Crime, Broken Families, and Punishment*

Emeline Bezin† Thierry Verdier‡ Yves Zenou§

March 25, 2021

Abstract

We develop a two-period overlapping generations model in which both the structure of the family and the decision to commit crime are endogenous and the dynamics of moral norms of good conduct (honesty trait) is transmitted intergenerationally by families and peers. Having a father at home might be crucial to prevent susceptible boys from becoming criminals, as this facilitates the transmission of the honesty trait against criminal behavior. By "destroying" biparental families and putting fathers in prison, we show that more intense crime repression can backfire at the local level because it increases the possibility that criminals' sons become criminals themselves. Consistent with sociological disorganization theories of crime, the model also explains the emergence and persistence of urban ghettos characterized by a large proportion of broken families, high crime rates, and high levels of peer socialization, which reinforce criminal activities. Finally, we discuss the efficiency of segregation, family and education policies in terms of long-term crime rates.

JEL Classification: K42, J15, Z13.

Keywords: Crime, social interactions, cultural transmission, social disorganization theory, broken families, urban ghettos, segregation.

---

*We thant the editor and three anonymous referees for helpful comments. This research was undertaken under the project TECTACOM. The financial support from the European Research Council (ERC) grant number 324004 is gratefully acknowledged.

†Paris School of Economics and IRES (UCL). France. E-mail: emeline.bezin@gmail.com.

‡Paris School of Economics and Ecole des Ponts ParisTech, PUC-Rio and CEPR. Email: thierry.verdier@ens.fr.

§Monash University, Australia, and CEPR. E-mail: yves.zenou@monash.edu.

1 Introduction

Over the past 30 years, incarceration in the United States has increased 500 percent so that it is now the world leader with almost 2.3 million individuals involved in the full American criminal justice system (Prison Policy Initiative, 2020).[1] During the same period, the United States has experienced a sharp reduction in crime, which suggests a negative relationship at the *macro* level between crime and incarceration. However, at the *micro* level, especially at the *neighborhood* level, this is less true. There is evidence showing that, locally, increasing incarceration may, in fact, increase rather than decrease crime (see e.g, Goffman, 2014, and the whole social disorganization theory that we survey in Section 1.1 below).

In this study, we argue that, locally, incarceration can backfire by having a positive effect on crime because it puts fathers in prison, away from their family, and therefore can increase the probability that their children become criminals. Indeed, by "destroying" biparental families, incarceration threatens the earning power of the remaining family members and makes children growing up in these broken families much more vulnerable to crime influence and negative community experiences (or peer effects). This, in turn, encourages their criminal participation. Moreover, when the residential location of these families is taken into account, this may explain why some neighborhoods end up with a large proportion of broken families and a very high crime level. This is in accordance to a recent study by Chetty et al. (2020) who show that the share of broken families/single-parent households significantly predicts intergenerational mobility, particularly for African American individuals in the United States.

We develop a two-period overlapping generations model in which half of the population are male and the other half are female. Each individual lives for two periods. In the first period, which corresponds to childhood, individuals do not make economic choices but are subject to socialization. They belong to either single-mother or biparental families and can inherit (through both vertical and horizontal transmissions) either the "honest" cultural trait or the "dishonest" cultural trait. Both traits refer to "crime ethics", where being "dishonest" means having a "bad" crime ethic. A person with a bad crime ethic is more

[1] This includes 1,833 state prisons, 110 federal prisons, 1,772 juvenile correctional facilities, 3,134 local jails, 218 immigration detention facilities, 80 Indian Country jails as well as military prisons, civil commitment centers, state psychiatric hospitals, and prisons in the U.S. territories.

inclined to commit crime in the next period than a person endowed with the "honest" trait. At the beginning of the second period, each child becomes an adult and his or her parent(s) die. Then, men and women are matched to form a household. After matching with a female, given the trait he has inherited in the first period, each male decides whether to commit crime. If a male is not a criminal or if he is a criminal and is not caught, then he forms a biparental family. If he is a criminal and gets arrested, then he spends some time in prison and therefore his wife raises their offspring alone as a single mother. Then, each (biparental and monoparental) family exerts a socialization effort to influence their offspring to adopt the honesty trait. In this context, we analyze the dynamics of the proportion of honest individuals in the population and the long-run crime rate.

Our first contribution to the literature is to highlight the interplay of two elements. The first one is the classical Beckerian *deterrence effect*, indicating that an increase in p, the probability of being caught, reduces crime by decreasing the expected returns from criminal activities. The second is a *social disorganization effect*, recognizing the fact that an increase in incarceration disrupts the family structure, which, in turn, has a negative long-term impact on the transmission of a culture of honesty in the population. The importance of this last effect depends on how widespread is the initial culture of honesty in the society (third element).

Our second contribution is to analyze a policy that aims at reducing overall crime. We show that the effectiveness of an incarceration policy (i.e., increasing p) depends on q_0, the initial proportion of honest individuals in the population. If q_0 is small, any rise in incarceration (higher p) reduces long-run crime. This is the standard Beckerian model. However, when q_0 is sufficiently large, incarceration policies can backfire because a non-monotonic relationship between long-run crime and the degree of crime repression emerges. Specifically, we show that if the level of incarceration p is already sufficiently high, then increasing p further *actually* increases long-run crime in the population. Indeed, when q_0 is sufficiently high, increasing p "destroys" the structure of families by increasing monoparental families as the expense of biparental families. Because this effect is stronger than the deterrence effect, an increase in p raises the long-run crime level.

Our third contribution is to consider the spatial consequences of crime and social disorganization by endogenizing the location choices of families. Each individual has to reside in one of two neighborhoods in the city. All individuals bid for housing and we analyze the resulting urban equilibrium. We show that two urban equilibria are possible. In the segregated equilibrium, all honest families live in one neighborhood, while all dishonest families reside in the other neighborhood. In the integrated equilibrium, half of the honest families reside in one neighborhood, while the other half reside in the other neighborhood. We show that spatial segregation strengthens social disorganization and vice versa. In particular, we demonstrate that depending on the initial conditions, we can end up in the long run with a segregated equilibrium in which, in one neighborhood, the crime rate is high and most families are broken, whereas the opposite is true in the other neighborhood.

Kleiman (2009) argues that simply locking up more people for lengthier terms (as has been done in recent years in the United States) is not a workable crime control strategy. As Kleiman shows, "zero tolerance" is nonsense: there are always more offenses than there is punishment. Is there an alternative to brute force? In this study, we argue that, at the *neighborhood* level, deterrence can be counterproductive to reduce crime if it is *not* implemented with a social policy that helps single-mother families "educate" their children into pro-social behaviors. It is the failure to form and maintain intact families that explains the incidence and the persistence of high crime in a neighborhood. Chetty et al. (2018) construct an Opportunity Atlas (i.e., a publicly available atlas of children's outcomes in adulthood by Census tract using anonymized longitudinal data covering nearly the entire U.S. population), which suggests that children growing up in certain areas face lower social mobility than children living in other neighborhoods. Our paper provides a microfoundation for such mechanism, suggesting that incarceration policies, coupled with the socialization of children with peers of single-parent families, lead to highly persistent crime rates at the local level, and, consequently, low social mobility. As a result, an effective crime policy should also take into account the impact of incarceration on the structure of families, and how this affects their urban mobility and the criminal behavior of future generations.

1.1 Related literature

Our work is directly related to the so-called *social disorganization theory* initiated by Shaw and McKay (1942). This line of research investigates the relationship between the social organization of neighborhoods (or communities), the process of the growth of large cities, and the evolution of crime behavior. According to this theory, crime appears in communities characterized by social disorganization and is perpetuated through a process of cultural transmission whereby criminal norms and values are transmitted from generation to generation. Social disorganization theory can then explain the variations in criminal offending and delinquency, across both time and space, by examining differences in institutions (e.g., family, school, neighborhood, church, friend networks).[2]

To explain criminal behavior, we focus on two key aspects of social disorganization theory: the *structure of the family* and *location*. There are models aiming at explaining the variation in crime within cities based on the spatial variation of police forces, land prices, and access to formal jobs (Freeman et al., 1996; Glaeser and Sacerdote, 1999; Zenou, 2003; Verdier and Zenou, 2004; Decreuse et al., 2015; Gaigné and Zenou, 2015; O'Flaherty and Sethi, 2015; Deng and Sun, 2017; Galiani et al., 2018).

Our study is the first to incorporate social disorganization in an economic model of crime and to provide a new mechanism that explains the spatial variation of crime based on the family structure, peers, and the local intergenerational cultural transmission processes. We show under which conditions there is a steady-state equilibrium with segregation in which one neighborhood is characterized by high levels of crime, broken families, and a large proportion of single-mother families and another neighborhood in which the opposite is true. In particular, we provide some conditions under which an *urban ghetto* emerges and persists over time.

Our paper is also related to the large literature suggesting that peer effects are important in criminal activities (Glaeser et al., 1996; Ludwig et al., 2001; Calvó-Armengol and Zenou, 2004; Kling et al., 2005; Ballester et al., 2010; Bayer et al., 2009; Damm and Dustmann, 2014; Cortés et al., 2020; Lee et al., 2021).

In our study, peers also play an important role since they determine whether someone adopts the honesty trait and thus whether he or she is more likely to commit crime and form a biparental family.[3] Compared to this literature, we add two dimensions related to criminal activities: the evolution of a culture of honesty (that affects the intrinsic willingness to commit crime) and how the structure of the family responds and also determines criminal behaviors.

Finally, our paper is related to the literature on cultural transmission initiated by Bisin

[2]For a recent overview of social disorganization theory, see Porter et al. (2016) and Kubrin and Wo (2016).

[3]Contrary to the standard peer-effect literature, children are not influenced by other children of the same age but by adults in the community.

and Verdier (2000, 2001), who argue that the transmission of a particular trait is the outcome of a socialization inside (parents) and outside the family (peers and role models). This model has been extended in different directions (see, e.g., Hauk and Saéz-Martí, 2002; Bisin et al., 2011; Büchel et al., 2014; Panebianco, 2014; Prummer and Siedlarek; 2014; Cheung and Wu, 2018; Verdier and Zenou, 2018; Della Lena and Dindo, 2019). The closest paper to ours is that of Bethencourt and Kunze (2014), who also model the dynamics of honesty in the Bisin-Verdier framework to explain criminal behavior in the United States but have more a macro perspective. Contrary to our model, they do not introduce the family structure as the key determinant of honesty transmission and do not study the impact of incarceration on crime, which is the main focus of our paper. Furthermore, they do not have a model with two neighborhoods where crime is spatially differentiated and where urban ghettos with local cultures of crime emerge in equilibrium.

2 Some stylized facts and ethnographic studies

In this section, we discuss some stylized empirical facts and ethnographic studies to motivate three salient elements that our framework will incorporate: crime repression, the connection between incarceration, family and crime behavior, and the spatial dimension of crime and family social disorganization.

2.1 Facts

2.1.1 Fact 1: Parental incarceration, family structure, and crime behavior

It is estimated that over 800,000 children have a parent in prison on any given day in EU countries and 1 of every 12 American children, more than 5.7 million kids under age 18, have experienced parental incarceration at some point during their lives (Ghost, 2016). In the U.S., from 1980 to 2000, the number of children with incarcerated father has increased from 351 000 to 2.1 million (i.e. to 3% of all US children). Black children born in 1990 had a 25% chance to have a father in prison.

With respect to family structure, there has been an increase in the rates of mother-headed households[4] (Bureau of the Census, 2016), and children growing up in single-mother families

[4]During 1960-2016, the percentage of children living only with their mother nearly triples from 8 to 23 percent.

are at greater risk of developing behavioral problems (Barber and Eccles, 1992; Dornbusch et al., 1985; Kellam et al., 1977) and engage in a variety of high-risk behaviors such as crime and delinquency (Stern et al., 1984; Turner et al., 1991; Florsheim et al., 1998).[5]

A quantitative literature in sociology suggests a significant impact of paternal incarceration on the fabric of family life, the modes of parenting, and the consequences on children's behaviors and socio-economic well-being. In particular, paternal incarceration is shown to be positively affecting juvenile delinquency (Wilderman, 2010; Murray and Farrington, 2005).[6]

In addition, there is a literature that provides casual evidence of substantial intergenerational associations in criminal behavior, i.e., incarcerating a delinquent father can raise the likelihood of delinquency of his son. For example, Hjalmarsson and Lindquist (2012) show that sons with criminal fathers have 2.06 times higher odds of having a criminal conviction than those with noncriminal fathers and one additional paternal sentence increases sons' convictions by 32 percent. Wildeman and Andersen (2017) exploit variation from a reform that decreased the risk of incarceration for some crimes and find that parental incarceration increases criminal behavior for boys. Finally, Dobbie et al. (2018) estimate the causal effects of parental incarceration on their children's convictions. For children from the most disadvantaged families, they show that parental incarceration increases teen convictions (i.e., criminal convictions of the children between the ages of 15 and 17) by 10.2 percentage points

2.1.2 Fact 2: Spatial aspects of crime and family structure

It is well documented that crime is highly concentrated in a limited number of areas within cities (Anselin et al., 2000; Brantingham 2016). In U.S. metropolitan areas, after controlling for education, crime rates are higher in central cities than in suburbs. For instance, between 1985 and 1992, crime victimizations averaged 0.409 per household in central cities compared with 0.306 per household in suburbs (Bearse, 1996, Figure 1).[7] There is also evidence that central cities have more single-mother families than suburbs. Indeed, between 1970 and 1994, the percentage of children living in single-mother households rose from 12.8% to 30.8%. The

[5]For example, by using data from AddHealth, Cobb-Clark and Tekin (2014) find that adolescent boys engage in more delinquent behavior if there is *no* father figure in their lives. Adolescent girls' behavior is largely independent of the presence (or absence) of their fathers.

[6]For a review of this literature, see Wildeman (2020).

[7]While, for the more recent period (1990-2008), crime rates have substantially declined in large U.S. cities, central cities are still characterized by higher rates of violent crime than suburbs (Kneebone and Raphael, 2011).

rate of increase was particularly sharp among *inner-city minority families* (Bureau of the Census, 1994) for whom the scarcity of employment opportunities made it more difficult for men to fulfill their expected roles as fathers and husbands (Wilson, 1987).

2.1.3 Ethnographic studies

The salient connection between crime repression, family and crime at the local level is also well illustrated by an abundant amount of ethnographic studies in sociology and criminology. Most qualitative works emphazise the range of emotional and behavioral stress children have, the financial loss, the social stigma, the poor visiting arrangements, and the care problems that families often endure during and after parental incarceration (Boswell and Wedge, 2002; Katz, 2002; Braman, 2004). These difficulties have obvious negative consequences for the capacity of these families to cope with parenting and child management issues.

In the Online Appendix A, we outline different ethnographic studies that show that incarceration reduces in various ways the parenting capacity of the family of the imprisoned member and this is strongly associated with children's antisocial behaviors and later propensity to enter into delinquency and crime. In particular, Goffman (2014) shows how, at the neighborhood level, repression and incarceration against African American communities can be ineffective and can increase rather than decrease crime.

3 Benchmark model

Consider a two-period overlapping generations model populated by a continuum of agents with a mass equal to two. One half of the population are male and the other half are female. Each individual lives for two periods. In the first period, which corresponds to childhood, individuals do not make economic choices but are subject to socialization. They belong to either single-mother or biparental families and can inherit (through both vertical and horizontal transmissions) either the "honest" trait or the "dishonest" trait. Being "dishonest" means having a "bad" crime ethic and being "honest" means having a moral norm of good conduct.[8] At the beginning of the second period, each child becomes an adult and the parents die. Men and women are matched to form a household. To keep the population of

[8] We use the terms honesty/dishonesty as the trait parents influence but, clearly, our analysis would go through with any trait/ability related to criminal versus non-criminal earnings.

men and women constant, we assume that each household has two children, a boy and a girl. After matching with a female, given the trait he has inherited in the first period, each male has to decide whether to become a criminal. For simplicity, we assume that female workers always choose legal activities. If a male is not a criminal or if he is a criminal and is not caught, then he forms a biparental family. If he is a criminal and is arrested, then he spends some time in prison; therefore, his wife raises their offspring alone as a single mother. Clearly, the structure of the family depends on the father's criminal behaviors.[9] This simple mechanism implies that family disruption increases with criminality. Then, each family exerts a socialization effort to influence its offspring to adopt the honesty trait.

Let us first analyze the decision of male individuals in the second period.

3.1 Criminal decision

The male population can be of two types: they can either be "honest" (type h), have a good crime ethic, and thus bear a psychological cost K when they engage in criminal activities or be "dishonest" (type d), have a bad crime ethic, and thus bear no psychological cost when committing crime. Let β be the proceeds from crime, p, the probability of being arrested, σ, the cost of punishment, w, earnings in the legal labor market,[10] and θ, an idiosyncratic component that captures the (inverse) individual ability to carry out criminal activities. The variable θ is *uniformly* distributed on the support $[0, 1]$. It is important to differentiate between K and θ, where the former captures the *moral cost* of committing crime only for "honest" individuals who have a good crime ethic, while the latter refers to an (inverse) *ability* to commit crime for the whole population. An honest person with a very low θ may be more likely to commit crime than a dishonest person with a very high θ as long as K is not too high.[11]

An individual of type d with a given θ chooses to engage in criminal activities if and only

[9] A father who is arrested does not have to spend all his time in prison in period 2. It suffices that he spends a sufficiently long time in prison during the socialization period of his children. Empirical evidence suggests that adult men spend a discontinuous time in prison, first, in different local prisons and then, over time, in federal prisons because they tend to commit more serious crimes. See, for example, Goffman (2014).

[10] We assume that the legal wage is the same for all individuals. Here, we mainly focus on low-skilled workers who all earn the same *minimum wage*.

[11] Instead of having θ, the inverse ability of committing crime, we could have assumed that individuals have different proceeds from crime, i.e., β is drawn from a continuous distribution. In this case, if we further assume that the benefits from crime is independent of the probability to be incarcerated (so that the expected gain from crime becomes $\beta - p\sigma$), then the two formulations are equivalent.

if $(1-p)\beta - p\sigma - \theta > w$. Therefore, θ^d, the proportion of type$-d$ individuals who engage in criminal activities, is given by

$$\theta^d = (1-p)\beta - p\sigma - w. \tag{1}$$

Similarly, an individual of h with a given θ chooses to engage in criminal activities if and only if $(1-p)\beta - p\sigma - K - \theta > w$. Thus, θ^h, the proportion of type$-h$ individuals who engage in criminal activities, is equal to

$$\theta^h = (1-p)\beta - p\sigma - w - K. \tag{2}$$

We have $\theta^d > \theta^h$, which means that honest individuals have a lower probability of engaging in criminal activities. We assume throughout that $\beta - w - K > 0$ so that $\theta^h > 0$ and $\theta^d > 0$ for some $p \in [0,1]$. For any $p > \frac{\beta - w - K}{\beta + \sigma} \equiv \overline{p}_1$ (resp. $p > \frac{\beta - w}{\beta + \sigma} \equiv \overline{p}_2$), no honest (dishonest) individual engages in criminal activities.

Let q_t be the proportion of honest individuals in the economy at time t. Define $C : [0,1] \times [0,1] \to [0,1]$ such that $C_t := C(q_t, p)$, where C_t is the proportion of criminal individuals at time t and $C(q_t, p)$ is given by

$$C(q_t, p) = \begin{cases} q_t\theta^h + (1-q_t)\theta^d, & \forall p \in [0, \overline{p}_1] \\ (1-q_t)\theta^d, & \forall p \in [\overline{p}_1, \overline{p}_2] \\ 0, & \forall p \geq \overline{p}_2 \end{cases} \tag{3}$$

Naturally, $C(q_t, p)$ is decreasing in q_t since the higher the proportion of honest individuals in the population, the lower is the level of crime in the economy. Further, $C(q_t, p)$ is decreasing in p because of the *deterrence effect*; in other words, the higher the incarceration rate p,[12] the lower is the expected gain from criminal activities and thus the lower is the crime rate $C(q_t, p)$. By using the values of θ^d and θ^h defined in (1) and (2), we obtain

$$C(q_t, p) = \begin{cases} -q_tK + (1-p)\beta - p\sigma - w, & \forall p \in [0, \overline{p}_1] \\ (1-q_t)\left[(1-p)\beta - p\sigma - w\right], & \forall p \in [\overline{p}_1, \overline{p}_2] \\ 0, & \forall p \geq \overline{p}_2 \end{cases}, \tag{4}$$

[12]To refer to p, we use the "probability of being arrested" and "incarceration rate" interchangeably because we assume that someone who is arrested is automatically incarcerated.

where $\overline{p}_1 = (\beta - w - K) / (\beta + \sigma)$ and $\overline{p}_2 = (\beta - w) / (\beta + \sigma)$.

3.2 Dynamics of traits and crime

The way in which individuals adopt h (the honesty trait) and d (the dishonesty trait) is modeled as follows. The child becomes honest if both parents' socialization effort succeeds *and* the role model met randomly by the child is honest.[13] Symmetrically, the child becomes dishonest if both socialization within the family fails *and* the role model is dishonest. However, if parents and society send conflicting messages about honesty (e.g., socialization by parents succeeds but the role model met is dishonest), the child is matched a second time with a role model (also met randomly) and adopts his or her trait.[14] This pattern of socialization is different from the Bisin–Verdier framework in two ways. First, whereas in Bisin and Verdier (2000, 2001), parents are typically imperfectly altruist and always prefer to transmit their own trait to their children, we assume that all parents consider the honesty trait to be preferable to the dishonesty trait. Second, we assume that a parent cannot transmit an honesty trait by him- or herself because there is strong evidence of contextual and peer effects in crime.[15] Hence, the social environment in which young individuals are brought up should play a more important role than simply parents in the individual decision to commit crime.

Recall that there are two types of families, $k = S, B$, where $k = S$ stands for a single-mother family and $k = B$ stands for a biparental family. Therefore, the probability P_t^{hk} of a child of a parent from a type$-k$ family ($k = S, B$) becoming type h (i.e., honest) at time

[13] In Section 4 below, we endogenize the location choices of all agents in a city. In that case, the meeting and matching with peers are *not* anymore random as they depend on the traits of the parents and on the location choice.

[14] Following Saéz-Martí and Sjögren (2008), we could have assumed that the cultural transmission was *biased* so that the probability of adopting a trait when learning from society (i.e., peers) is a non-linear function of popularity, which implies that more popular traits are copied with a probability higher than their population share (*frequency dependency*). Our main results would have been the same but without the fact that the individual needs to meet their peers twice in case of an unsuccessful transmission. In both cases, peer effects (horizontal transmission) are more important than parent socialization effort (vertical transmission) in transmitting a given trait if peers with that trait are in majority in the population.

[15] If we had only one match as in Bisin and Verdier (2001), then the role of the family would always be dominating.

t is given by:[16]

$$P_t^{hk} = 1 - P_t^{dk} = \tau_t^k q_t + \tau_t^k (1 - q_t) q_t + (1 - \tau_t^k) q_t^2 = q_t \left[2\tau_t^k (1 - q_t) + q_t \right], \tag{5}$$

where τ_t^k is the socialization effort of a type$-k$ parent and P_t^{dk} is the probability of a child of a parent from a type$-k$ family ($k = S, B$) becoming type d (i.e., dishonest) at time t. Indeed, a child can become type h if (i) the parent of type k is successful in transmitting the honesty trait (which occurs with a probability equal to the socialization effort τ_t^k) *and* the child (randomly) meets in the population a role model who is honest (which occurs with a probability q_t);[17] (ii) the type-k parent is successful in transmitting the honesty trait and the child meets first a dishonest role model, which occurs with a probability $1 - q_t$, (conflicting messages about honesty) but then the child is matched a second time with an honest role model; or (iii) the parent of type k is unsuccessful in transmitting the honesty trait (which occurs with a probability $1 - \tau_t^k$) and the child meets first an honest role model (conflicting messages about honesty) and then meets again an honest role model.

More generally, equation (5) states that the trait transitions depend on the probability of getting two positive influences; in particular, it depends on the probability of meeting *at least* one honest individual. This implies that children are influenced by the traits of the adults and not by their behavior. Indeed, the probability of meeting an honest individual is equal to the proportion of honest individuals in the community; similarly, the probability of meeting a dishonest individual is equal to the proportion of dishonest individuals. This implies that children are matched with members of the community *before* the outcome of law enforcement is realized.

Let us now model the parents' choice. All parents (of both types) value the honesty trait for their children. Let V^h (resp. V^d) be the gain of having a child of type h (resp. d) with $V^h > V^d$, $V^i < 1$, $\forall i \in \{h, d\}$. We do not have the superscript k in V^i because we assume that $V^{hS} = V^{hB} = V^h$ and $V^{dB} = V^{dS} = V^d$. In other words, the utility (disutility) of having a child of type h (type d) is the same for both types of parents. This assumption is made for simplicity and does not affect any of our results.

[16]Here we make the assumption that a dishonest father who is not caught and an honest father have the same impact on the transmission of the honest trait of their children. We relax this assumption in Section 3.5 below by allowing fathers to be role models for their children so that unarrested dishonest parents have a direct negative influence on their children while the opposite is true for honest parents.

[17]Observe that peer influences (i.e., horizontal transmission) of a child are not generated from socializing with children of the same age but from adults in the community.

We assume that the structure of the family influences children's socialization into values that influence criminal behaviors. In particular, single-mother families bear a higher socialization cost than biparental families because of their time constraints. Moreover, when fathers reside with their children, incarceration removes them from the household and incapacitates them from the labor market, depriving their families of a potential source of income, which raises the cost of socialization. Incarceration not only limits these contributions, but also threatens the earning power of remaining family members, who may sacrifice family and socialization to perform tasks previously carried out by the incarcerated father (Lynch and Sabol, 2004) or struggle to cover expenses associated with his incarceration such as legal representation, as well as maintaining contact through phone calls and visits (Comfort, 2008).

Let $c(\tau_t^S) = c^S \left(\tau_t^S\right)^2/2$ be the individual socialization cost of a *single-mother* family exerting effort τ_t^S and $c(\tau_t^B) = c^B \left(\tau_t^B\right)^2/2$ be the individual socialization cost of a *biparental* family exerting effort τ_t^B. We assume that $c^S > c^B > 1$. A parent from a type$-k$ family chooses his or her socialization effort τ_t^k at time t to maximize:[18]

$$u^k = P_t^{hk} V^h + P_t^{dk} V^d - c^k \frac{\left(\tau_t^k\right)^2}{2}. \tag{6}$$

By using (5), it is easily verified that the optimal socialization effort of a type-k family is given by

$$\tau_t^k = 2q_t(1-q_t)\Delta^k, \tag{7}$$

where $\Delta^k = \left(V^h - V^d\right)/c^k$. Observe that

$$\frac{\partial \tau_t^k}{\partial q_t} \gtreqqless 0 \Leftrightarrow q_t \lesseqqgtr \frac{1}{2}. \tag{8}$$

This fact implies that if, at time t, the majority of the people in the male population are dishonest (honest), then an increase in q_t, the proportion of honest individuals, leads to an increase (a decrease) in the parent's socialization effort. In other words, and using the

[18] In (6), we assume that parents only care about the honesty of their children. In the working paper version of this paper (Bezin et al., 2020; Section 3.4), we extend this utility function to incorporate the fact that parents care about the honesty as well as the welfare of their children. The results are similar with one main difference: the value of parenting increases when incarceration increases, as parents want to prevent their kids from going to jail.

terminology of Bisin and Verdier (2001), when $q_t < 1/2$, the socialization activities inside and outside the family are *cultural complements*, while when $q_t > 1/2$, they are *cultural substitutes*. Indeed, when $q_t < 1/2$ ($q_t > 1/2$), parents have more (less) incentive to socialize their children to the honesty trait, the more widely dominant is this trait in the population since *both* vertical and horizontal transmissions are needed for a trait to be successfully transmitted in the first place (see (5)). When the two send contradictory messages, the individual needs to meet other role models to determine which trait he or she will adopt. Therefore, when most people in the population are dishonest, the parent increases his or her effort with q_t, while the opposite is true when $q_t > 1/2$. In the context of crime and traits, the intuition is as follows. When the proportion of honest individuals is small (i.e., $q < 1/2$), a family cannot count on outside help. When q increases, families can see that the outside help is improving the chance that their offspring will be honest, and so they incease their socialization effort. On the contrary, when the proportion of honest individuals is high (i.e., $q > 1/2$), a family can count on outside help; thus, when q increases, families can just rely on outside help and, therefore, decrease their socialization effort.

The dynamics of the honesty trait h are then described by the following equation:

$$q_{t+1} = \underbrace{[1 - C(q_t, p)]}_{\text{fraction of non-criminals}} \; P_t^{hB} + \underbrace{(1-p)\, C(q_t, p)}_{\text{fraction of non-caught criminals}} \; P_t^{hB} + \underbrace{pC(q_t, p)}_{\text{fraction of caught criminals}} \; P_t^{hS}. \quad (9)$$

Indeed, there is a mass 1 of men in the population. Among them, $1 - C(q_t, p)$ are not criminals and $C(q_t, p)$ are criminals. Among the mass (or proportion) of criminals, $(1-p)\, C(q_t, p)$ of them are not arrested and $pC(q_t, p)$ are arrested. As a result, among the mass 1 of men, $1 - C(q_t, p) + (1-p)\, C(q_t, p)$ form biparental families, while $pC(q_t, p)$ form single-mother families. By using (5) and (7), this dynamic equation can be written as

$$q_{t+1} - q_t = q_t(1 - q_t) \left\{ 4q_t \,(1 - q_t) \left[\Delta^B - pC(q_t, p) \left(\Delta^B - \Delta^S \right) \right] - 1 \right\}, \quad (10)$$

where $C(q_t, p)$ is given by (4).

3.3 Steady-state equilibrium

We can now characterize the long-run cultural dynamics of honesty as a function of the probability p of arrest in society.

Assumption 1 *Suppose that: (i)* $\Delta^B > 1$ *and (ii)* $\left[\Delta^B - \frac{(\beta-w-K)^2}{4(\beta+\sigma)}\left(\Delta^B - \Delta^S\right)\right] < 1.$

Assumption $1(i)$ implies that $V^h - V^d > c^B$, which means that, for biparental families, the net benefit of having an honest child is larger than the cost of socialization per unit of effort. This implies that when p is low, and therefore the prevalence of single-mother families is low, socialization within the family is sufficiently effective to allow the survival of the honesty trait within the community (at least when the proportion of honest agents is initially high). Assumption $1(ii)$ implies that when the rate of single-parent families is high, socialization within families is not sufficiently effective to maintain the honesty trait within the community.[19]

Proposition 1 *Suppose that Assumption 1 holds.*

(i) If $p \in]\widehat{p}_1, \widehat{p}_2[$*, for any* $q_0 \in [0,1]$*, then the sequence* q_t *converges to* $q^* = 0$*.*

(ii) If $p \in [0, \widehat{p}_1] \cup [\widehat{p}_2, \overline{p}_2]$*, then there exists* $\underline{q} \in]0,1[$ *and* $\overline{q} \in]0,1[$ *for any* $q_0 \in \left[0, \underline{q}\right[$*, the sequence* q_t *converges to* $q^* = 0$*, while, for any* $q_0 \in \left[\underline{q}, 1\right]$*, the sequence* q_t *converges to* $q^* = \overline{q}$*.*

(iii) If $p \in [0, \widehat{p}_1]$*, then* $\partial\underline{q}/\partial p > 0$ *and* $\partial\overline{q}/\partial p < 0$*. If* $p \in [\widehat{p}_2, \overline{p}_2]$*, then* $\partial\underline{q}/\partial p < 0$ *and* $\partial\overline{q}/\partial p > 0$*.*

Consider first part (i) of Proposition 1. When p, the probability of being arrested, takes intermediate values, the number of monoparental families $pC(q_t, p)$ is at its maximum. As a consequence, the transmission of the honesty trait in the population is the weakest and the unique stable steady-state equilibrium has no honest individuals in the population ($q^* = 0$).

Consider now part (ii) of Proposition 1 where p can take either low or high values. We show that the steady-state value of q^* depends on the initial conditions as illustrated in Figure 1. As we know, for low or high values of p, the number of monoparental families $pC(q_t, p)$ is low. This is true for $p \in [0, \widehat{p}_1]$ because there is little chance that criminal fathers will be caught. This is also true for $p \in [\widehat{p}_2, \overline{p}_2]$ because most fathers are deterred from crime. In both cases, incarceration has a "neutral" effect on the structure of families and cultural transmission is mostly driven by the direct peer socialization effect (oblique transmission) and the endogenous family socialization effort (vertical transmission).

[19] All proofs can be found in the Mathematical Appendix.

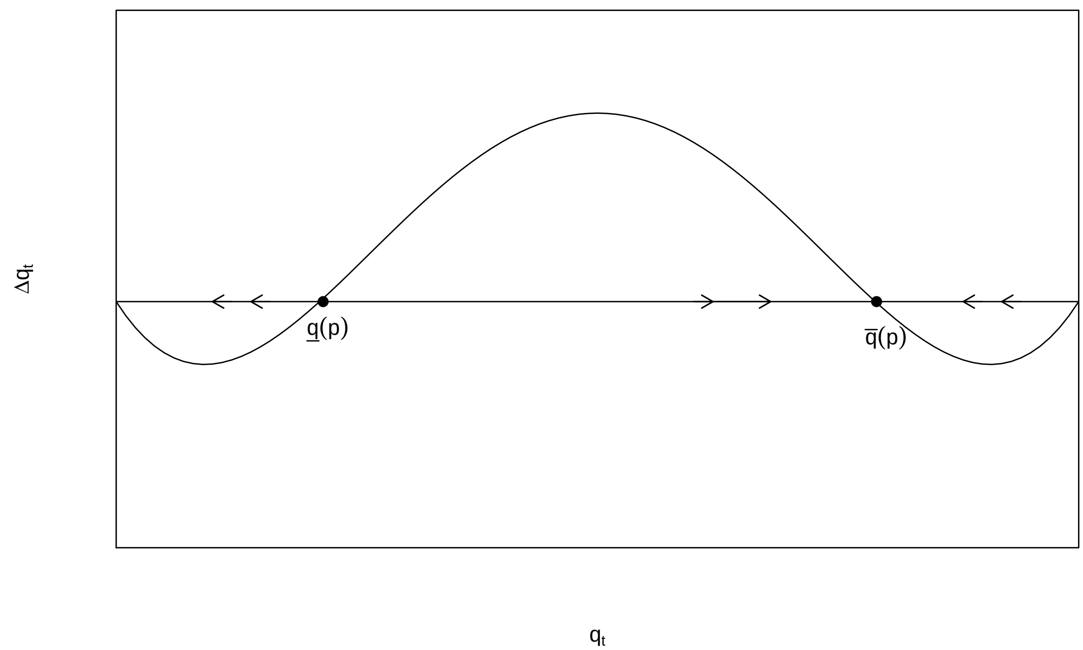

Figure 1: Dynamics of honesty with low or high incarceration rates, $p \in [0, \hat{p}_1] \cup [\hat{p}_2, \bar{p}_2]$

Consequently, if, at time $t = 0$, the proportion of honest individuals is low, the direct peer socialization effect and cultural complementarity effect between vertical and oblique socializations prevent a culture of honesty from taking off and the unique stable equilibrium is such that $q^* = 0$. Conversely, when q starts at a sufficiently high value (i.e., $q_0 > \underline{q}$), then the economy converges to an interior solution $q^* = \overline{q}\,(p) > \underline{q}$. Indeed, when q_t is above the threshold value $\underline{q}$, the cultural complementarity effect tends to push for an increase in q_t. Since, at the same time, fewer people are committing crime, more biparental families are formed and this reinforces the diffusion of the honesty trait in society. As q_t continues to grow, the cultural substitutability effect between family and peers starts to kick in and this leads to a stabilization of q_t to $q^* = \overline{q}$. As a result, $\underline{q}$ is a *tipping point* since when the proportion of honest individuals is below (above) $\underline{q}$, the sequence q_t converges to $q^* = 0$ ($q^* = \overline{q}$).

More generally, this proposition highlights the interaction between the *deterrence effect* of p and its impact on the family structure. When p is low or very high, there are few single-mother families and thus the transmission of the honesty trait is mostly driven by the interaction between the family socialization and peer effects. When p takes intermediate values, many individuals decide to commit crime and these are likely to be arrested. This leads to a relatively large proportion of broken families and a weak transmission of the honesty trait.

Proposition B1 in Online Appendix B.1 provides some comparative statics results of the interior stable equilibrium $q^* = \overline{q}$, and of the tipping point $\underline{q}$ that characterizes the lower bound of the basin of attraction of $q^* = \overline{q}$. We find that when w, the outside opportunity in the legal market, or σ, the cost of the punishment, increases, or β, the proceeds from crime, decreases, the steady-state proportion of honest individuals in the population rises and the size of its basin of attraction increases.

In our framework, although *incarceration* (i.e., an increase in p) always reduces *short-run crime* $C(q_t, p)$ (see (4)), it may have an ambiguous effect on *long-run crime*, $C(q^*, p)$, because of its impact on long-run honesty through the structure of the family. To highlight the impact of incarceration on long-run crime, let us define the functions $\underline{q}(p)$ and $\bar{q}(p)$ on $[0,1] \times [0,1] \rightarrow]0,1[$ implicitly defined by $q_{t+1} - q_t = 0$ (see (10)).

Proposition 2 *Suppose that Assumption 1 holds.*

(i) $\forall q_0 < \underline{q}(0)$, *an increase in incarceration (i.e., higher* p*) reduces long-run crime.*

(ii) $\forall q_0 > \max\{\underline{q}(\hat{p}_1), \underline{q}(\hat{p}_2)\}$, *an increase in incarceration has a non-monotonic impact on long-run crime.*

(iii) $\forall q_0 \in [0,1]$, *strong incarceration policies (i.e.,* $p \geq \bar{p}_2$*) minimize long-run crime.*

When the initial proportion of honest agents, q_0, is low, the bad peer effects are so strong that whatever the structure of the family, the community cannot maintain a culture of honesty. In such a case, an increase in incarceration (higher p) always reduces long-run crime since only the deterrence effect is at stake. This case is equivalent to the classical punishment effect in the Beckerian model. However, when q_0 is higher, our predictions differ from those of the Beckerian model since the structure of the family affects the culture of honesty and thus long-run crime. Indeed, by destructuring families, incarceration can have a detrimental effect on long-run crime.

These results are illustrated in Figure 2, which depicts the relationship between incarceration policies and long-run crime. When q_0 is low (dotted line), any rise in incarceration (higher p) reduces long-run crime. This corresponds to the standard Beckerian model.

When q_0 becomes larger (solid curve), the impact of incarceration on long-run crime is non-monotonic. In particular, Figure 2 shows that around $\hat{p}_1$, *increasing incarceration can backfire since it raises long-run crime.* This is because when q_0 is sufficiently high, increasing p, the probability of being arrested, "destroys" the structure of families by increasing mono-parental families as the expense of biparental families. Because this effect is stronger than the deterrence effect, increasing p raises long-run crime. Furthermore, in Figure 2, $\hat{p}_1$ and $\hat{p}_2$ are the two bifurcation points, meaning that slightly increasing $\hat{p}_1$ or slightly decreasing $\hat{p}_2$ leads to a large increase in long-run crime.

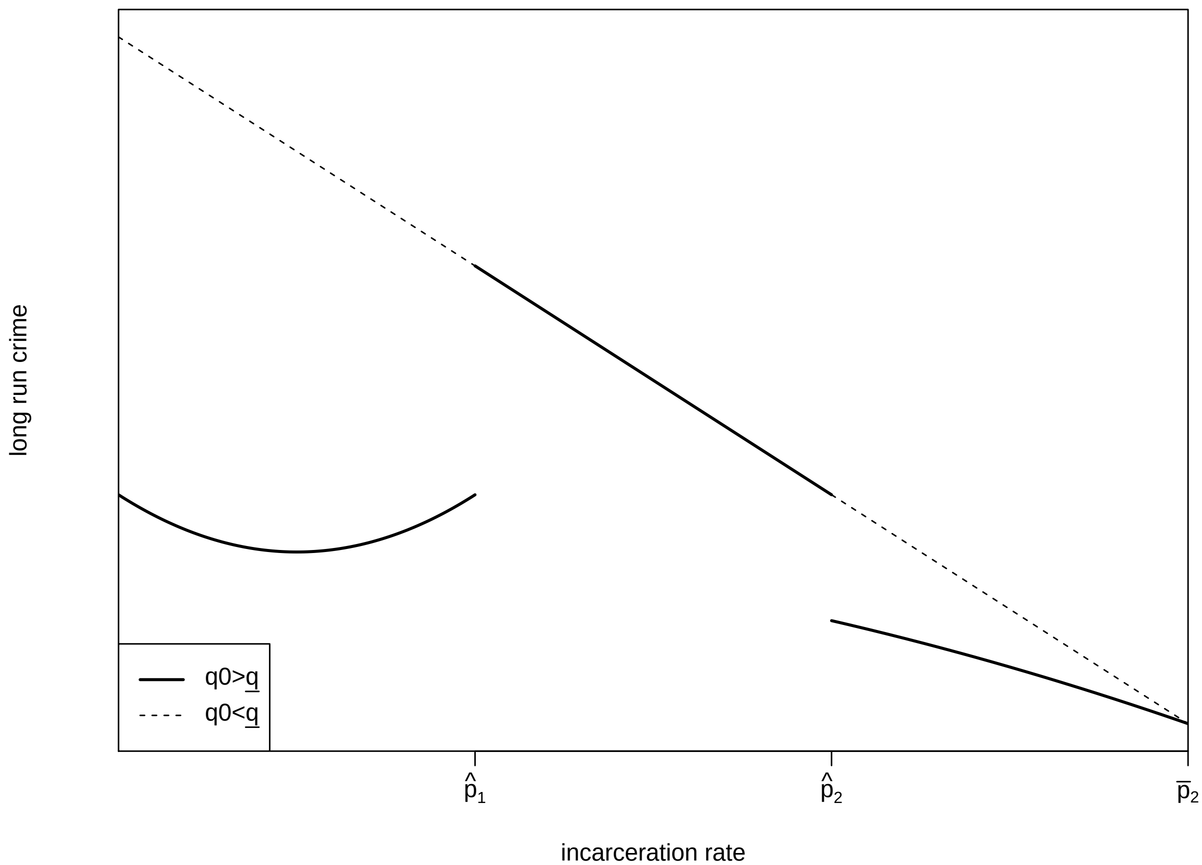

Figure 2: Long-run crime as a function of the incarceration rate p.

3.4 Education versus incarceration

So far, education was not included in our model. Evidence shows that education has a large impact on crime. For example, in 1997, 75 percent of state and 59 percent of federal prison inmates in the US did not have a high school diploma (Harlow, 2003). In the United Kingdom, incarceration rates among men ages 21-25 were more than eight times higher for those without an education qualification (i.e., dropouts) relative to those with a qualification (Machin et al., 2011). Using a human capital-based model with estimations on US data, Lochner (2004), Lochner and Moretti (2004), and Fella and Gallipoli (2014) show that more schooling reduce most types of adult crimes.[20]

Following this literature, we extend our model to include education, which increases the opportunity cost of crime (i.e., the wage in the legal market) of criminals. Assume that the probability of being arrested increases with the number of police officers P. For simplicity, suppose that $p = P$ and that each police officer has a cost equal to 1. The government has a fixed budget R, which is divided between spending on crime repression p and providing education g (i.e., $R = p + g$). As stated above, education increases legal income since more education implies higher productivity, and the wage rate w in the legal sector is $w = w(g)$, with $w'(g) > 0$, $w''(g) < 0$, $w(0) = 0$.[21]

By using $g = R - p$, the proportion of honest and dishonest individuals who engage in criminal activities can now be written as

$$\theta^h = \max\{\beta - K - p(\beta + \sigma) - w(R - p), 0\}$$
$$\theta^d = \max\{\beta - p(\beta + \sigma) - w(R - p), 0\}.$$

Definition 1 *An incarceration policy p is efficient if it minimizes long-run crime.*

In other words, an efficient policy p solves

$$\min_p C(q^*(p), p), \tag{11}$$

where q^* is a function of p and is implicitly defined by Proposition 1. We cannot solve this minimization problem explicitly because we do not have an explicit form for q^*. Instead, we

[20]See Lochner (2011) for an overview of the effects of education on crime.

[21]Here, education spending g directly influences the criminal opportunity cost (i.e., the wage). Another possible direction would have been to assume that g directly affects the socialization of the children.

compare the crime rate under two incarceration policies: high and low. When incarceration is *high* (i.e., $p = R$), the whole budget is spent on incarceration; on the contrary, when incarceration is *low* (i.e., $p < R$), some money is spent on education.

We consider two cases. First, assume that the budget R is quite high. At $R = p$, we have $\beta - p(\beta+\sigma) - w(R-p) = \beta - R(\beta+\sigma)$, meaning that for any $R \geq \frac{\beta}{(\beta+\sigma)}$, $C(q_t, R) = 0$, $\forall q_t$. In particular, $C(q^*, R) = 0$. Hence, when the government's budget R is sufficiently high, trivially repressive policies are always efficient in the sense that they minimize long-run crime.

Second, consider the case when R takes lower values, i.e., $R \leq \frac{\beta-K}{(\beta+\sigma)}$.[22] When the returns to education are sufficiently high, i.e., $w'(0) > \beta + \sigma$, repression can have a detrimental impact on short-run crime. In this case, incarceration has a negative impact on both short-run and long-run crime (through the decrease in education and increase in incarcerated fathers). The trade-off between the positive effect on incentives at time t and negative effect on the transmission of cultural traits in the future disappears. Hence, this trivially shows that incarceration is not efficient.

Let us now examine in detail the dynamics of honesty under education. Proposition B2 in the Online Appendix B.2 shows that, when the budget R takes lower values and the initial proportion of honest individuals is sufficiently high, incarceration policies matter for the long-run level of honesty q^*. As above, incarceration policies have a positive impact on honesty by decreasing crime at time t and a negative one by increasing the proportion of incarcerated parents. However, compared to the case without education and no budget constraint (previous section), the *positive impact on crime is reduced* since it entails an additional opportunity cost, as it requires a reduction in spending on education. In particular, when public resources R are limited, whatever the incarceration rate, the crime rate is high; hence, any increase in p has a strong negative impact on the proportion of single-parent families. In that case, the *family disorganization effect* exceeds the *deterrence effect* and incarceration negatively impacts on long-run honesty.

Moreover, Proposition B3 in the Online Appendix B.2 shows that, when the initial proportion of honest individuals is sufficiently large and the budget R takes intermediate values, incarceration policies may have deep (negative) consequences on community culture because both the crime rate and the incarceration rate are high, which maximizes family disorgani-

[22] A similar reasoning applies for the case when $R \in [\frac{\beta-K}{(\beta+\sigma)}, \frac{\beta}{(\beta+\sigma)}]$. To ease the exposition, we skip this case and the results are available upon request.

zation.

To generate more general results on the efficient repression policy, we resort to numerical simulations for *any value* of p, namely we solve (11). Figure 3 displays the long-run proportion of honest individuals (q^*) and long-run crime rate $C(q^*)$ as a function of the incarceration policy p for an initially *low* proportion of honest agents, i.e., $q_0 = 0.1$. Figure 4 provides similar results but for an initially *high* proportion of honest agents, i.e., $q_0 = 0.6$. In both cases, $R = 0.2$.

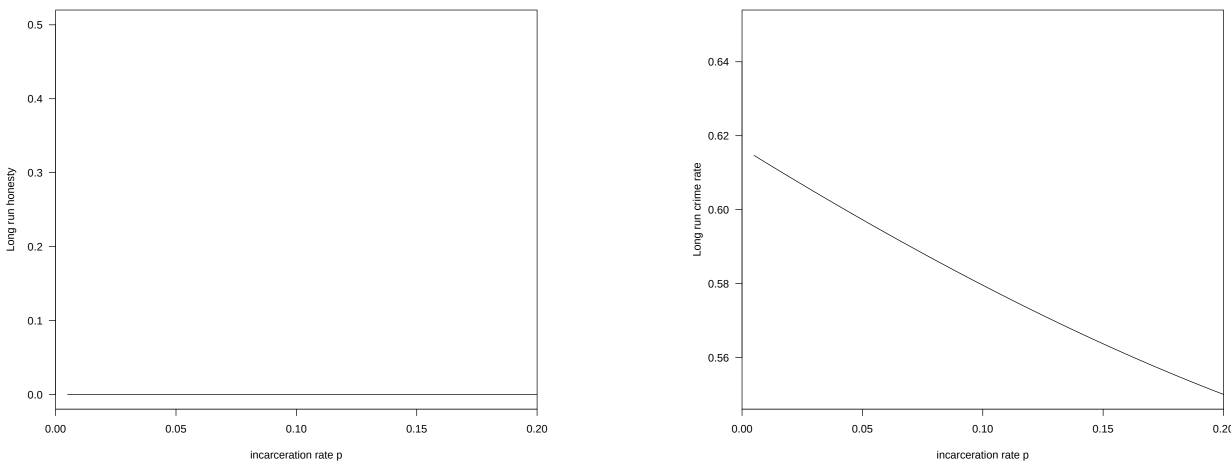

Figure 3: Long-run honesty (left) and the crime rate (right) as a function of the incarceration rate p for $q_0 = 0.1$. We set $w(g) = ag/(1 + g)$ and the parameters are such that $a = 0.5$, $\beta = 0.07$, $K = 1/7$, $\Delta = 0.01$, $c^B = 0.01$, $c^S = 2$, $\sigma = 0.05$, and $R = 0.2$.

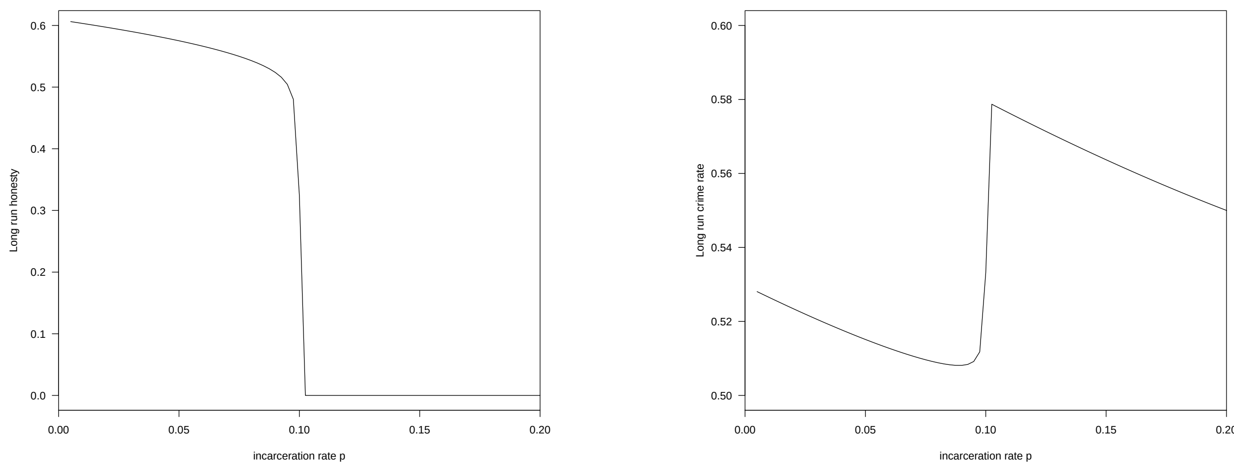

Figure 4: Long-run honesty (left) and the crime rate (right) as a function of the incarceration rate p for $q_0 = 0.6$. We set $w(g) = ag/(1 + g)$ and the parameters are such that $a = 0.5$ $\beta = 0.07$, $K = 1/7$, $\Delta = 0.01$, $c^B = 0.01$, $c^S = 2$, $\sigma = 0.05$, and $R = 0.2$.

As stated in Proposition B3, when the initial proportion of honest individuals is low (Figure 3), incarceration policies minimize both the short-run and the long-run crime rates, meaning that $q^* = 0$ and long-run crime is equal to zero.

When q_0 is high (Figure 4), the effect of an incarceration policy on crime is non-monotonic and may backfire. This example illustrates the result of Proposition B3. Indeed, the crime rate without incarceration ($p = 0$) is lower than that with maximal incarceration (i.e., repressive policies when $p = R$). When the government budget is low, the positive short-run effect of incarceration is mitigated as the legal income falls, and the opportunity cost of crime goes down. In fact, the negative effect of incarceration outweighs the positive effect of incarceration in the short run and strong incarceration increases long-run crime compared with the case of no incarceration. Moreover, zero incarceration is not efficient. The efficient incarceration policy then corresponds to an intermediate level of p, given by 0.10 in Figure 4. Figure 2 also shows that when incarceration is costless, the lowest value of long-run crime is obtained when p peaks. In Figure 4, this is not possible because the budget $R = 0.2$ is not sufficiently large for the deterrence effect to outweigh the social disorganization effect.[23]

[23]In Section B.3 of the Appendix, we consider another policy that provides financial help in the form of a subsidy to single-parent families. We show that this policy can reduce long-term crime under specific conditions.

3.5 "Good" versus "bad" fathers as role models

So far, our model has mainly focused on one channel through which paternal incarceration could affect child crime − household structure. But there are many other channels, through which parental incarceration could affect child crime, some of which would even decrease child crime. For example, if the father serves as a negative role model, then the father's imprisonment may prevent his child to mimic his behavior because imprisonment breaks the link between fathers and sons. Moreover, a father may be abusive, alcoholic, etc. In all these cases, removing this negative influence could reduce child crime.

In Section B.4 of the Online Appendix B, we formally study this model. In Proposition B6, we now find that, when the initial level of honesty is sufficiently high, the population may end up fully honest (i.e., $\overline{q}_2(p)$ can now be equal to 1), which was not possible in the benchmark model. The reason is that bi-parental families with an honest father are endowed with a better socialization technology than in the benchmark model since they do not only transmit honesty through family socialization efforts but also through fathers' role modeling. If, initially, the share of these types of families is sufficiently high and the father's role modeling channel is efficient, then the transmission of honesty at the population level is very strong so that the fraction of honest individuals converges to one.

We also show that, contrary to the benchmark model, when there is room for fathers to be role models, the impact of repressive policies varies with the initial culture of honesty. Typically, a rise in p now has two distinct impacts on the dynamics of honesty. On the one hand, it (negatively) affects the transmission of honesty by increasing the proportion of single-parent families. On the other hand, it has an impact on the long-run honesty by reducing the proportion of bi-parental families with "bad" fathers. The overall impact of p on steady-state honesty then depends on which type of family is more efficient in transmitting honesty.

The relative efficiency of these two different types of family depends, in turn, on q_t, the fraction of honest individuals in the population, because children born in single-parent families are much more subject to population composition effects (i.e., role models outside the family). Indeed, bi-parental families exert a higher socialization effort (because of their lower socialization cost). However, they also have a lower chance of transmitting honesty through role models outside the family (i.e., peers), since the child does not only mimic individuals from his neighborhood but also his father who is dishonest.[24]

4 Broken families, segregation, and crime

In this section, we provide a mechanism that explains why some neighborhoods end up with high crime, low housing prices, broken families, and the dissemination of a culture of crime, while other neighborhoods end up with the opposite. In particular, we would like to highlight the importance of broken families in the formation and persistence of ghettos.

4.1 The model

Let us return to the benchmark model without education and without parents as role models. We assume a city with two residential areas (or neighborhoods) indexed by $l = 1, 2$. The population of the city is a continuum of families of mass 2.

The timing is as follows. As before, the child is subject to socialization in the first period. At the end of the first period, each child has inherited a cultural trait $i = h, d$. At the beginning of the second period, when the child is now an adult, a male matches with a female and they decide in which neighborhood to reside. We assume that each family lives in one house and that the inelastic supply of houses within a residential area is normalized to 1. Therefore, each neighborhood encompasses a mass of 1 families. As in Verdier and Zenou (2004), we assume that each individual makes his or her location decision without knowing the value of his or her θ (i.e., the idiosyncratic ability to commit crime). Then, the types are revealed and individuals decide to commit crime or not. The assumption that types are revealed only after location choices have been made accounts for the relative inertia of the land market compared with the crime market. Obviously, individuals make quicker decisions in terms of crime than in terms of residential location. Next, the structure of the family is determined ($k = S, B$). Finally, at the end of the second period, each family exerts a socialization effort to influence its offspring to adopt the honesty trait.[25]

[24]In Section B.6 of the Online Appendix B, we extend our benchmark model in another direction by modeling the idea that some individuals may be wrongfully arrested and convicted. This is well-documented. We show that this will reinforce our results since there will be more single-mother families than in the benchmark model and, therefore, high levels of incarceration can break even more families and yield higher crime rates.

[25]Observe that, in the benchmark model, peer effects were independent of the family structure. Here, we endogenize the relationship between family structure and peer effects by letting families deciding their

Let $Q_t = q_{1,t} + q_{2,t}$ be the mass of honest agents. The bid rent for a parent of type (i, n), i.e., a parent of type $i = h, d$ in neighborhood $n = 1, 2$, at time t is denoted by $\rho^i_{n,t}$. Without loss of generality, we impose that $q_{1,t} \geq q_{2,t}$ (i.e., the proportion of agents endowed with the honesty trait h is higher in neighborhood 1) and $\rho^i_{2,t} = 0$ (the land rent is zero in neighborhood 2). The utility of a parent of type $k = S, B$ living in neighborhood n is still given by (6), that is,

$$u^k_n = P^{hk}_{n,t} V^h + P^{dk}_{n,t} V^d - c^k \frac{\left(\tau^k_{n,t}\right)^2}{2}. \tag{12}$$

Utility u^k_n is clearly affected by location n through $q_{n,t}$, the proportion of honest individuals in neighborhood n (see (5)). In particular, parental effort $\tau^k_{n,t}$ depends on the neighborhood in which the family lives since $\tau^k_{n,t}$ is strongly affected by $q_{n,t}$.

Let us now calculate the *expected* utility of an individual of type (i, n) before the revelation of θ. To simplify the presentation, we skip the time index. We have

$$U^i_n = \int_0^{\theta^i} \left[(1-p)\,\beta - p\sigma - K\,\mathbf{1}_{i=h} - \theta\right] d\theta + \int_{\theta^i}^1 w d\theta - \rho^i_{n,t} + \int_0^{\theta^i} \left[p u^S_n + (1-p)\, u^B_n\right] d\theta + \int_{\theta^i}^1 u^B_n d\theta,$$

where $\mathbf{1}_{i=h}$ is an indicator function equal to one if the parent is of type h and zero otherwise. This expected utility function has two parts. The first part (i.e., the first two terms on the right-hand side of U^i_n) is the expected utility of individual i, who does not know his or her θ^i, in terms of criminal behavior. The second part (i.e., the last two terms on the right-hand side of U^i_n) of this utility function corresponds to the expected utility from being a parent without knowing the type of family. In the Online Appendix C, we determine $\rho^i_{n,t}$, the bid rent for a parent of type $i = h, d$ residing in neighborhood $n = 1, 2$ (see equation (C11)).

4.2 Equilibrium

To obtain the urban equilibria, we need to know who is eager to bid more for land in a particular neighborhood. Following the literature (Fujita, 1989; Zenou, 2009), the urban equilibrium is defined as follows.

Definition 2 *At any date t and given Q_t, the urban configuration, characterized by $\rho^{i*}_{1,t}$, $q^*_{1,t}$, $\tau^{S*}_{1,t}$, $\tau^{B*}_{1,t}$, $\tau^{S*}_{2,t}$, $\tau^{B*}_{2,t}$, is an equilibrium if no one wants to move and change their location choice. The highest bidders for neighborhood 1 are individuals with trait h.*

residential location.

In any urban equilibrium, both neighborhoods need to be equally attractive to each type$-i$ parent, for $i = h, d$. In the Online Appendix C, we determine the bid rent differential, $\Delta\rho \equiv \rho^h - \rho^d$ between honest and dishonest families, which is given by (C13). We then show that two different types of urban equilibria may exist.

Definition 3

(*i*) *An urban equilibrium is* ***segregated*** *at time t if all "honest" families reside in neighborhood 1 or all "dishonest" families reside in neighborhood 2, i.e.,* $q_{1,t} = Q_t$ *and* $q_{2,t} = 0$ *if* $Q_t < 1$ *and* $q_{1,t} = 1$ *and* $q_{2,t} = Q_t - 1$ *if* $Q_t > 1$.

(*ii*) *An urban equilibrium is* ***integrated*** *at time t if half of "honest" families reside in neighborhood 1 and the other half in neighborhood 2, i.e.,* $q_{1,t} = q_{2,t} = Q_t/2$.

We have the following result:

Proposition 3 *When* $Q_t < 1$*, the unique stable urban equilibrium is segregated. When* $Q_t > 1$*, the unique stable urban equilibrium is integrated.*

The agents who have higher incentives to live in neighborhood 1 are those willing to bid more for land to reside in a location with a higher proportion of honest agents. However, individuals who exert a higher socialization effort (honest agents) do not always benefit from living with a higher proportion of honest agents. This depends on whether parents' socialization effort (vertical transmission) and peer effects (oblique or horizontal transmission) are *complements* (which is true when $Q_t < 1$) or *substitutes* (when $Q_t > 1$). Indeed, we saw in Section 3.2 (see equation (8)) that when the proportion of honest individuals is low (high), parents have more (less) incentive to socialize their children to the honesty trait the more widely dominant this trait is in the population. Therefore, when $Q_t < 1$ (few honest individuals around), any increase in the proportion of honest agents benefits more parents who exert a higher socialization effort. As a result, honest parents have a strong incentive to segregate themselves in neighborhood 1 by bidding away parents with the dishonesty trait who end up in neighborhood 2. This leads to the fact that the unique stable urban equilibrium is *segregated*. When $Q > 1$, any increase in the proportion of honest agents benefits more those parents who exert a lower socialization effort (i.e. dishonest parents). As a result, to reside with honest families, dishonest parents are willing to bid more for land than honest parents and thus the unique stable urban equilibrium is *integrated.*

The following proposition then characterizes the long run structure of the culture of honesty and crime rates in the city.

Proposition 4 *Suppose that the probability of being apprehended, p, is such that $p \in [0, \widehat{p}_1[\cup]\widehat{p}_2, \overline{p}_2]$. Define $q_{max} = \max\{\underline{q}, \frac{1}{2}\}$.*

(i) If $Q_0 \in \left[0, \underline{q}\right[$, then in the long run, there is no spatial pattern of social disorganization, no honesty trait in either neighborhood, i.e., $q_1^* = q_2^* = 0$, and the crime rate is high such that $C_1 = C_2 = \theta^d$.

(ii) If $Q_0 \in \left[\underline{q}, 2q_{max}\right[$, then in the long run, social disorganization is spatially differentiated, there is segregation in terms of the honesty trait, and crime is much higher in neighborhood 2, i.e., $q_1^* = \overline{q}$, $q_2^* = 0$, and

$$C_1 = \overline{q}\theta^h + (1 - \overline{q})\theta^d < C_2 = \theta^d.$$

(iii) If $Q_0 \in [2q_{max}, 2]$, then in the long run, there is no spatial pattern of social disorganization, a high frequency of the honesty trait, and a low crime rate in both neighborhoods, i.e., $q_1^* = q_2^* = \overline{q}$, and

$$C_1 = C_2 = \overline{q}\theta^h + (1 - \overline{q})\theta^d.$$

In this proposition, we focus on the case when the probability of being apprehended, p, takes intermediate values, i.e., $p \in [0, \widehat{p}_1[\cup]\widehat{p}_2, \overline{p}_2]$, which corresponds to part (ii) of Proposition 1.[26] When there is no location choice, there are two stable steady-state equilibria depending on the initial conditions. If q_0 is low (i.e., $q_0 < \underline{q}$), then $q^* = 0$, while if it is sufficiently high (i.e., $q_0 > \underline{q}$), then $q^* = \overline{q} < 1$.

When we introduce location choices, the results differ because now crime becomes spatially differentiated. In Proposition 4, in case (i), we show that when Q_0, the initial proportion of honest families, is very low, meaning that the initial proportion of honest individuals is low in both neighborhoods (i.e., $q_{1,0} < \underline{q}$ and $q_{2,0} < \underline{q}$), then there is a unique long-term

[26]The other case in Proposition 1 is uninteresting since the unique stable steady-state equilibrium is such that $q^* = 0$ (i.e., there are no honest families in the long run).

equilibrium for which $q_1^* = q_2^* = 0$. Similar but opposite results are obtained in case (*iii*) of Proposition 4. Indeed, when the initial proportion of honest individuals in both neighborhoods is sufficiently high (i.e., $q_{1,0} > \underline{q}$ and $q_{2,0} > \underline{q}$), then the steady-state proportion of honest individuals is $q_1^* = q_2^* = \overline{q}$, and the crime rate is low in both neighborhoods. In other words, when the culture of honesty is sufficiently widespread at the beginning, there is no spatial disorganization.

Segregation emerges only when $Q_0 \in \left[\underline{q}, 2q_{max}\right[$ (case (*ii*) of Proposition 4). Consider first the case when the tipping point $\underline{q} < \frac{1}{2}$, meaning that $Q_0 \in \left[\underline{q}, 1\right[$. Since $Q_0 < 1$, we have shown in Proposition 3 that the unique stable urban equilibrium is segregated. In that case, in neighborhood 1, there will be $q_1^* = \overline{q}$ honest individuals in the steady state, while, in neighborhood 2, nobody will be honest (i.e., $q_2^* = 0$). Consider now the case when $\underline{q} > \frac{1}{2}$ and thus $Q_0 \in \left[\underline{q}, 2\underline{q}\right[$. This means that Q_0 can be greater than 1, so that we may start with spatial integration where the proportion of honest families is the same in each neighborhood. However, this initial spatial integration allocation is not stable because honest (biparental) families decrease their socialization efforts (parents' socialization efforts and peer effects are substitutes) up to the point where, in some period t, $Q_t < 1$ and we end up with a segregated equilibrium.

Proposition 4 thus shows that when Q_0 takes intermediate values, there is a unique stable steady-state equilibrium at which, in neighborhood 1, land rent is high, most families are biparental, and the crime rate is very low, while in neighborhood 2, land rent is low, families are broken (the father is absent), and the crime rate is very high. This result highlights the fact that *spatial segregation strengthens social disorganization* and vice versa, explaining why there is a collapse of the community in some neighborhoods, which leads to high-crime rates. This finding explains why criminal youths from broken families tend to live in high-crime neighborhoods.

More generally, Proposition 4 provides a microfoundation of the equilibrium urban models a la Benabou (1993) in which segregation occurs because of positive externalities. In Benabou (1993), the higher is the number of educated individuals in a neighborhood, the lower is the cost of acquiring education. As a result, segregation occurs between high- and low-educated families because of these positive externalities in terms of educational costs. Here, we take a dynamic perspective and provide a microfoundation of these positive externalities by showing that segregation only emerges and persists when the initial proportion of families with the honesty trait is neither too low nor too high.

In particular, urban ghettos emerge in our model because when, initially, few people have the honesty trait, many adults become criminals, which increases the proportion of fatherless families in the neighborhood. The high proportion of broken families together with the criminal behavior of other children lead to the demise of the local community and eventually to urban ghettos with high-crime rates.

4.3 Efficiency versus equity

Let us now turn to normative assessments and consider efficiency and equity. Again, the objective of a social planner is to minimize long-run crime.

Definition 4 (Efficiency with reallocation policies) *For a given Q_0, the first-best (or efficient) allocation of individuals into the two neighborhoods is the one that minimizes total long-run crime.*

Definition 5 (Equity) *For a given Q_0, an equitable allocation of agents is the one that minimizes the difference in crime between the two neighborhoods.*

Proposition B8 in the Online Appendix B.5 studies the efficiency of the market solution and determines whether segregation is efficient and/or it conflicts with equity. We show that, when $Q_0 \in \left[0, \underline{q}\right[$ (resp. $Q_0 \in [2q_{max}, 2]$), the market solution does not lead to spatial segregation and there is high crime and no honesty culture everywhere (resp. low crime and the same degree of honesty in the two neighborhoods). A relocation policy clearly does not improve the long-run crime rate in either case. The market solution is efficient and obviously matches equity as there is no spatial differentiation.

When $Q_0 \in [\underline{q}, 2\underline{q}[$, the market solution leads to segregation, which is efficient. Indeed, in neighborhood 1, one gets $q_{1,0} > \underline{q}$, while in neighborhood 2, $q_{2,0} < \underline{q}$, where $\underline{q}$ is the tipping point. As a result, in neighborhood 1, the steady-state equilibrium of honest families is equal to $q_1^* = \overline{q}$, while in neighborhood 2, we have $q_2^* = 0$. If the planner wanted to integrate the two populations, he or she would not succeed. In neighborhood 2, he or she cannot affect the steady-state equilibrium for which $q_2^* = 0$, while in neighborhood 1, he or she will reduce the proportion of honest people below the tipping point, which will lead to $q_1^* = 0$. This clearly provides a worse global long-term crime outcome than the segregation outcome. As a result, segregation is efficient but efficiency conflicts with equity because each neighborhood can only encompass a mass 1 of families and $Q_0 < min\left\{2\underline{q}, 1\right\}$.

Finally, when $Q_0 \in \left[2\underline{q}, 2q_{max}\right[$ is equivalent to $Q_0 \in \left[2\underline{q}, 1\right[$,[27] the unique steady-state equilibrium is segregation since $Q_0 < 1$. This is not efficient because the planner could integrate the two neighborhoods by having $q_{1,0} > \underline{q}$ and $q_{2,0} > \underline{q}$, which would lead to $q_1^* = q_2^* = \overline{q}$. Obviously, integration is equitable in this case and efficiency matches with equity.

Now, consider a planner who cannot reallocate people between the two locations but who rather implements a policy that affects the model's parameters. In particular, assume that the planner can reduce the socialization cost of single-mother families through a non-costly subsidy δ.[28]

Proposition B9 in the Online Appendix B.5 studies such a policy. We show that, when Q_0 is very low, the culture of honesty disappears in the long run and everybody is dishonest. Hence, both the subsidy policy and the reallocation policy are inefficient since there is no impact on long-run honesty. When the reallocation policy is not efficient, segregation is efficient because reallocating families would only lead to a worse outcome. On the contrary, a subsidy policy δ has a positive impact on honesty because it increases the proportion of honest agents in neighborhood 1 from $\overline{q}$ to $\overline{q}^\delta$.

Conversely, when $Q_0 \in \left[2\underline{q}^\delta, 2q_{max}^\delta\right[$,[29] the reallocation policy is more efficient than the subsidy policy. Indeed, consider that $\underline{q}^\delta < \frac{1}{2}$ so that $q_{max}^\delta = \frac{1}{2}$. Thus, $Q_0 \in \left[2\underline{q}^\delta, 1\right[$.[30] In that case, the long-run equilibrium is such that there is segregation where all honest people live in neighborhood 1 ($q_1^* = \overline{q}$ and $q_1^* = 0$). The subsidy policy would keep segregation but increase the proportion of honest families in neighborhood 1 from $\overline{q}$ to $\overline{q}^\delta$. On the contrary, since $Q_0 > 2\underline{q}$, the reallocation policy would move people around so that we end up with spatial integration where, in each neighborhood, there is a proportion $\overline{q}$ of honest families. In terms of the reduction in total crime, this is better than a subsidy policy since $2\overline{q} > \overline{q}^\delta$ (because $2\overline{q} > 1$). As a result, the reallocation policy is more efficient in reducing total crime

[27] This case is only relevant when $q_{max} = \frac{1}{2} > \underline{q}$.

[28] Suppose that the maximal subsidy is such that $\Delta^S = \Delta^B$, which means that, at the maximal subsidy, the planner completely removes the negative effect of family disruption on crime. We could equivalently consider an increase in w, increase in β, or increase in σ. Nevertheless, we do not consider changes in p since p is restricted to a given interval.

[29] $\underline{q}^\delta$ and q_{max}^δ are the equivalent to $\underline{q}$ and q_{max} when the single family socialization subsidy is δ.

[30] The other case when $\underline{q}^\delta > \frac{1}{2}$ is impossible since it leads to $q_{max}^\delta = \underline{q}^\delta$, which implies that $Q_0 \in \left[2\underline{q}^\delta, 2\underline{q}^\delta\right[$.

than the subsidy policy.

4.4 Cultural disorganization, neighborhood and education policies

Before concluding, we would like to understand the effectiveness of spatial education policies in reducing crime. In the popular press, it has often been argued that we should increase education expenditures (e.g. improve school quality, hire better teachers, etc.) mostly in poor neighborhoods so that kids have more work opportunities and, thus, will be less likely to revert to crime. We examine this issue in this section by looking at our model when the crime rate differs between neighborhoods 1 and 2 (i.e., case (ii) of Proposition 4).

As in Section 3.4, suppose that the wage w is a function of a public education investment g, so that $w = w(g)$, with $w'(g) > 0$, $w''(g) < 0$, $w(0) = 0$. Consider a public education policy, which consists in increasing g to $g' > g$.

Let $\underline{q}^g$ (resp. $\underline{q}^{g'}$) and $\overline{q}^g$ (resp. $\overline{q}^{g'}$) be the steady-state equilibria of Proposition 1 when the education policy is not implemented (resp. implemented). From Proposition B1, we know that, because public education increases the wage rate w, we have $\underline{q}^{g'} < \underline{q}^g$ and $\overline{q}^{g'} > \overline{q}^g$

Proposition 5 *Suppose that* $Q_0 \in \left[\underline{q}^g, 2q^g_{max}\right[$*, where* $q^g_{max} = \max\{\underline{q}^g, \frac{1}{2}\}$*. Then,*

(i) if the initial mass of honest agents is low, i.e., $Q_0 < \max\{2\underline{q}^{g'}, 1\}$*, the education policy is more effective in reducing crime in neighborhhood 1. A common education policy for both neighborhoods strengthens the spatial disparity in crime;*

(ii) if the initial mass of honest agents is high, i.e., $Q_0 > \max\{2\underline{q}^{g'}, 1\}$*, the education policy is more effective in reducing crime in neighborhood 2. A common education policy for both neighborhoods breaks the spatial disparity in crime.*

Proposition 5 reveals that education policies may be more effective in some neighborhoods than in others and not necessary in the poor crime-ridden neighborhood. We show that the neighborhood that should be targeted by the policy crucially depends on the culture of honesty and the size of the investment in the education policy. Indeed, when the initial culture of honesty is leading to spatial segmentation, i.e., $Q_0 \in \left[\underline{q}^g, 2q^g_{max}\right[$ and the education policy is not strong enough to induce spatial integration, i.e., $Q_0 < \max\{2\underline{q}^{g'}, 1\}$, then the policy is *less* effective in the high-crime neighborhood, which is populated by only dishonest agents. In such a neighborhood, more education reduces the incentives of dishonest agents to commit crime. Conversely, in the other neighborhood, which has both honest and dishonest agents, the education policy not only affects the incentives of exerting crime but *also*, by reducing short-run crime, lowers the fraction of single-parent families and, thus, positively impacts on the transmission of the culture of honesty.

In the socially disorganized neighborhood (neighborhood 2), education policies affect crime only through the deterrence effect while, in the low-crime neighborhood, education policies also reduces crime by promoting the culture of honesty. As a result, such policies are more effective in neighborhood 1 in reducing crime.

By contrast, when education investment policies are strong enough and the initial culture of honesty is not too weak, i.e., $Q_0 > \max\{2\underline{q}^{g'}, 1\}$, the education policy is more effective in reducing crime in the socially disorganized neighborhood because it reduces crime and thus family disruption, which positively impacts on the transmission of honesty in the population. If the initial mass of honest agents is not too small, the positive effect of education on the dynamics of honesty is then high enough to induce a rise in the culture of honesty. This, in turn, modifies the incentives to reside in each neighborhood because it induces dishonest agents to mix with honest individuals. Integration, in turn, promotes the rise of the culture of honesty in neighborhood 2, which leads to a substantial reduction in the crime rate in that area. Hence, an education policy prevents the cultural disorganization in neighborhood 2. This qualitative change, which occurs only in neighborhood 2, strongly contributes to the reduction of crime in that neighborhood, which makes education policies in that area very effective.

5 Concluding remarks and policy implications

Policymakers have been recognizing for a long time the connection between the breakdown of American families and various social problems. The unfolding debate over welfare reform, for instance, has been shaped by the wide acceptance in recent years that children born into single-parent families are much more likely than children of intact families to fall into poverty and welfare dependence themselves in later years. These children thus face a daunting array of problems. Without an understanding of the root causes of criminal behavior (i.e., how criminals are formed), it is difficult to understand why whole sectors of society, particularly in urban areas, are being torn apart by crime.

We develop a dynamic model that provides a mechanism linking higher incarceration rates to a higher number of single-parent families, less cultural control on crime, and thus a higher crime rate. Then, we study whether parental investments or peer effects are more relevant. We also investigate whether the neighborhood is the more relevant unit of socialization and whether high house prices in lower-crime neighborhood prevent families from moving to these more desirable neighborhoods.

Our study indicates that policies aimed at reducing crime by focusing on keeping families intact may be better served to improve parenting practices, especially attachment, monitoring, and involvement. Parenting education programs encourage parents to take notice what their children are doing, praise good behavior, state house rules clearly, and make rewards and punishments contingent on children's behavior. A number of programs have demonstrated success in reducing children's antisocial behavior.[31]

Our model also suggests that, at least for some parameter values, public spending and private investment must be concentrated in the most impoverished areas. Money must be mainly spent on programs physically located in underclass neighborhoods, run by people with ties to the neighborhoods they intend to serve. This policy has the effect of targeting programs for the underclass, while also strengthening minority agencies or creating new agencies within very poor neighborhoods. These agencies not only provide services, but can also provide jobs for neighborhood residents, which may strengthen the structure of biparental families in these neighborhoods. For example, *place-based policies* could be a good way of improving poor neighborhoods.[32] Implementing these types of policies creates jobs and thus facilitates the flow of job information in depressed neighborhoods. As employment opportunities increase and better-funded local agencies become centers for social action, the pressure on working- and middle-class residents to flee should decrease.

Another place-based policy that directly targets crime is the so-called *hot spots policy*. The general idea is that crimes are clustered at specific locations in the city, i.e., hot spots, which are small geographic units. Thus, police forces will try to use the location of predicted crime to target their efforts. Of course, this may also generate arrests caused by police

[31] For an overview and evaluation of existing programs focusing on family support, see Farrington and Welsh (2005).

[32] For an overview of place-based policies, see Neumark and Simpson (2015).

presence, hence the direction of causality may not be fully clear (see, e.g. Wu and Lum, 2017). In addition, there are displacement effects of policing efforts (Weisburd et al., 2006) so the net effect of this policy is unclear (Lazzati and Menichini, 2016).[33] We could analyze this policy in the spatial model of Section 4, which would imply that the probability of apprehension p would be neighborhood specific.

Furthermore, in this paper (Section 3.4), education has been mainly assumed to increase the wage in the legal sector since more education implies higher productivity, and thus higher wage. One possible extension is that the government optimally chooses public education under some budget constraint (as, for example, in Verdier and Zenou, 2018) to influence the transmission of honesty[34] and, thus, the long-term crime level in the economy. We leave these two interesting extensions for future research.

References

[1] Anselin, L., Cohen, J., Cook, D., Gorr, W., and G. Tita (2000), "Spatial analyses of crime," *Criminal Justice* 4, 213–262.

[2] Ballester, C., Calvó-Armengol, A. and Y. Zenou (2010), "Delinquent networks," *Journal of the European Economic Association* 8(1), 34–61.

[3] Barber, B.L. and J.S. Eccles (1992), "Long-term influence of divorce and single-parenting on adolescent family- and work-related values, behaviors, and aspirations," *Psychological Bulletin* 222, 108–126.

[4] Bayer, P., Hjalmarsson, R. and D. Pozen (2009), "Building criminal capital behind bars: Peer effects in juvenile corrections," *Quarterly Journal of Economics* 124(1), 105–147.

[5] Bearse, P.M. (1996), "Community, youth crime and deterrence: A quantitative theoric analysis," Unpublished manuscript, University of North Carolina at Greensboro.

[6] Becker, G.S. (1968) "Crime and punishment: An economic approach," *Journal of Political Economy* 76(2), 169-217.

[33] For overviews on hot spots policies, see Braga et al. (2012) and Weisburd and Telep (2014).
[34] We partly address this issue in Section 4.4.

[7] Benabou, R. (1993), "Workings of a city: Location, education, and production," *Quarterly Journal of Economics* 108, 619-652.

[8] Bethencourt, C. and L. Kunze (2014), "On the intergenerational nature of criminal behavior," Munich Personal RePEc Archive (MPRA) Working Paper No. 58344.

[9] Bezin, E., Verdier, T. and Y. Zenou (2020), "Crime, broken families, and punishment," ssrn.com/abstract=3200773.

[10] Bisin, A. and T. Verdier (2000), "Beyond the melting pot: Cultural transmission, marriage, and the evolution of ethnic and religious traits," *Quarterly Journal of Economics* 115, 955-988.

[11] Bisin, A. and T. Verdier (2001), "The economics of cultural transmission and the dynamics of preferences," *Journal of Economic Theory* 97, 298-319.

[12] Bisin, A., Patacchini, E., Verdier, T. and Y. Zenou (2011), "Formation and persistence of oppositional identities," *European Economic Review* 55, 1046–1071.

[13] Boswell, G. and P. Wedge (2002), *Imprisoned Fathers and their Children*, London: Jessica Kingsley Publishers.

[14] Braga, A.A., Papachristos, A.V. and D.M. Hureau (2012), "Hot spots policing effects on crime," *Campbell Systematic Reviews*, Article 8.

[15] Braman, D. (2004), *Doing Time on the Outside: Incarceration and Family Life in Urban America*, Ann Arbor: University of Michigan Press.

[16] Brantingham, P.J. (2016), "Crime diversity," *Criminology* 54, 553–586.

[17] Büchel, B., Hellmann, T. and M.M. Pichler (2014), "The dynamics of continuous cultural traits in social networks," *Journal of Economic Theory* 154, 274–309.

[18] Bureau of the Census (2016), "America's families and living arrangements," Washington, DC: U.S. Department of Commerce.

[19] Calvó-Armengol, A. and Y. Zenou (2004), "Social networks and crime decisions. The role of social structure in facilitating delinquent behavior," *International Economic Review* 45(3), 939-958.

[20] Chetty, R., Friedman, J., Hendren, N., Jones, M.R. and S.R. Porter (2018), "The opportunity Atlas: Mapping the childhood roots of social mobility," NBER Working Paper No. 25147.

[21] Chetty, R., Hendren, N., Jones, M.R. and S.R. Porter (2020), "Race and economic opportunity in the United States: An intergenerational perspective," *Quarterly Journal of Economics* 135(2), 711–783.

[22] Cheung, M.-W. and J. Wu (2018), "On the probabilistic transmission of continuous cultural traits," *Journal of Economic Theory* 174, 300–323.

[23] Cobb-Clark, D.A. and E. Tekin (2014), "Fathers and youths' delinquent behavior," *Review of Economics of the Household* 12, 327-358.

[24] Comfort, M. (2008), *Doing Time Together: Love and Family in the Shadow of the Prison*, Chicago, University of Chicago Press.

[25] Cortés, D., Friebel, G. and D. Maldonado (2020), "Crime and education in a model of information transmission," *Annals of Public and Cooperative Economics* 91(1), 71–93.

[26] Damm, A.P. and C. Dustmann (2014), "Does growing up in a high crime neighborhood affect youth criminal behavior?" *American Economic Review* 104(6), 1806-1832.

[27] Decreuse, B., Mongrain, S. and T. van Ypersele (2015), "Property crime and private protection allocation within cities: Theory and evidence," CEPR Discussion Paper No. 10707.

[28] Della Lena, S. and P. Dindo (2019), "On the evolution of norms in strategic environments," Ca' Foscari University of Venice, Working Paper No. 16.

[29] Deng, L. and Y. Sun (2017), "Criminal network formation and optimal detection policy: The role of cascade of detection," *Journal of Economic Behavior and Organization* 141, 43–63.

[30] Dobbie, W., Grönqvist, H., Niknami, S., Palme, M. and M. Priks (2018), "The intergenerational effects of parental incarceration," NBER Working Paper No. 24186.

[31] Dornbusch, S.M., Cardsmith, J., Bushwall, S., Ritter, P., Leiderman, H., Hastorf, A. and R. Gross (1985), "Single parents, extended households, and control of adolescents," *Child Development* 56, 326-341.

[32] Farrington, D.P. and B.C. Welsh (2005), "Randomized experiments in criminology: What have we learned in the last two decades?" *Journal of Experimental Criminology* 1, 9-38.

[33] Fella, G. and G. Gallipoli (2014), "Education and crime over the life cycle," *Review of Economic Studies* 81, 1484–1517.

[34] Florsheim, P., Tolan, P. and D. Gorman-Smith (1998), "Family relationships, parenting practices, the availability of male family members, and the behavior of inner-city boys in single-mother and two-parent families," *Child Development* 69, 1437-1447.

[35] Freeman, R.B. (1996), "Why do so many young American men commit crimes and what might we do about it?" *Journal of Economic Perspectives* 10, 24-42.

[36] Fujita, M. (1989), *Urban Economic Theory*, Cambridge: Cambridge University Press.

[37] Gaigné, C. and Y. Zenou (2015), "Agglomeration, city size and crime," *European Economic Review* 80, 62–82.

[38] Galiani, S., Lopez Cruz, I.G. and G. Torrens (2018), "Stirring up a hornets' nest: Geographic distribution of crime," *Journal of Economic Behavior and Organization* 152, 17–35.

[39] Glaeser, E.L. and B.I. Sacerdote (1999), "Why is there more crime in cities," *Journal of Political Economy* 107, S225-S258.

[40] Glaeser, E.L., Sacerdote, B.I. and J.A. Scheinkman (1996), "Crime and social interactions," *Quarterly Journal of Economics* 111(2), 507-548.

[41] Goffman, A. (2014), *On the Run. Fugitive Life in an American City*, Chicago: The University of Chicago Press.

[42] Ghost, K. (2018), "Families and mass incarceration," In: *Criminal Justice Involvement of Families in Child Welfare*, Spring, Center of Advanced Studies in Child Welfare, School of Social Work, University of Minnesota.

[43] Harlow, C.W. (2003), "Education and correctional populations," U.S. Department of Justice, Bureau of Justice. Statistics

[44] Hauk, E. and M. Saéz-Martí (2002), "On the cultural transmission of corruption," *Journal of Economic Theory* 107, 311–335.

[45] Katz A. (2002), *Parenting Under Pressure: Prison,* Young Voice Publisher.

[46] Kellam, S.G., Ensminger, M.E. and R.J. Turner (1977), "Family structure and the mental health of children," *Archives of General Psychiatry* 34, 1012-1022.

[47] Kleiman, M.A. (2009), *When Brute Force Fails. How to Have Less Crime and Less Punishment*, Princeton: Princeton University Press.

[48] Kling, J.R., Ludwig, J. and L.F. Katz (2005), "Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment," *Quarterly Journal of Economics* 120(1), 87-130.

[49] Kneebone, E. and S. Raphael (2011), "City and suburban crime trends in Metropolitan America," The Brooking Institution, Washington D.C.

[50] Kubrin C.E. and J.C. Wo (2016), "Social disorganization theory's greatest challenge: Linking structural characteristics to crime in socially disorganized communities," In: A.R. Piquero (Ed.), *The Handbook of Criminological Theory*, New York: John Wiley.

[51] Lazzati, N. and A. Menichini (2016), "Hot spot policing: A study of place-based strategies for crime prevention," *Southern Economic Journal* 82(3), 893–913.

[52] Lee, L.-F. , Liu, X., Patacchini, E. and Y. Zenou (2021), "Who is the key player? A network analysis of juvenile delinquency," *Journal of Business and Economic Statistics*, forthcoming.

[53] Lochner, L. (2004), "Education, work, and crime: A human capital approach," *International Economic Review* 45, 811–843.

[54] Lochner, L. (2011), "Nonproduction benefits of education: Crime, health, and good citizenship," In: E.A. Hanushek, S.J. Machin and L. Woessmann (Eds.), *Handbook of the Economics of Education, Vol. 4*, Amsterdam: Elsevier, pp. 183–282.

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

[56] Ludwig, J., Duncan, G.J. and P. Hirschfield (2001), "Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment," *Quarterly Journal of Economics* 116(2), 655-679.

[57] Lynch, J.P. and W.J. Sabol (2004), "Assessing the effects of mass incarceration on informal social control in communities," *Criminology and Public Policy* 3, 267-294.

[58] Machin, S., Marie, O. and S. Vujic (2011), "The crime reducing effect of education," *Economic Journal* 121, 463–484.

[59] Murray, J. and D.P. Farrington (2005), "Parental imprisonment: Effects on boys' antisocial behaviour and delinquency through the life-course," *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 46, 1269–1278.

[60] Neumark, D. and H. Simpson (2015), "Place-based policies," In: G. Duranton, V. Henderson and W. Strange (Eds.), *Handbook of Regional and Urban Economics*, Vol. 5B, Amsterdam: Elsevier Publisher, pp. 1197-1288.

[61] O'Flaherty, B. and R. Sethi (2015), "Urban crime," In: G. Duranton, V. Henderson and W. Strange (Eds.), *Handbook of Regional and Urban Economics, Vol. 5B*, Amsterdam: Elsevier Publisher, pp. 1519-1621.

[62] Panebianco, F. (2014), "Socialization networks and the transmission of interethnic attitudes," *Journal of Economic Theory* 150, 583–610.

[63] Porter, J.R., Capellan, J. and S. Chintakrindi (2016), "Social disorganization theory," In W.G. Jennings (Ed.), *The Encyclopedia of Crime and Punishment,* London: Wiley-Blackwell.

[64] Prison Policy Initiative (2020), "Mass Incarceration: The Whole Pie 2020," https://www.prisonpolicy.org/reports/pie2020.html.

[65] Prummer, A. and J.-P. Siedlarek (2017), "Community leaders and the preservation of cultural traits," *Journal of Economic Theory* 168, 143–176.

[66] Saéz-Martí, M. and A. Sjögren (2008), "Peers and culture," *Scandinavian Journal of Economics* 100(1), 73–92.

[67] Shaw, C.R. and H.D. McKay (1942), *Juvenile Delinquency and Urban Areas*, Chicago: University of Chicago Press.

[68] Stern, M., Northman, J.E. and M.R. Van Slyck (1984), "Father absence and adolescent problem behavior: Alcohol consumption, drug use and sexual activity," *Adolescence* 19, 301-312.

[69] Turner, R.A., Irwin, C.E. and S.G. Millstein (1991), "Family structure, family processes, and experimenting with substances during adolescence," *Journal of Research on Adolescence* 1, 93-106.

[70] Verdier, T. and Y. Zenou (2004), "Racial beliefs, location and the causes of crime," *International Economic Review* 45, 731-760.

[71] Verdier, T. and Y. Zenou (2018), "Cultural leader and the dynamics of assimilation," *Journal of Economic Theory* 175, 374–414.

[72] Weisburd, D., Wycko , L. A., Ready, J., Eck, J. E., Hinkle, J.C. and F. Gajewski (2006), "Does crime just move around the corner? A controlled study of spatial displacement and diffusion of crime control benefits," *Criminology* 44, 549–592.

[73] Weisburd, D. and C. Telep (2014), "Hot spots policing: What we know and what we need to know, "*Journal of Contemporary Criminal Justice* 30, 200?220.

[74] Wildeman, C. (2020), "The intergenerational transmission of criminal justice contact," *Annual Review of Criminology* 3, 217–244.

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

[76] Wilson, W.J. (1987), *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*, Chicago: University of Chicago Press.

[77] Wu, X. and C. Lum (2017), “Measuring the spatial and temporal patterns of police proactivity, ” *Journal of Quantitative Criminology* 33, 915–934.

[78] Zenou, Y. (2003), “The spatial aspects of crime,” *Journal of the European Economic Association* 1, 459-467.

[79] Zenou, Y. (2009), *Urban Labor Economics*, Cambridge: Cambridge University Press.

Mathematical Appendix

Proof of Proposition 1:

Let us denote by $f(q_t)$, the function defined on $[0,1] \to [0,1]$ and given by

$$f(q_t) = q_t(1-q_t)\left\{4q_t\left(1-q_t\right)\left[\Delta^B - pC(q_t)\left(\Delta^B - \Delta^S\right)\right] - 1\right\}. \tag{13}$$

The stationary equilibria of the economy for which $q_t = q_{t+1} = q$ are such that $f(q) = 0$.

First, we have $f(0) = f(1) = 0$.

Second, if there exists some $q_t \neq 0, 1$, then solving $f(q_t) = 0$ must lead to

$$4q_t\left(1-q_t\right)\left[\Delta^B - pC(q_t)\left(\Delta^B - \Delta^S\right)\right] - 1 = 0,$$

which is equivalent to

$$q_t\left(1-q_t\right)\left[\Delta^B - pC(q_t)\left(\Delta^B - \Delta^S\right)\right] = \frac{1}{4}.$$

Let us denote by $h(q_t)$ the function defined on $[0,1] \to [0,1]$ and given by

$$h(q_t) = q_t(1-q_t)\left[\Delta^B - pC(q_t)\left(\Delta^B - \Delta^S\right)\right]. \tag{14}$$

Step 1. We show that the function $h(q_t)$ admits a unique maximum on $[0,1]$.

We have

$$h'(q_t) = (1-2q_t)\left[\Delta^B - pC(q_t)\left(\Delta^B - \Delta^S\right)\right] - q_t(1-q_t)pC'(q_t)\left(\Delta^B - \Delta^S\right).$$

By using (4), we have

$$h'(q_t) = (1-2q_t)\left[\Delta^B - pC(q_t)\left(\Delta^B - \Delta^S\right)\right] + q_t(1-q_t)pK\left(\Delta^B - \Delta^S\right).$$

The function $h'(q_t)$ is a polynomial of order two, which is concave. We have

$$h'(0) = \Delta^B - pC(0)\left(\Delta^B - \Delta^S\right) > 0$$

and

$$h'(1) = -\left[\Delta^B - pC(1)\left(\Delta^B - \Delta^S\right)\right] < 0,$$

so that there exists a unique

$$q_m = \frac{-\left[\Delta^B - pK\left(\Delta^B - \Delta^S\right)\right] + \sqrt{D}}{3K\left(\Delta^B - \Delta^S\right)},$$

with

$$D = \left[\Delta^B - pK\left(\Delta^B - \Delta^S\right)\right]^2 + 3K\left(\Delta^B - \Delta^S\right)\left[\Delta^B - p\theta^d\left(\Delta^B - \Delta^S\right)\right]$$

such that $h'(q_m) = 0$. This implies that the function $h\left(q_t\right)$ reaches a global maximum at $q = q_m$. We deduce that there exists $\underline{q} \leq \overline{q} \neq 0, 1$ such that $f(\underline{q}) = f(\overline{q}) = 0$ if and only if $h(q_m) \geq \frac{1}{4}$.

Figures 5 and 6 respectively depict the function h for $h(q_m) > \frac{1}{4}$ and $h(q_m) < \frac{1}{4}$.

Step 2. We show that there exists $\hat{p}_1$ such that for any $p \leq \hat{p}_1$, $h(q_m) \geq \frac{1}{4}$ (i.e., $h(q;p) - \frac{1}{4} = 0$ has two solutions) and for any $p^* > p > \hat{p}_1$, $h(q_m) < \frac{1}{4}$, $p^* \leq \bar{p}_1$ (i.e., $h(q;p) - \frac{1}{4} = 0$ has no solution).

Let us define

$$u(p) \equiv h(q_m(p), p).$$

We have $h(q_m) \geq \frac{1}{4}$ if and only if

$$u(p) \geq \frac{1}{4}.$$

Let us differentiate u with respect to p. We obtain

$$\begin{aligned}\frac{du}{dp} &= \frac{\partial h\left(q_m(p);p\right)}{\partial q_m}\frac{dq_m}{dp} + \frac{\partial h\left(q_m(p);p\right)}{\partial p} \\ &= q_m(p)\left[1 - q_m(p)\right]\left(\Delta^B - \Delta^S\right)\left[q_m(p)K - \left(\beta - 2p(\beta + \sigma) - w\right)\right].\end{aligned}$$

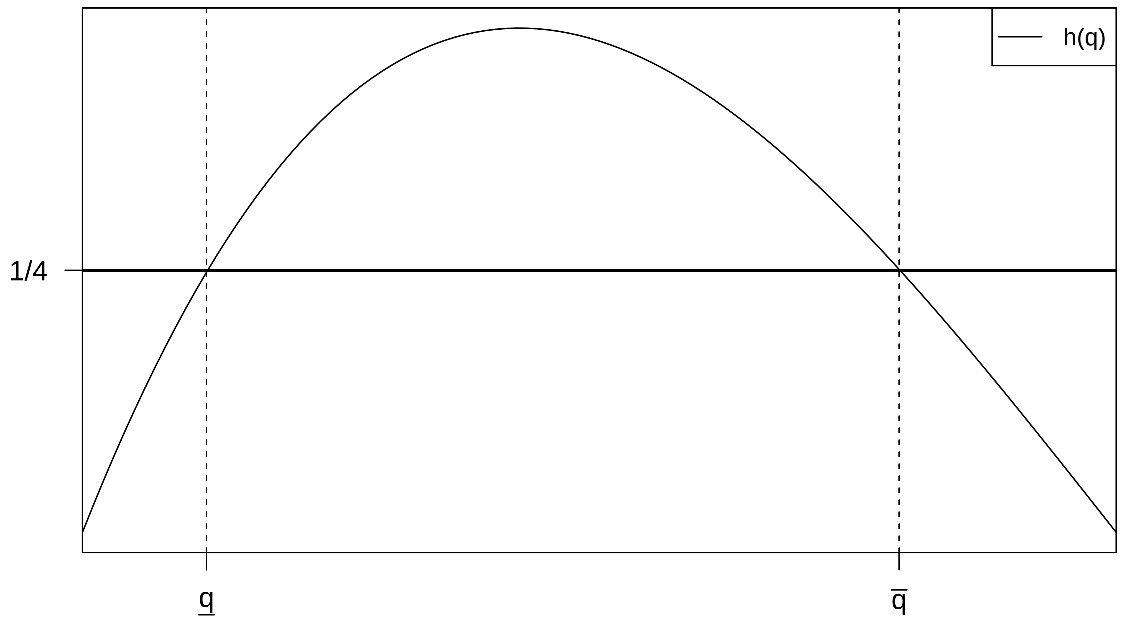

Figure 5: Case $h(q_m) > \frac{1}{4}$.

We have $\frac{du}{dp} > 0$ if and only if

$$q_m(p) > \frac{\beta - 2p(\beta+\sigma) - w}{K} \equiv \widetilde{q}.$$

Since q_m is the maximum of $h\left(.\right)$, we deduce that $q_m(p) < \widetilde{q}$ if and only if $h'(\widetilde{q}) < 0$. Let us study this condition. We have

$$\begin{aligned} h'(\widetilde{q}) &= \frac{\left(K - 2\left[\beta - 2p(\beta+\sigma) - w\right]\right)\left[\Delta^B - p^2\left(\beta+\sigma\right)\left(\Delta^B - \Delta^S\right)\right]}{K} \\ &+ \frac{p\left(\Delta^B - \Delta^S\right)\left(\beta - 2p(\beta+\sigma) - w\right)^2}{K}. \end{aligned}$$

This is a polynomial function of p with coefficients associated with the squared term p^2 equal to $-\left(\beta+\sigma\right)\left[2(\beta - w) + K\right]$, which is negative. Thus, this function $h'(\widetilde{q})$ is concave.

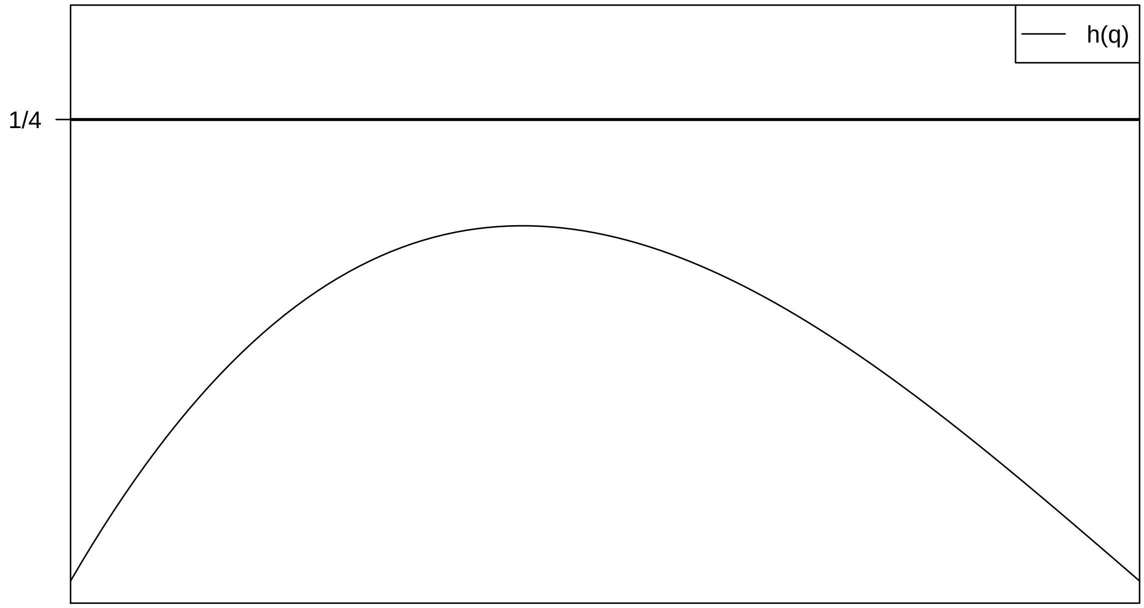

Figure 6: Case $h(q_m) < \frac{1}{4}$.

Furthermore, we know that at $p = (\beta - w - K) / [2(\beta + \sigma)]$, $\widetilde{q} = 1$; hence, $q_m(p) < \widetilde{q}$ and the polynomial is negative. Further, at $p = (\beta - w) / [2(\beta + \sigma)]$, $\widetilde{q} = 0$, so that $q_m(p) > \widetilde{q}$, which implies that the polynomial is positive. As a result, there exists a unique $\widetilde{p}$ such that for any $p \leq \widetilde{p}$, $q_m(p) \leq \widetilde{q}$, which implies that

$$\frac{dh(q_m(p); p)}{dp} \leq 0,$$

while for any $p \geq \widetilde{p}$, $q_m(p) \geq \widetilde{q}$, which implies that

$$\frac{dh(q_m(p); p)}{dp} \geq 0.$$

At $p = 0$, we have $h(\widetilde{q}(0); 0) = \frac{1}{4}\Delta^B$, which is higher than 1/4 from part (i) of Assumption 1.

Second, note that $\forall q_t$,

$$h(q_t) \leq q_t(1-q_t)\left[\Delta^B - p\left(-K+\beta-p(\beta+\sigma)-w\right)\left(\Delta^B-\Delta^S\right)\right] \equiv h^1(q_t).$$

Note that $h^1(\frac{1}{2}) \geq h^1(q_m) > h(q_m)$. Define $h^1(\frac{1}{2}) \equiv u^1(p)$. The latter inequality implies $u^1(p) > u(p)$, $\forall p \in [0, \bar{p}_1]$. The function $u^1(p)$ is a convex function of p, which reaches a minimum at $p = \frac{-K+\beta-w}{2(\beta+\sigma)} < \bar{p}_1$. Suppose that

$$u^1(\frac{-K+\beta-w}{2(\beta+\sigma)}) < \frac{1}{4},$$

which is equivalent to item (ii) of Assumption 1. We deduce that

$$u\left(\frac{-K+\beta-w}{2(\beta+\sigma)}\right) < u^1\left(\frac{-K+\beta-w}{2(\beta+\sigma)}\right) < \frac{1}{4}.$$

This implies that for $p \in [0, \overline{p}_1[$, there exists $\widehat{p}_1$ such that for any $p \leq \widehat{p}_1$, the equation $h(q;p) - \frac{1}{4}$ has two solutions. There exists $p^* \leq \bar{p}_1$ such that for any $p \in]\widehat{p}_1, p^*]$, the equation $h(q;p) - \frac{1}{4}$ has no solution.

Step 3. We show that there exists $\hat{p}_2 \in]p^*, \bar{p}_2[$ such that for any $p^* < p \leq \hat{p}_2$, $h(q;p) - \frac{1}{4} = 0$ has no solution, and for any $\bar{p}_2 \geq p > \hat{p}_2$, $h(q;p) - \frac{1}{4} = 0$ has two solutions.

Let us study the function $h\left(.\right)$ on the interval $[\min\{p^*, \overline{p}_1\}, \overline{p}_2]$. First, we have $h\left(q_m(\overline{p}_2); \overline{p}_2\right) = \frac{1}{4}\Delta^B$, which is higher than $\frac{1}{4}$ from part (i) of Assumption 1.

For any $p < \bar{p}_1$, we know from Step 2 that $u(\tilde{p}) < \frac{1}{4}$ and for any $\tilde{p} < p < \bar{p}_1$, the function u is increasing.

For $p \in [\bar{p}_1, \bar{p}_2]$, by using similar arguments to those used in Step 1, we can show that $h(q_t) = q_t(1-q_t)\left[\Delta^B - (1-q_t)p\theta^d\left(\Delta^B - \Delta^S\right)\right]$ admits a unique maximum q_m on the interval $[0,1]$. Moreover,

$$\frac{dh\left(q_m(p);p\right)}{dp} = q_m(p)\left[1-q_m(p)\right]^2\left[-\beta+w+2p(\beta+\sigma)\right]\left(\Delta^B-\Delta^S\right) > 0,$$

since $p > \overline{p}_1 = \frac{\beta-w}{(\beta+\sigma)}$.

We deduce that there exists a unique $\widehat{p}_2$ such that for any $[\min\{p^*, \overline{p}_1\}, \overline{p}_2]$, the equation

$h(q;p) - \frac{1}{4}$ has two solutions. For any $p \in]\widehat{p}_1, \widehat{p}_2[$, $h(q;p) - \frac{1}{4}$ has one solution. The function u is depicted in Figure 7.

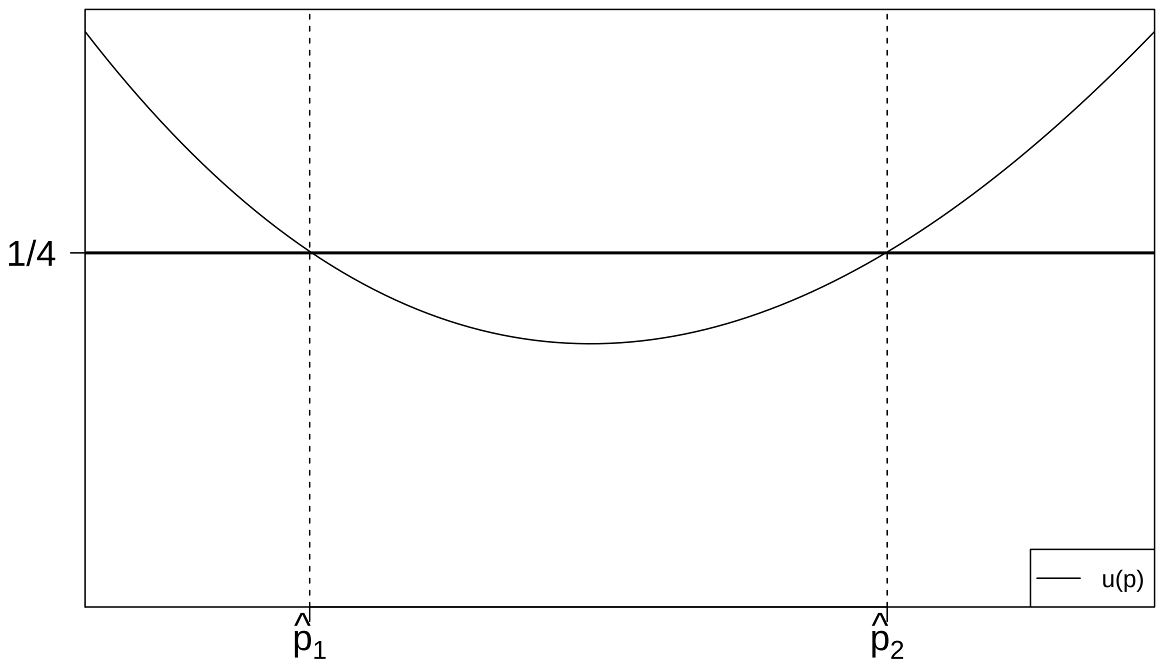

Figure 7: The function $u \equiv h(q_m(p), p)$.

Finally, for $p \in]\widehat{p}_1, \widehat{p}_2[$, the equation $f(q) = 0$ has two solutions $q = 0$ and $q = 1$ with $f'(0) < 0$ and $f'(1) > 0$, meaning that 0 is stable and 1 is unstable. From the continuity of $f(.)$ (and because q is bounded), we deduce that for any $q_0 \in [0, 1]$, the dynamic system globally converges to 0.

For $p \in [0, \widehat{p}_1] \cup [\widehat{p}_2, \bar{p}_2]$, the equation $f(q) = 0$ has four solutions: $q = 0$, $q = 1$, and two interior solutions $\underline{q}$ and $\overline{q}$ with $f'(0) < 0$, $f'(1) > 0$, $f'(\underline{q}) > 0$, and $f'(\overline{q}) < 0$. We conclude that for any $q_0 \in \left[0, \underline{q}\right[$, the sequence q_t converges to 0, while for any $q_0 \in [\underline{q}, 1]$, the sequence q_t converges to $\overline{q}$.

To show item (iii), let us do to some comparative statics exercise and look at the derivatives of $\underline{q}$ and $\overline{q}$ with respect to p. First consider the case $p \in [0, \widehat{p}_1]$. The implicit function theorem gives

$$\frac{\partial \overline{q}}{\partial p} = -\frac{\partial h/\partial p \mid_{\overline{q}}}{\partial h/\partial q \mid_{\overline{q}}},$$
$$\frac{\partial \underline{q}}{\partial p} = -\frac{\partial h/\partial p \mid_{\underline{q}}}{\partial h/\partial q \mid_{\underline{q}}}.$$

We know from the above that $\partial h/\partial q \mid_{\overline{q}} < 0$ while $\partial h/\partial q \mid_{\underline{q}} > 0$. The sign of $\partial \overline{q}/\partial p$ is the sign of $\partial h/\partial p \mid_{\overline{q}}$. The sign of $\partial \underline{q}/\partial p$ is the opposite sign of $\partial h/\partial p \mid_{\underline{q}}$. We also know from the above that $\partial h/\partial p$ is the sign of $qK - (\beta - 2p(\beta+\sigma) - w)$. Further,

$$\frac{\partial h}{\partial p} < 0,$$

since we showed that $q_m(p)K - (\beta - 2p(\beta+\sigma) - w) < 0$, $\forall p \in [0, \widehat{p}_1]$.

• First consider the derivative of $\underline{q}$ with respect to p. We know that $\underline{q} < q_m(p)$ so that we can deduce that $\underline{q}K - (\beta - 2p(\beta+\sigma) - w) \leq 0$ which is equivalent to

$$\frac{\partial \underline{q}}{\partial p} > 0.$$

• Second, consider the derivative of $\overline{q}$ with respect to p. Note that at $p = \hat{p}_1$, we have (i) $q_m(\hat{p}_1) = \overline{q} \mid_{\hat{p}_1}$ and (ii) $\frac{\partial h}{\partial p} \mid_{q_m} < 0$ implying

$$\frac{\partial \overline{q}}{\partial p} \mid_{\hat{p}_1} < 0,$$
$$\Leftrightarrow \overline{q} \mid_{\hat{p}_1} K - (\beta - 2p(\beta+\sigma) - w) < 0.$$

We use a reductio ad absurdum argument to show that $\partial \overline{q}/\partial p < 0 \ \forall p \in [0, \widehat{p}_1]$.

Suppose that there exists $p^* \in [0, \widehat{p}_1]$ such that $\partial \overline{q}/\partial p|_{p^*} > 0$. Given that $\frac{\partial \overline{q}}{\partial p} \mid_{\hat{p}_1} < 0$, the continuity of $\overline{q}$ in p and the continuity of the derivative $\partial \overline{q}/\partial p$ in p, it implies that there exists some $p < \hat{p}_1$ such that $\partial \overline{q}/\partial p = 0$, equivalent to $\overline{q}K - \left(\beta - 2p^{'}(\beta+\sigma) - w\right) = 0$. Denote by p^{**} the value of p which globally minimizes $\overline{q}$ on the interval $[0, \hat{p}_1]$. By definition, $\overline{q} \mid_{p^{**}} < \overline{q} \mid_{\hat{p}_1}$ equivalent to $\overline{q} \mid_{p^{**}} K - \left(\beta - 2p^{'}(\beta+\sigma) - w\right) < \overline{q} \mid_{\hat{p}_1} K - \left(\beta - 2p^{'}(\beta+\sigma) - w\right) < 0$ which contradicts the above. We deduce that $\forall p \in [0, \widehat{p}_1]$,

$$\frac{\partial \overline{q}}{\partial p} < 0.$$

A similar reasoning leads to

$$\frac{\partial \underline{q}}{\partial p} > 0,$$
$$\frac{\partial \overline{q}}{\partial p} < 0,$$

$\forall p \in [\widehat{p}_2, 1]$. ∎

Proof of Proposition 3: There is segregation if and only if type-h individuals are willing to bid more than type-d parents to live with individuals of their type (i.e., to live in neighborhood 1). Formally, the segregated equilibrium exists and is unique if and only if

$$\Delta\rho(q_1) > 0 \quad \forall\, q_1 \geq \frac{Q}{2}.$$

In this kind of model, the symmetric equilibrium always exists. We must check whether it is stable. Stability is defined as follows: for a small rise (resp. decrease) in the proportion of type-h agents in neighborhood 1, agents of type d (resp. h) are willing to bid more than agents of type h (resp. d). Formally, the symmetric equilibrium is stable if and only if

$$\frac{d\Delta\rho}{dq_1}|_{Q/2} < 0.$$

The symmetric equilibrium is unique if and only if

$$\frac{d\Delta\rho}{dq_1} \leq 0 \quad \forall\, q_1 \geq \frac{Q}{2}.$$

The function $\Delta\rho$ is positive for any q_1 if and only if

$$\begin{aligned} & q_1(1-q_1) > q_2(1-q_2), \\ & \Leftrightarrow q_1 - \frac{1}{2} > \frac{1}{2} - q_2, \\ & \Leftrightarrow q_1 - \frac{1}{2} > \frac{1}{2} - Q + q_1, \\ & \Leftrightarrow 1 - Q < 0. \end{aligned}$$

We deduce that

$$\begin{aligned} \Delta\rho(q_1) > 0, \quad \forall\, q_1 \geq \frac{Q}{2} \quad & \Leftrightarrow Q < 1, \\ \Delta\rho(q_1) < 0, \quad \forall\, q_1 \geq \frac{Q}{2} \quad & \Leftrightarrow Q > 1. \end{aligned}$$

This proves the result. ∎

Proof of Proposition 4:

(i) Suppose that $Q_0 \in \left[0, \underline{q}\right[$, i.e., the initial proportion of honest families is low. First, from Proposition 3, at time $t = 0$, since $Q_0 < 1$, the unique urban equilibrium is *segregated* and we have $q_{1,0} \in \left[0, \underline{q}\right[$and $q_{2,0} = 0$. After the socialization choices have been made, we have $q_{2,1} = f(q_{2,0}) = 0$, and since $q_{1,0} < \underline{q}$, $q_{1,1} = f(q_{1,0}) < f(\underline{q}) = \underline{q}$. We deduce that for all $t \geq 0$, $q_{1,t} < \underline{q}$, $q_{2,t} = 0$. For any $q_{1,t} \in \left[0, \underline{q}\right[$, the sequence $q_{1,t}$ is decreasing and converges to 0 (see the proof of Proposition 1). Hence, in the long run, we have $q_1^* = 0$ and $q_2^* = 0$.

(ii) Suppose now that $Q_0 \in \left[\underline{q}, 2\underline{q}\right[$.

If $Q_0 < 1$, the unique urban equilibrium at time $t = 0$ is *segregated* so that $q_{1,0} \in \left[\underline{q}, 1\right[$ and $q_{2,0} = 0$. After the socialization choices have been made, we have $q_{2,1} = f(q_{2,0}) = 0$, and since $q_{1,0} \geq \underline{q}$, $q_{1,1} = f(q_{1,0}) \geq f(\underline{q}) = \underline{q}$. From the arguments developed in the proof of Proposition 1, we have that for any $q_{1,t} \in \left[\underline{q}, \overline{q}\right]$ (resp. $[\overline{q}, 1]$) the sequence $q_{1,t}$ is increasing (resp. decreasing) and converges to $\overline{q}$ (resp. $\overline{q}$). In the long run, we thus have $q_1^* = \overline{q}$ and $q_2^* = 0$.

If $Q_0 > 1$, the unique urban equilibrium at time $t = 0$ is *integrated* so that $q_{1,0} = q_{2,0} = \frac{Q_0}{2}$. After the socialization choices have been made, we have $q_{1,1} = q_{2,1} = f(q_{2,0}) < f(\frac{Q_0}{2}) < \underline{q}$. From the arguments developed in the proof of Proposition 1, we have that for any $q_{1,t}, q_{2,t} \in$

$\left[0, \underline{q}\right[$, the sequences $q_{1,t}$ and $q_{2,t}$ are decreasing. We deduce that there exists some t such that $q_{1,t} + q_{2,t} = Q_t < 1$ (segregated equilibrium) and we are exactly in the above case so that $q_1^* = \overline{q}$ and $q_2^* = 0$.

(iii) Suppose that $Q_0 \in \left[2\underline{q}, 2\right]$. By using similar reasoning, we can deduce that the sequences $q_{1,t}$ and $q_{2,t}$ are increasing for $q_{1,t} = q_{2,t} \in \left[2\underline{q}, 2\overline{q}\right]$ and decreasing for any $q_{1,t} = q_{2,t} \in [2\overline{q}, 2]$. We deduce that, in the long run, we have $q_1^* = q_2^* = \overline{q}$. ∎

Proof of Proposition 5:

Part (i): Suppose that $Q_0 \in \left[\underline{q}^g, 2q^g_{max}\right[$, where $q^g_{max} = \max\{\underline{q}^g, \frac{1}{2}\}$, and $Q_0 < \max\{2\underline{q}^{g'}, 1\}$. Given Proposition 4, we deduce that under the education policy $q_1^* = \overline{q}^{g'}$ and $q_2^* = 0$, which implies that the long-run crime rate under the policy are given by:

$$C_1^{g'} = \overline{q}^{g'}\theta^h + (1 - \overline{q}^{g'})\theta^d,$$
$$C_2^{g'} = \theta^d,$$

where

$$\theta^h = (1-p)\beta - p\sigma - K - w(g'),$$
$$\theta^d = (1-p)\beta - p\sigma - w(g').$$

Let us perform the reduction in the long-run crime rate induced by the policy in neighborhoods 1 and 2. We have:

$$\Delta C^1 = C_1^{g'} - C_1^g = w(g) - w(g') + K(\overline{q}^g - \overline{q}^{g'}),$$
$$\Delta C^2 = C_2^{g'} - C_2^g = w(g) - w(g').$$

We have:

$$\Delta C^1 < \Delta C^2 \quad \Leftrightarrow \overline{q}^g < \overline{q}^{g'}.$$

Given Proposition B1 and the fact that $w'(g) > 0$, we deduce that:

$$\frac{\partial \overline{q}^g}{\partial g} > 0,$$

so that $\overline{q}^g < \overline{q}^{g'}$.

Part (*ii*): Suppose that $Q_0 \in \left[\underline{q}^g, 2q^g_{max}\right[$, where $q^g_{max} = \max\{\underline{q}^g, \frac{1}{2}\}$, and $Q_0 > \max\{2\underline{q}^{g'}, 1\}$. Given Proposition 4, we deduce that under the education policy $q_1^* = q_2^* = \overline{q}^{g'}$, which implies that the long-run crime rates under the policy are given by

$$C_1^{g'} = C_2^{g'} = \overline{q}^{g'}\theta^h + (1 - \overline{q}^{g'})\theta^d.$$

Let us perform the reduction in the long-run crime rate induced by the policy in neighborhoods 1 and 2. We have:

$$\Delta C^1 = C_1^{g'} - C_1^g = w(g) - w(g') + K(\overline{q}^g - \overline{q}^{g'}),$$
$$\Delta C^2 = C_2^{g'} - C_2^g = w(g) - w(g') - \overline{q}^{g'}.$$

We have

$$\Delta C^1 > \Delta C^2 \quad \Leftrightarrow K\overline{q}^g > 0.$$

This completes the proof. ∎