e.g. the ability of the police to handle each case effectively), so that the probability of being apprehended decreases with the overall level of crime in the economy, i.e. $p(\tilde{x})$ with $\partial p(\tilde{x})/\partial\tilde{x} < 0$. 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 as it only enlarges the size of the intergenerational effect or, put differently, the length of the transition process to a stable equilibrium. For example, suppose that the share of individuals with the undesired trait is increasing throughout the transition towards a new steady state. Then, the resulting increase in the aggregate crime rate reduces the apprehension probability, which in turn increases the incentive to commit crimes and thus, the share of dishonest individuals in society and so on. Hence, the speed of convergence to the new equilibrium will be lower.

Formally, consider the basic model framework with the transmission process defined by equations (1)-(2) and the parents maximization problem defined in section 2.2. Imagine now that the probability of being apprehended is given by $p = p(\tilde{x}) = p(1-\tilde{x})$, $p \in (0,1)$. Note that the parents' optimal fraction of time spent on criminal activities is then:

$$x_t^i = \max\left\{0, \frac{1}{\gamma^i}\left(g - w - p(1-\tilde{x}_t)f - \beta(1-S(\mu_t))\Delta V\right)\right\} \tag{A.6}$$

where the aggregate share of criminal activities in period t is given by

$$\tilde{x}_t = \big(g - w - p(1-\tilde{x}_t)f - \beta\Delta V(1-S(\mu_t))\big)\left(\frac{1-\mu_t}{\gamma^H} + \frac{\mu_t}{\gamma^D}\right) \tag{A.7}$$

Clearly, a higher aggregate crime rate encourages individuals to devote more time to commit crime as it reduces the expected costs. Simplifying equation (A.7), it is easy to see that,

$$\tilde{x}_t = \big(g - w - pf - \beta\Delta V(1-S(\mu_t))\big)\frac{\left(\frac{1-\mu_t}{\gamma^H} + \frac{\mu_t}{\gamma^D}\right)}{1 - pf\left(\frac{1-\mu_t}{\gamma^H} + \frac{\mu_t}{\gamma^D}\right)}. \tag{A.8}$$

Hence, in this version of the model the aggregate share of criminal activities tends to be lower than in the basic model (cf. equation 8). A higher aggregate crime rate will also affect the dynamics of types in the population. Now, the change in the fraction of D-types is given by:

$$\mu_{t+1} = S(\mu_t) + \big[g - w - pf - \beta\Delta V(1-S(\mu_t)) + pf\tilde{x}_t\big]\left[\frac{\mu_t}{\gamma^D} + \frac{1-\mu_t}{\gamma^H}\right](1-S(\mu_t)) \tag{A.9}$$

and the change in the fraction of D-types can be obtained from (A.9) and (A.8) as:

$$\triangle\mu_{t+1} = S(\mu_t) - \mu_t + \tilde{\Gamma}(\mu_t) \tag{A.10}$$

with

$$\tilde{\Gamma}(\mu_t) = \big[g - w - pf - \beta\Delta V(1-S(\mu_t))\big]\frac{\left[\frac{\mu_t}{\gamma^D} + \frac{1-\mu_t}{\gamma^H}\right]}{1 - pf\left(\frac{1-\mu_t}{\gamma^H} + \frac{\mu_t}{\gamma^D}\right)}(1-S(\mu_t)) \tag{A.11}$$

It is straight forward to see that $\tilde{\Gamma}(\mu_t) \geq 0$ for all $\mu_t \in [0,1]$. In particular, $\triangle\mu_{t+1} = \bar{x}^p > 0$ if $\mu_t = 0$ and $\triangle\mu_{t+1} = 1$ if $\mu_t = 1$.[41] Moreover, note that these dynamics are very similar to the one in the basic model described by equations (6) and (7). So, regarding the existence of steady states, it follows immediately from these observations that:

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

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

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

[41] $\bar{x}^p \equiv \frac{1}{(\gamma^H - pf)}(g - w - pf - \beta\Delta V)$.

- *For* $S'(1) > \frac{1}{1-(g-w-pf)/(\gamma^D-pf)}$, *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* $S'(1) < \frac{1}{1-(g-w-pf)/(\gamma^D-pf)}$ *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* $\mu(t,\mu_0) \to 1 \ \forall \mu_0$.

The formal proof is equivalent to proposition 1.

Therefore, similar to the basic model, multiple equilibria with low or high crime rates may exit in the long-run and both types, honest and dishonest, can survive. We can conclude that all results obtained in the basic model hold when we consider an endogenous apprehension probability. The unique difference is the size of both the crime rate and the population of dishonest individuals. The negative effect of the aggregate crime rate on the apprehension probability increases the incentive to commit crime, which results in a larger intergenerational effect through an increase in the share of dishonest individuals.

Compliance with ethical standards:

Funding: Grant PID2019-107161GB-C33 from Spanish Ministry of Science and Technology (R&D Excellence Program), Spain.

Conflicts of interest/Competing interests: No conflict of interest

Availability of data and material: Not applicable

Code availability: Not applicable

References

Abdulkadiroglu, A., Pathak, P.A., Schellenberg, J., Walters, C.R., 2020. Do parents value school effectiveness? American Economic Review 110, 1502–1539.

Acemoglu, D., Jackson, M.O., 2017. Social norms and the enforcement of laws. Journal of the European Economic Association 15, 245–295.

Agostinelli, F., Doepke, M., Sorrenti, G., Zilibotti, F., 2020. It takes a village: The economics of parenting with neighborhood and peer effects. Cowles Foundation Discussion Paper No. 2228 .

Altonji, J.G., Hayashi, F., Kotlikoff, L.J., 1997. Parental altruism and inter vivos transfers: Theory and evidence. Journal of Political Economy 105, 1121–1166.

Barseghyan, L., Clark, D., Coate, S., 2019. Peer preferences, school competition and the effects of public school choice. American Economic Journal: Economic Policy 11, 124–158.

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.

Becker, G.S., Murphy, K.M., 2003. Social economics: Market behavior in a social environment. Belknap Press of Harvard University Press .

Bell, B., Costa, R., Machin, S., 2022. Why does education reduce crime? Journal of Political Economy 130, 732–765.

Besemer, S., Ahmad, S.I., Hinshaw, S.P., Farrington, D.P., 2017. A systematic review and meta-analysis of the intergenerational transmission of criminal behavior. Aggression and Violent Behavior 37, 161–178.

Bezin, E., Moizeau, F., 2017. Cultural dynamics, social mobility and urban segregation. Journal of Urban Economics 99, 173–187.

Bezin, E., Verdier, T., Zenou, Y., 2021. Crime, broken families, and punishment. American Economic Journal: Microeconomics , forthcoming.

Bhuller, M., Dahl, G.B., Loken, K.V., Mogstad, M., 2018. Intergenerational effects of incarceration. AEA Papers and Proceedings 108, 234–240.

Bhuller, M., Dahl, G.B., Loken, K.V., Mogstad, M., 2021. Measuring the intergenerational effects of incarceration. in Incarceration and Generation, Volume 2, ed. by Silvia Gomes, Maria de Carvalho and Vera Duarte, Springer Nature forthcoming.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change in informational cascades. Journal of Political Economy 100, 992–1026.

Bisin, A., Patacchini, E., Verdier, T., Zenou, Y., 2011. Formation and persistence of oppositional identities. European Economic Review 55, 1046–1071.

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.

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

Bruze, G., 2018. Intergenerational mobility: New evidence from consumption data. Journal of Applied Econometrics 33, 580–593.

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

Chalfin, A., Deza, M., 2018. The effect of parental education on children's drug and alcohol use. AEA Papers and Proceedings 108, 373–378.

Chalfin, A., Deza, M., 2019. The intergenerational effects of education on delinquency. Journal of Economic Behavior & Organization 159, 553–571.

Chalfin, A., McCrary, J., 2017. Criminal deterrence: A review of the literature. Journal of Economic Literature 55, 5–48.

Chetty, R., Hendren, N., Kline, P., Saez, E., 2014. Where is the land of opportunity? The geography of intergenerational mobility in the United States. Quarterly Journal of Economics 129, 1553–1623.

Christopoulou, R., Jaber, A., Lillard, D.R., 2013. The inter-generational and social transmission of cultural traits: Theory and evidence from smoking behavior. NBER Working Paper Series No. 19304 .

Cummings, E., Davies, P.T., 2002. Effects of marital conflict on children: Recent advances and emerging themes in process-oriented research. Journal of Child Psychology and Psychiatry 43, 31–63.

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

Dobbie, W., Grönqvist, H., Niknami, S., Palme, M., Priks, M., 2018. The intergenerational effects of parental incarceration. NBER Working Paper No. 24186 .

Dogan, S., Conger, R., Kim, K., Masyn, K.E., 2007. Cognitive and parenting pathways in the transmission of antisocial behavior from parents to adolescents. Child Development 78, 335–349.

Duflo, E., 2001. Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. American Economic Review 91, 795–813.

Duflo, E., Dupas, P., Kremer, M., 2021. The impact of free secondary education: Experimental evidence from Ghana. NBER Working Paper No. 28937 .

Dustmann, C., Landerso, R., 2021. Child's gender, young fathers' crime, and spillover effects in criminal behavior. Journal of Political Economy 129, 3261–3301.

Farrington, D.P., Welsh, B.C., 2005. Randomized experiments in criminology: What have we learned in the last two decades? Journal of Experimental Criminology 1, 9–38.

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

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

Frimmel, W., Halla, M., Paetzold, J., 2019. The intergenerational causal effect of tax evasion: Evidence from the commuter tax allowance in Austria. Journal of the European Economic Association 17, 1843–1880.

Gaigné, C., Zenou, Y., 2015. Agglomeration, city size and crime. European Economic Review 80, 62–82.

Galiani, S., Cruz, I.L., Torrens, G., 2018. Stirring up a hornets' nest: Geographic distribution of crime. Journal of Economic Behavior & Organization 152, 17–35.

Giordano, P.C., 2010. Legacies of crime: A follow-up of the children of highly delinquent girls and boys. Cambridge studies in criminology (Cambridge University Press) .

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

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

Hartley, R.P., Lamarche, C., Ziliak, J.P., 2022. Welfare reform and the intergenerational transmission of dependence. Journal of Political Economy 130, 523–565.

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.

Hjalmarsson, R., Lindquist, M.J., 2019. The causal effect of military conscription on crime. The Economic Journal 129, 2522–2562.

Houser, D., List, J.A., Piovesan, M., Samek, A., Winter, J., 2016. Dishonesty: From parents to children. European Economic Review 82, 242–254.

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.

Laferrere, A., Wolff, F.C., 2006. Microeconomic models of family transfers, in: Kolm, S.C., Ythier, J.M. (Eds.), Applications. Elsevier. volume 2 of *Handbook of the Economics of Giving, Altruism and Reciprocity*, pp. 889–969.

Laub, J.H., Nagin, D.S., Sampson, R.J., 1998. Trajectories of change in criminal offending: Good mariages and the desistance process. American Sociological Review 63, 225–238.

Lee, L.F., Liu, X., Patacchini, E., Zenou, Y., 2021. Who is the key player? A network analysis of juvenile delinquency. Journal of Business and Economic Statistics 39, 849–857.

Lindquist, M.J., Zenou, Y., 2014. Key players in co-offending networks. CEPR Discussion Paper No. 8889 .

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.

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

Massenkoff, M., Rose, E.K., 2020. Family formation and crime. Working Paper .

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

Michaeli, M., Wu, J., 2022. Fighting polarization with (parental) internalization. Journal of Economic Behavior & Organization 194, 124–138.

Murphy, F.X., 2019. Does increased exposure to peers with adverse characteristics reduce workplace performance? Evidence from a natural experiment in the US army. Journal of Labor Economics 37, 435–466.

Norris, S., Pecenco, M., Weaver, J., 2021. The effects of parental and sibling incarceration: Evidence from Ohio. American Economic Review 111, 2926–63.

O'Flaherty, B., Sethi, R., 2015. Urban crime. In G. Duranton, J. V. Henderson and W. C. Strange (Eds.), Handbook of Regional and Urban Economics. Amsterdam: Elsevier Publications. 5B, 1519–1621.

Olds, D.L., Henderson, C.R., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., Pettitt, L., Sidora, K., Morris, P., Powers, J., 1998. Long-term effects of nurse home visitation on children's criminal and antisocial behavior: 15-year follow-up of a randomized controlled trial. Journal of the American Medical Association 280, 1238–1244.

Oreopoulos, P., Page, M.E., Stevens, A.H., 2006. The intergenerational effects of compulsory schooling. Journal of Labor Economics 24, 729–760.

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

Patacchini, E., Zenou, Y., 2012. Juvenile delinquency and conformism. Journal of Law, Economic, and Organization 28, 1–31.

Reiss, A.J., Farrington, D.P., 1991. Advancing knowledge about co-offending: 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.

Stevenson, M., 2017. Breaking bad: Mechanisms of social influence and the path to criminality in juvenile jails. The Review of Economics and Statistics 99, 824–838.

Sutherland, E.H., Cressey, D.R., 1966. Principles of criminology. Chicago, IL: Lippincott .

Thompson, O., 2014. Genetic mechanisms in the intergenerational transmission of health. Journal of Health Economics 35, 132–146.

Varvarigos, D., 2020. Cultural transmission, education-promoting attitudes, and economic development. Review of Economic Dynamics 37, 173–194.

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

Wildeman, C., Andersen, S.H., 2017. Paternal incarceration and children's risk of being charged by early adulthood: Evidence from a Danish policy shock. Criminology 55, 32–58.