DISCUSSION PAPER SERIES

DP13014

CRIME, BROKEN FAMILIES, AND PUNISHMENT

Emeline Bezin, Thierry Verdier and Yves Zenou

PUBLIC ECONOMICS

ISSN 0265-8003

CRIME, BROKEN FAMILIES, AND PUNISHMENT

Emeline Bezin, Thierry Verdier and Yves Zenou

Discussion Paper DP13014
Published 29 June 2018
Submitted 29 June 2018

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

CRIME, BROKEN FAMILIES, AND PUNISHMENT

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 a culture of honesty 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 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 and high crime rates. This is because for children who come from these broken families, negative community experiences (peer effects) further encourage their criminal participation. Finally, we discuss the efficiency of location and family policies on long-term crime rates.

JEL Classification: K42, J15, Z13

Keywords: crime, Social interactions, neighborhood segregation

Emeline Bezin - emeline.bezin@gmail.com
Paris School of Economics and IRES (UCL)

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

Yves Zenou - yves.zenou@monash.edu
Monash University and CEPR

Acknowledgements
This research was undertaken under the project TECTACOM. The financial support from the European Research Council (ERC) grant number 324004 is gratefully acknowledged.

Crime, Broken Families, and Punishment*

Emeline Bezin[†] Thierry Verdier[‡] Yves Zenou[§]

June 21, 2018

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 a culture of honesty 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 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 and high crime rates. This is because for children who come from these broken families, negative community experiences (peer effects) further encourage their criminal participation. Finally, we discuss the efficiency of location and family policies on long-term crime rates.

JEL Classification: K42, J15, Z13.

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

*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

In the United States, nearly four in 10 births are to unmarried women (Ventura, 2009) and the proportion of children under age 18 living in mother-only families rose from 8% in 1960 to 23% in 2010 (U.S. Census Bureau, 2010). Overall, 30% of U.S. children are estimated to spend some time living in stepfamilies (Bumpass et al., 1995). If we break down these numbers by race, we see that African-American families have a much higher rate. In 2010, 48.5% of black children lived with only their mother (Figure 1), which is more than double the proportion of white children (18.3%) and much higher than that of Hispanic children (26.3%). Moreover, in 2016, the National Longitudinal Survey of Youth (Figure 2) showed that only 5% of youths who grew up in an intact married family had ever been arrested, followed by youths from married stepfamilies and families with intact cohabiting partners (8%), single divorced parent families (9%), and cohabiting stepfamilies and always single-parent families (13%). The results of this survey thus imply that 22% of adolescents living in a single-parent family have been arrested compared with 5% in intact married families,[1] showing a clear positive correlation between the family structure of adolescents and their criminal behavior.[2] The mechanism behind this correlation is, however, unclear, and there is strong evidence that this pattern is spatially differentiated (see Section 2.4).

In this study, we argue that having a father at home might be crucial to prevent susceptible boys from becoming criminals. This is because the transmission of a trait against criminal behavior (criminal ethics trait) is easier in a two-parent family. We also show that increasing the cost of committing crime (by increasing deterrence and imposing longer sentences) can backfire. This is because incarceration can "destroy" biparental families by putting fathers in prison, which, in turn, makes it more likely that their son will become a criminal. Indeed, incarceration threatens the earning power of the remaining family members and makes children growing up in these families much more vulnerable to crime. We also argue that children who come from these broken families tend to have negative community experiences (peer effects), which further encourages their criminal participation. This finding implies that location matters and may explain why some neighborhoods end up with

[1]See also Richters and Martinez (1993) who show that only 6% of children from stable, safe homes become delinquent. Meanwhile, 90% of children from homes rated as both unstable and unsafe (broken marriage or lack of supervision) become delinquent.

[2]See also the evidence in Section 2.3, which shows the positive correlation between the father's and son's involvement in criminal activities.

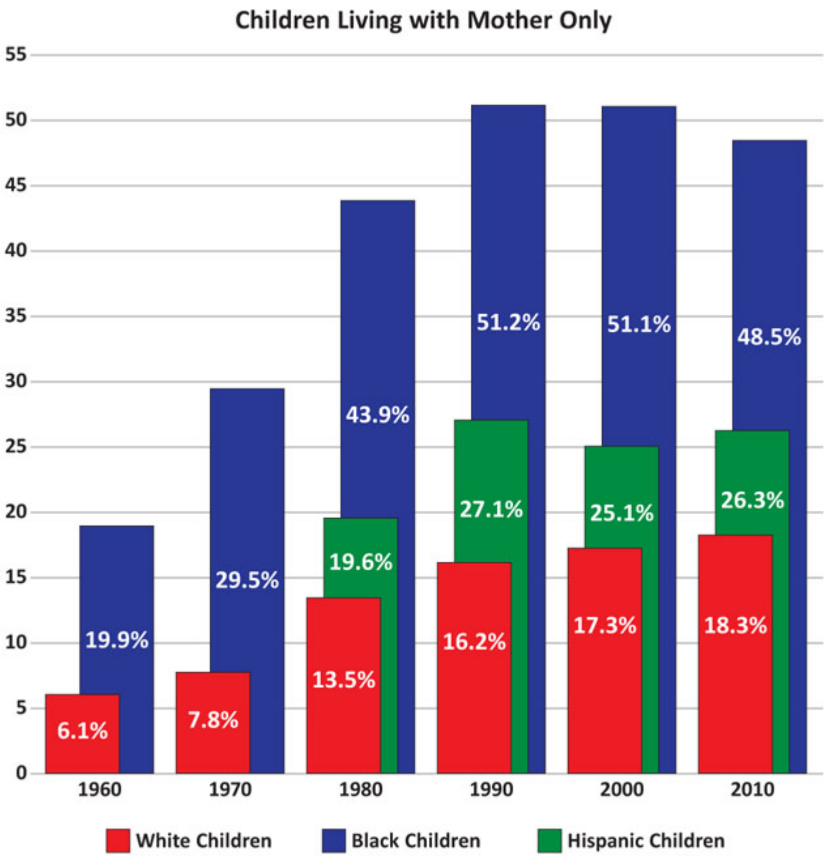

Figure 1: Percentage of children living with their mother only by year and race.
Source: U.S. Census Bureau, "Living Arrangements of Children Under 18," July 1, 2012.

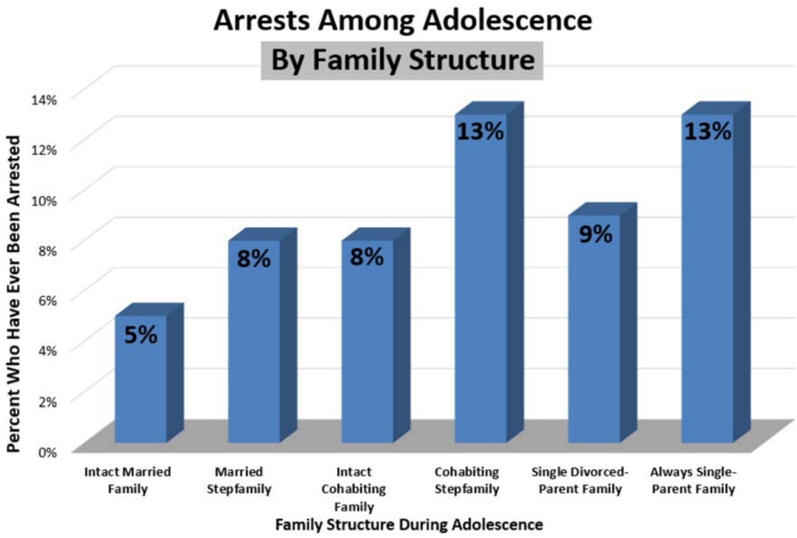

Figure 2: Arrests among adolescents by family structure.
Source: National Longitudinal Survey of Youth 2016.

a large proportion of broken families and a very high crime level.

In particular, 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" trait or the "dishonest" 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 *inclined* to commit crime in the next period. 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 three 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 last is demonstrating how widespread is the initial culture of honesty in society.

Specifically, we show that for intermediate levels of crime repression, the unique stable steady-state equilibrium is such that there are no honest individuals in the population (i.e., the crime level is very high). Conversely, when crime repression is weak or strong, then the steady-state proportion of honest individuals typically depends on the initial conditions of the culture of honesty in the population. In particular, if q_0, the initial proportion of honest individuals in the population, is sufficiently large, then the long-run steady state is composed of a combination of honest and dishonest individuals and the crime rate takes an intermediate value.

The intuition for these results is as follows. When the initial proportion of honest individuals in the population is low, many fathers commit crime and end up in prison. Families are then mostly monoparental and have, therefore, a high cost of transmitting the honesty trait. Consequently, in the long run, the culture of honesty completely disappears from the population.

Conversely, when the initial proportion of honest individuals is large, the steady-state outcome mainly depends on the intensity of crime repression. When crime repression is weak, there are few incarcerations of young men and few single-mother families, which makes the honesty trait more likely to be transmitted. When, however, crime repression increases, the *deterrence* and *social disorganization* effects act in opposite directions. On the one hand, because of deterrence, few male individuals decide to commit crime and, thus, more non-criminal fathers contribute to biparental families. On the other hand, those who decide to commit crime are more likely to be arrested, which leads to more broken families. The net impact of crime repression on the structure of the family is, therefore, unclear. The cultural dynamics of the honesty trait depend on the relative strengths of the deterrence versus the social disorganization effects, which, in turn, depend on the initial proportion of honest individuals in the population.

Our second contribution is to analyze a policy that aims to reduce crime overall. 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 and there is a non-monotonic relationship between long-run crime and the degree of crime repression. 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.

We also analyze a policy that subsidizes single-mother families and show that this policy can be efficient (in the sense that it minimizes crime) only if q_0, the initial proportion of honest individuals in the population, is sufficiently large. This policy can also backfire when q_0 is sufficiently large since an increase in the subsidy has a non-monotonic impact on long-run crime.

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 land 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 at which, in one neighborhood, the crime rate is high and most families are dishonest and comprising single mothers, whereas the opposite is true in the other neighborhood.

In this respect, we are able to provide a microfoundation of the equilibrium urban models a la Benabou (1993) in which segregation occurs because of positive externalities (the higher the number of educated individuals in a neighborhood, the lower is the cost of being educated). Here, we have a dynamic perspective and show that positive externalities only exist when the initial proportion of honest families is sufficiently low (cultural complementarity between parents' socialization effort and proportion of honest families in a neighborhood). In this case, we show that parents with the honesty trait residing in one neighborhood increase their socialization effort and are able to bid away the parents with the dishonesty trait to the other neighborhood. This leads to a segregated equilibrium at which crime and the structure of the family are spatially differentiated.

We finally compare an efficient policy, which aims to minimize crime, with an equitable spatial policy, which aims to minimize the crime difference between the two neighborhoods. We determine when the market solution is efficient and when efficiency conflicts with equity. We demonstrate that in some cases, segregation can be efficient because it minimizes total crime.

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.[3] As Kleiman shows, "zero tolerance" is nonsense: there are always more offenses than there is punishment.[4] Is there an alternative to brute force? In this study, we argue that the deterrence capacity of the justice system is insufficient 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 of high crime in a neighborhood. As a result, an effective crime policy should also take into account the impact of incarceration on the structure of families and how this then affects the criminal behavior of future generations.

The rest of the paper unfolds as follows. In the next section, we discuss our contribution to the related literature. In Section 3, we present our benchmark model, determine the long-run Equilibrium, and analyze a policy to reduce total crime. In Section 4, we extend our model to introduce location choices. Finally, Section 5 concludes. In Online Appendix A, we provide all the proofs of the results stated in this paper. In Online Appendix B, we determine the urban equilibrium when location choices are endogenous.

[3] Over the past 30 years, incarceration in the United States has increased 500 percent so that it is now the world leader with 2.2 million people in the nation's prisons and jails (U.S. Department of Justice, 2014).

[4] In the standard crime literature (Becker, 1968; Garoupa, 1997; Polinsky and Shavell, 2000), punishment is seen as an effective tool for reducing crime.

2 Related literature

Our study is related to different literatures.

2.1 Social interactions and crime

A growing body of the empirical literature in economics suggests that peer effects are important in criminal activities. In the economics literature, Glaeser et al. (1996) show that the number of social interactions is highest in petty crimes and moderate in more serious crimes. Ludwig et al. (2001) and Kling et al. (2005) study the relocation of families from high- to low-poverty neighborhoods by using data from the Moving to Opportunity experiment. They find that this policy reduces juvenile arrests for violent offences by 30–50% relative to the control group. Bayer et al. (2009) consider the extent to which juvenile offenders serving time in the same correctional facility influence each other's subsequent criminal behavior. They also find strong evidence of learning effects in criminal activities since exposure to peers with a history of committing a particular crime increases the probability that an individual who has already committed the same type of crime recidivates that crime.[5]

More recently, Damm and Dustmann (2014) and Corno (2017) investigate the influence of friends on crime.[6] The former exploit a Danish natural experiment that randomly allocates parents of young children to neighborhoods with different shares of youth criminals, while the latter uses data collected among the homeless. Both find strong peer effects in crime.[7]

From a theoretical viewpoint, Glaeser et al. (1996) were among the first to develop a crime social interaction model in which criminals are grouped into conformists (i.e., they copy what their neighbors do) and non-conformists (i.e., they decide their criminal activities by themselves). The authors show that criminal interconnections act as a social multiplier on aggregate crime. Along this line of research, Calvó-Armengol and Zenou (2004), Ballester et al. (2006, 2010), and Patacchini and Zenou (2012) consider criminal activities in general social network structures and show that the location in the social network of each criminal matters not only through direct connections, but also through connections of connections and so forth.

In our study, peers also play an important role since they determine whether someone

[5]See also Stevenson (2017) who finds similar results.

[6]See also Bernasco et al. (2017) who find significant social interaction effects on crime.

[7]See also Patacchini and Zenou (2012) and Liu et al. (2012) who use a more structural approach.

adopts the honesty trait and thus whether he or she is more likely to commit crime and form a biparental family. Compared with 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.

2.2 Social disorganization, family structure, and crime

Our work is also 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.[8] In particular, it tries to explain why high rates of delinquency persist in certain areas for many years, independent of changes in the composition of the population. 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. Shaw and McKay (1942) postulate that three structural factors, namely low socioeconomic status, ethnic heterogeneity, and residential mobility, disrupt a community's social organization, which in turn explains the spatial variations in the rates of crime and delinquency.[9]

The structural characteristics of a community such as stage of urbanization and degree of family disruption affect the capacity of the community to impose formal and informal controls on its members and outsiders. 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, church, friend networks). These institutions are traditionally responsible for the establishment of organized and cooperative relationships among groups within the local community. This organization is then linked to the bond or "sense of belonging" one might feel to his or her community, which decreases the likelihood of involvement in criminal or delinquent behaviors that might negatively affect that

[8] The term "social disorganization" was first introduced by Hall (1904) to set out an ethical explanation of crime.

[9] Both the systemic model of community crime (Bursik and Grasmick, 1993) and, more recently, collective efficacy theory (Sampson, 2012) have refined the original social disorganization approach, spurring a period of intensive empirical investigation. In particular, the systemic model emphasizes the role of neighborhood-based social network ties in generating informal social control capacity.

community.[10]

To explain criminal behavior, we focus on two key aspects of social disorganization theory: the structure of the family and location. With respect to the family structure aspect, substantial evidence points out the increased rates of mother-headed households (Bureau of the Census, 1994). Moreover, children growing up in single-mother families 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).[11]

In our model, male individuals decide whether to become criminals based on the benefits and costs of crime as well as their degree of honesty that they adopt from their parents and peers (cultural transmission). Depending on whether the father is a criminal and has been arrested, these choices determine the types of families in which children are brought up and socialized to honesty values and norms (single-mother or biparental families). This family structure influences the transmission of the honesty trait to their offspring since biparental families have more time to spend with their children and show them more attention than single-mother families (Florsheim et al., 1998). This fact, in turn, affects the honesty trait for the next generation of young male individuals, who then decide whether to become criminals and so forth.

Ours is the first study that explicitly captures a key aspect of social disorganization theory within an economic framework and investigates how the structure of the family interacts with criminal decisions to explain the long-run evolution of honesty values and crime behaviors.

2.3 Transmission of crime and the father–son correlation in criminal behaviors

Based on works on anthropology and sociology (Boyd and Richerson, 1985; Cavalli-Sforza and Feldman, 1981), the theoretical literature initiated by Bisin and Verdier (2000, 2001) argues that the transmission of a particular trait (e.g., religion, ethnicity, social status) is the outcome of socialization *inside* the family (parents) and/or *outside* the family (peers or

[10] For a recent overview of social disorganization theory, see Porter et al. (2016).

[11] 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.

role models).[12]

In addition, a large body of empirical research[13] provides evidence of substantial intergenerational associations in criminal behavior. The key findings from this literature are that family background and parental criminality are among the strongest predictors of criminal activity, stronger even than one's own income and employment status. For a review of this literature, see Rowe and Farrington (1997), Thornberry (2009), and Hjalmarsson and Lindquist (2012).

Moreover, research on the intergenerational transmission of incarceration (Bhuller et al., 2018; Dobbie et al., 2018) tends to emphasize the strong association between fathers' involvement with the criminal justice system and sons' behavioral outcomes such as experiencing incarceration. Indeed, it is estimated that over 1 million children in EU countries and 2.7 million in the United States have a parent behind bars at some point during any one year (Glaze and Maruschak, 2010, Philbrick et al., 2014). As Bhuller et al. (2018) put it: "roughly one in every 50 children in the EU and one in every 28 children in the United States has a parent in prison in a given year." The mechanism behind this correlation is, however, not well identified. Some researchers (e.g., Murray and Farrington, 2005; Wildeman, 2010; Wildeman and Andersen, 2017) propose that boys who experience paternal incarceration are at an elevated risk of physically aggressive behavior and involvement with the criminal justice system at later stages. A father's incarceration thus directly affects his sons' behavioral problems over and above other types of disadvantages that may have piled up in families with criminal justice contact.

To account for the importance of peer effects in honesty norm transmission, we amend the cultural transmission model a la Bisin–Verdier in the following way. First, all agents agree that one trait (honesty) is better than the other (dishonesty). Second, for a child to be socialized to one trait, the parent's vertical transmission through his effort is insufficient. Indeed, both parents and society's influences matter in a complementary way for a child to acquire the honesty trait. When parents and society send conflicting messages about honesty, the son remains subject to outside socialization (i.e., he is matched a second time with a role model (also met randomly) and adopts the trait of that role model. This set-up allows a richer combination of substitutability and complementarity effects in cultural

[12] For an overview of this literature, see Bisin and Verdier (2011).

[13] See Case and Katz (1991), Williams and Sickles (2002), Duncan et al. (2005), and Hjalmarsson and Lindquist (2010, 2012, 2013).

transmission and offers somewhat different dynamics in the proportion of honest individuals in the population than in the standard Bisin–Verdier framework.

Furthermore, our model can provide a new mechanism that explains the positive correlation between a father's incarceration rate and involvement in crime and his son's incarceration probability and criminal behavior. In our model, this mechanism goes through the structure of families. Indeed, the more criminal the father, the more likely he is to be incarcerated, which implies that the son is more likely to grow up in a single-mother family, which, because of the higher socialization cost of honesty, means he is more likely to commit crime and be incarcerated himself.

2.4 Residential mobility, segregation, and the spatial patterns of crime

Crime is highly concentrated in a limited number of areas within cities. For instance, in U.S. metropolitan areas, after controlling for education, crime rates are much higher in central cities than in suburbs. 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).[14] More generally, U.S. *central cities* have higher crime and unemployment rates, higher population densities, and larger black populations than their corresponding *suburban rings* (South and Crowder, 1997, Table 2). There is also strong 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 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). For example, Florsheim et al. (1998) show that inner-city boys in single-mother families are at a greater risk of developing behavioral problems than boys in two-parent families.

As already mentioned, from a theoretical viewpoint in sociology, social disorganization theory (see Section 2.2) explains why different neighborhoods may experience distinct crime rates based on the family disruption, friendship networks, and institutions such as school

[14] Grogger and Willis (2000, Table 2) also show that central cities are more crime-ridden than suburbs for most crimes. For instance, the mean murder rate in central cities is five times greater than that in suburbs and for property crimes they differ by a factor of two or three.

and church in each neighborhood. Some economics theories also explain 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).[15]

Our study also directly relates to these works by providing a new mechanism that explains the spatial variation of crime and honesty values based on the family structure and local intergenerational cultural transmission processes. In particular, 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 the opposite is true in another neighborhood. In other words, we provide some of the conditions under which an *urban ghetto* emerges and persists over time.

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. 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 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.[16] This simple mechanism implies that family disruption increases with criminal-

[15]For an overview, see O'Flaherty and Sethi (2015).

[16]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

ity.[17] 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

We assume that 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,[18] 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.

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

$$(1 - p)\beta - p\sigma - K - \theta > w.$$

---

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).

[17]For evidence of this issue, see Wildeman (2010), Geller et al. (2011, 2012), and Geller (2013).

[18]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*.

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,[19] 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

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

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. Symmetrically, the child becomes dis-

[19] 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.

honest 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. This pattern of socialization is different from the Bisin–Verdier framework in two ways. First, whereas in Bisin and Verdier (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 (see Section 2.1). 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 t is given by

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

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); (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.

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 $V^{dB} = V^{hS} = V^d$. 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.

Motivated by the empirical evidence (e.g., Wildeman, 2010; Geller, 2013), 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.[20] Incarceration not only limits these contributions, but also threatens the earning power of remaining family members, who may sacrifice work time 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).[21]

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

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

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

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

[20]Even when fathers are not living with their children, they often contribute financially in the form of child support (Geller et al., 2011) and maintain involvement in their children's lives and day-to-day routines (Swisher and Waller, 2008; Tach et al., 2010).

[21]Reeves and Howard (2013) find that parenting skills vary across demographic groups and that 44% of single mothers fall into the weakest category and only 3% in the strongest category. The weak parenting skills found among single parents in the study may be related not only to the lack of a second parent, but also to a lack of income and education. In particular, education stands out as the most critical factor in explaining poor parenting.

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

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

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 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.[22] 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 $q_t > 1/2$.

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

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

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) and denoting $\Delta q_t \equiv q_{t+1} - q_t$, this dynamic equation can be written as

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

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

[22] In the standard Bisin and Verdier (2001) model, either vertical or oblique transmission is sufficient to obtain successful socialization. Our set-up is closer to the extension of Bisin and Verdier (2001, Section 2.2.3) that highlights the complementarities between family and society role models.

To understand the drivers of cultural transmission, it is useful to rewrite the cultural dynamics equation by substituting (5) into (9). After rearrangement, we obtain

$$\Delta q_t = q_t(1-q_t)\left\{[1-pC(q_t,p)]\left(2\tau_t^B-1\right)+pC(q_t,p)\left(2\tau_t^S-1\right)\right\}. \tag{11}$$

Equation (11) describes a classical replicator dynamic equation reflecting how families and society contribute to the cultural evolution of the honesty trait. It highlights three main drivers of cultural change: (*i*) the direct role of *peer effects* (oblique transmission) in socialization to the honesty trait, (*ii*) the reaction of *family socialization* (vertical transmission) to the environment in which they live, and (*iii*) how the *family structure* is affected by crime and incarceration. Let us explain in more details how these three forces operate.

The first driver (peer effects) can be highlighted by considering the hypothetical case in which all families are of the same type and exert an exogenous socialization effort equal to a constant value τ. In such a case, equation (11) collapses to

$$\Delta q_t = q_t(1-q_t)\left(2\tau-1\right).$$

Given the direct role of peer effects in the socialization process, this shows that the rate of vertical transmission through the family has to be sufficiently large (larger than 1/2) to trigger a successful diffusion of the honesty trait in the population.

The second effect (family socialization) can be best illustrated by considering the optimal choice of the socialization effort of parents, abstracting from differences across families. With identical families, the optimal vertical transmission rate can be written as $\tau_t = 2q_t(1-q_t)\Delta$. As above, this equation reflects the fact that the socialization activities inside and outside the family are *cultural complements* for low values of q_t (i.e., when $q_t < 1/2$) and *cultural substitutes* for high values of q_t (i.e., when $q_t > 1/2$). In such a case, equation (11) can be written as

$$\Delta q_t = q_t(1-q_t)\left[4q_t(1-q_t)\Delta-1\right].$$

Assume $\Delta > 1$; in other words, there is a sufficiently high incentive for parents to socialize their offspring to the honesty trait. Then, two features emerge from this equation. First, with a low initial proportion q_0 of honesty traits in the population, the cultural complementarity effect prevents the diffusion of the honesty trait and, therefore, the long-run steady-state population will only be composed of non-honest individuals (i.e., the steady state $q^* = 0$ is stable). By contrast, with a sufficiently large initial q_0, owing to the same cultural complementarity effect, we will end up with a successful diffusion of the honesty trait. However, when q_t increases and crosses 1/2, family and peers become cultural substitutes, which provides a force inhibiting the further diffusion of the honesty trait in the population and thus prevents a steady state in which the whole population adopts the honesty trait. We will thus end up with a stable interior steady state $q_+^* \in (0,1)$, with a relatively large proportion of honest individuals.[23]

Finally, the third effect (family structure) is captured by the two terms inside the bracket of equation (11). The first term $[1 - pC(q_t, p)] \left(2\tau_t^B - 1\right)$ reflects the total effect of the socialization effort τ_t^B of biparental families,[24] while the second term $pC(q_t, p) \left(2\tau_t^S - 1\right)$ reflects the socialization effort τ_t^S of single-mother families.[25] The crime rate and incarceration create a compositional family effect on cultural transmission since the larger the number of monoparental families, the weaker is the transmission of the honesty trait in the population as $\tau_t^S < \tau_t^B$.

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

(i) $\Delta^B > 1$

(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

[23] Denote

$$q_-^* = \frac{1 - \sqrt{\frac{\Delta - 1}{\Delta}}}{2} \text{ and } q_+^* = \frac{1 + \sqrt{\frac{\Delta - 1}{\Delta}}}{2}.$$

It is then easily verified that for $q_0 < q_-^*$, the population converges to $q^* = 0$, while for $q_0 > q_-^*$, there will be a unique stable interior steady state equal to q_+^*.

[24] There are, indeed, $1 - pC(q_t, p)$ biparental families, i.e., $1 - C(q_t, p)$ families with non-criminal fathers and $(1 - p)C(q_t, p)$ families with a criminal father that has not been caught.

[25] There are, indeed, $pC(q_t, p)$ single-mother families, i.e., families for which the father is a criminal and has been caught.

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.

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 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* $q^* = \overline{q}$, *where* $\overline{q} \in]0, 1[$.

There is a non-monotonic inverted U-shaped relationship between the number of monoparental families $pC(q_t, p)$ and p, the probability of being arrested. Indeed, when p is low, incarceration is weak, while when p is high, crime is deterred. Hence, in both these cases, the number of monoparental families $pC(q_t, p)$ is low. Consequently, the effect of crime repression on the family structure is the highest for intermediate values of p.

Consider now 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 ($q^* = 0$). Figure 3 illustrates these dynamics and shows the other equilibrium for which $q^* = 1$ is unstable. Indeed, when $q^* = 1$, all men are honest and crime is low. However, because p is relatively high, young men committing crime end up in prison and, therefore, many single-mother families are formed. This, along with the fact that parents make little effort socializing their offspring to the honesty trait (see (8)), results in the reduction in the proportion of honest people in the population; hence, $q^* = 1$ is unstable. On the contrary, the equilibrium $q^* = 0$ is stable. In this case, no male is honest and most commit crime. Because p is relatively high, many single-mother families are formed, which, in turn, reinforces the fact that the dishonesty trait is adopted by individuals in the population.

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 illustrated in Figure 4. As we know, for low or high values of p, the number of monoparental families $pC(q_t, p)$

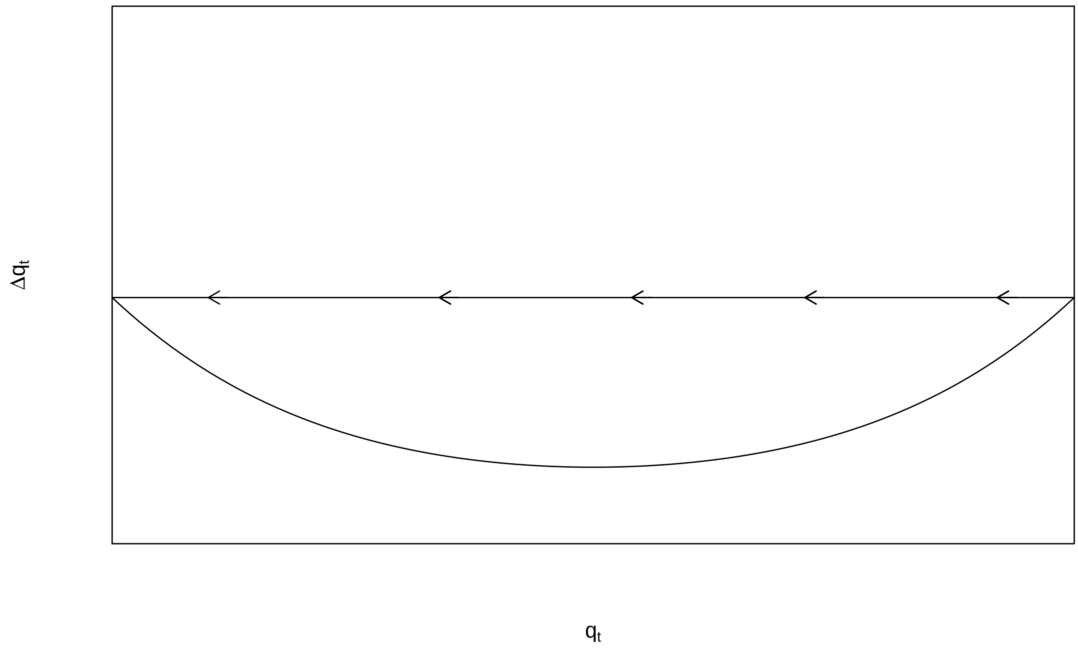

Figure 3: Dynamics of honesty with intermediate incarceration rates, $p \in]\hat{p}_1, \hat{p}_2[$

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).

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 $q^* = \overline{q} > \underline{q}$. 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.

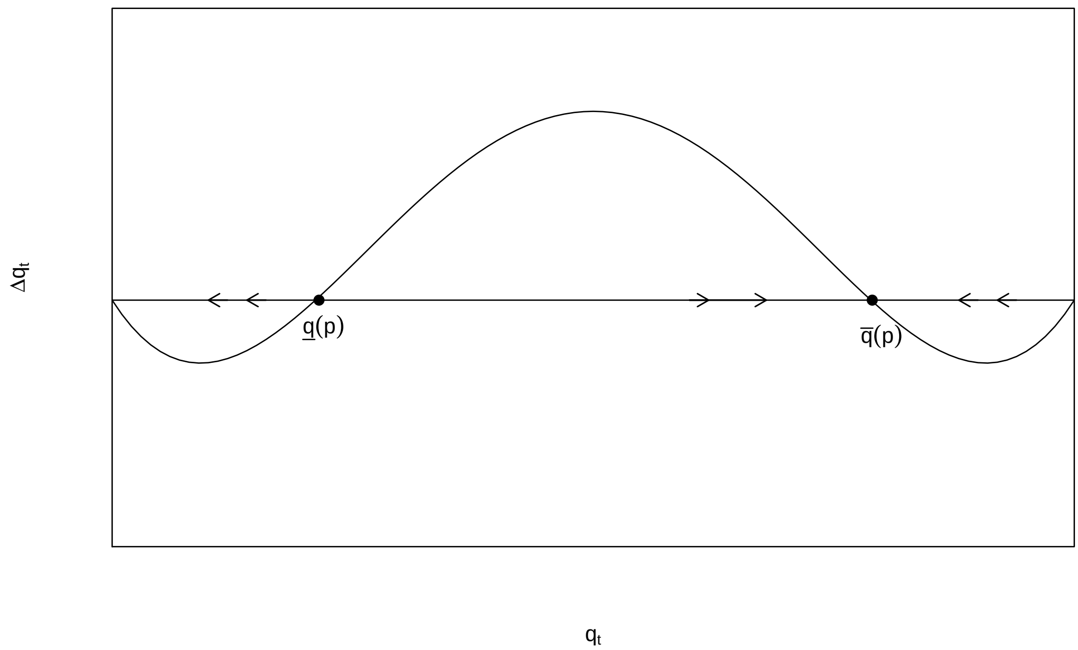

Figure 4: Dynamics of honesty with low or high incarceration rates, $p \in [0, \hat{p}_1] \cup [\hat{p}_2, \bar{p}_2]$

Proposition 2 *Suppose that Assumption 1 holds and that* $p \in [0, \widehat{p}_1] \cup [\widehat{p}_2, \overline{p}_2]$*. Then, an increase in* w *or* σ *or a decrease in* β*,* c^S*, or* c^B *increases* $\overline{q}$*, the long-run proportion of honest individuals in the population.*

Proposition 2 provides some comparative statics results of the interior equilibrium $q^* = \overline{q}$. 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. Indeed, when fewer incentives to commit crime exist, few individuals become criminals and, thus, more families are biparental. This facilitates the transmission of the honesty trait because it is less costly for these families to exert a socialization effort. At any time point, this increases q_t, which reinforces the transmission of the honesty trait since both parents and peers are more likely to influence the child.

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. We now focus on the impact of incarceration on long-run crime. For this purpose, let us define the functions $\underline{q}(p)$ and $\bar{q}(p)$ on $[0,1] \times [0,1] \rightarrow]0,1[$ implicitly defined by $\Delta q_t = 0$ (see (10)). These functions are represented in Figure 5, which depicts a bifurcation diagram of the cultural dynamics map as p varies. This figure illustrates the results stated in Proposition 1 where there can only be a stable interior steady-state equilibrium $q^* = \overline{q}(p)$ when p takes either low or high values.

We now turn to the long-run impact of p on the crime rate.

Proposition 3 *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. In particular, small increases in incarceration can have very strong negative effects on honesty. For instance, while an increase in incarceration from

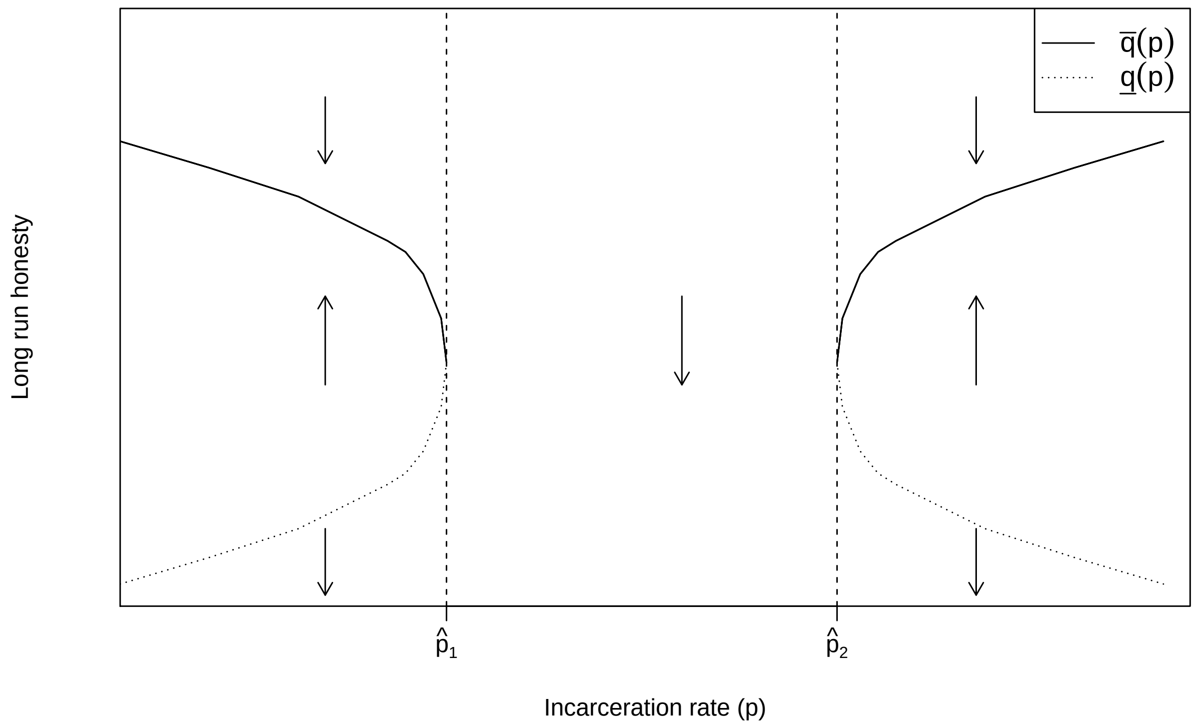

Figure 5: Bifurcation diagram as incarceration rate p varies

$\hat{p}_1 - \epsilon$ to $\hat{p}_1 + \epsilon$ weakly increases the deterrence effect (i.e., decreases the incentive to engage in criminal activities whatever the cultural trait of agents), it strongly reduces honesty (which falls to zero) and thus raises long-run crime substantially.

These results are illustrated in Figure 6, 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 6 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. Further, as in Figure 5, in Figure 6, $\hat{p}_1$ 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.

Finally, part (*iii*) of Proposition 3 shows that strong incarceration policies minimize crime whatever q_0 because when $p \geq \bar{p}_2$, even dishonest individuals have no incentive to commit crime and therefore both short-run and long-run crime are equal to zero.

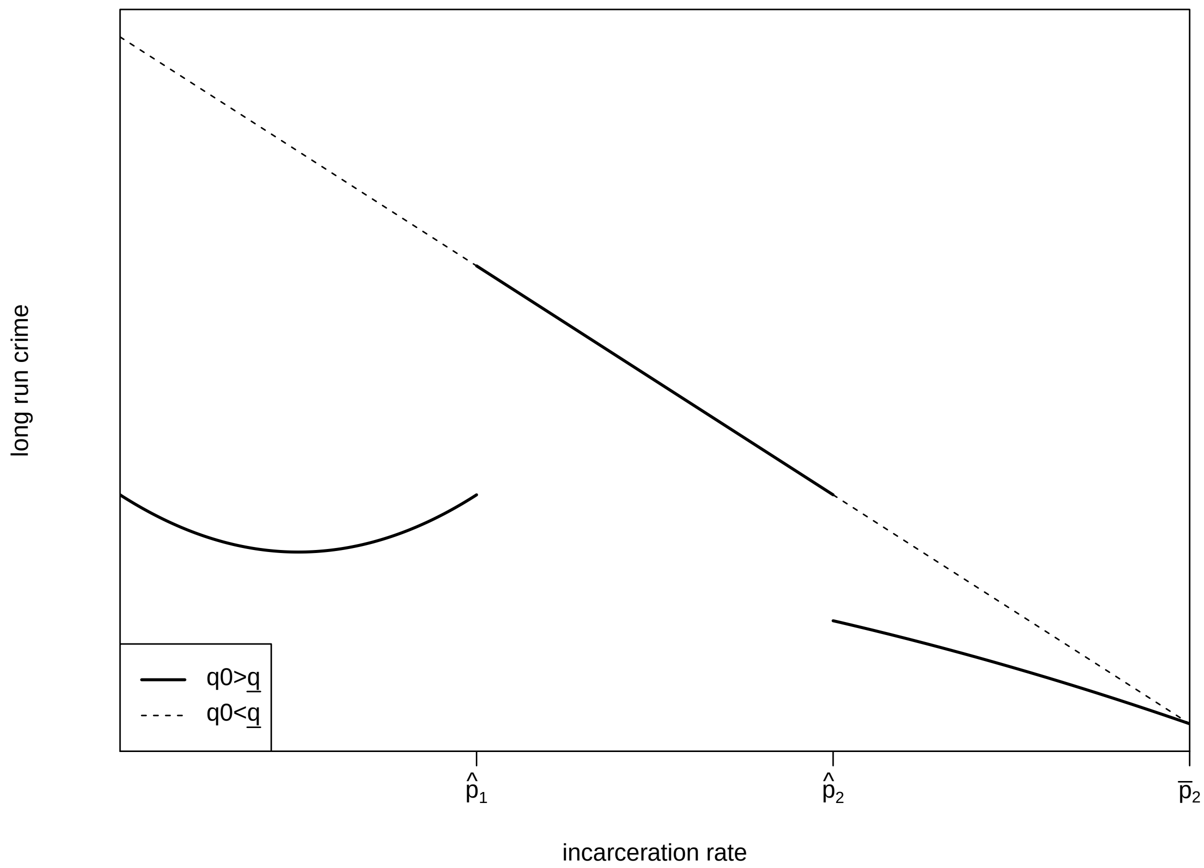

Figure 6: Long-run crime as a function of the incarceration rate p.

3.4 Costly incarceration policy

Figure 6 showed that when $p \geq \bar{p}_2$, long-run crime is minimized. In the real world, however, incarceration is costly, and increasing p to this level may be infeasible. For instance, in the United States, annual justice-related expenditure was \$260 billion in 2012.

We now extend the benchmark model to consider a costly incarceration policy. 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 a public good, say education g (i.e., $R = p + g$). Education increases legal income since more education implies higher productivity, and the wage rate w in the legal sector is $w = w(g)$, with $w'(g) > 0$, $w''(g) < 0$, $w(0) = 0$.

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

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

and the crime rate is still given by

$$C(q_t, p) = q_t\theta^h + (1 - q_t)\theta^d.$$

Let us now consider incarceration policies financed by a fixed budget of R dollars.

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

In other words, an efficient policy p solves

$$\min_p C(q^*(p), p), \tag{12}$$

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 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.

Definition 2

*(i) An incarceration policy is **repressive** if it minimizes short-run crime, i.e., $p = R$.*

(ii) An incarceration policy is ***permissive*** *if it does not minimize short-run crime, i.e., $p < R$.*

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.

However, implementing an incarceration policy that removes all incentives to commit crime seems highly unrealistic. This captures Kleiman's (2009) idea that "they are always more offenses than there is punishment."

Second, consider the case when R takes lower values, i.e., $R \leq \frac{\beta-K}{(\beta+\sigma)}$.[26] When the returns to education are sufficiently high, i.e., $-(\beta+\sigma)+w^{'}(0) > 0$, 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 examine the dynamics of honesty under costly incarceration.

Assumption 2

(i) $\quad -(\beta+\sigma) + w'(0) < 0$

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

Part (i) of Assumption 2 ensures that the policy $p = R$, where all the public budget is spent on repression, minimizes short-run crime. This allows us to concentrate on the trade-off between the short-run and long-run effects of repression on crime in what follows. Part (ii) of Assumption 2 is the counterpart of Assumption 1 when incarceration is costly. We consider the situation where the parent's socialization effort is insufficiently effective to maintain the culture of honesty when the rate of single-parent families is high.

Proposition 4 *Suppose that the public budget satisfies $R < \frac{\beta-K}{2(\beta+\sigma)}$ and that Assumption 2 holds. Then, there exist two thresholds $q^1_{\min} \in (0,1)$ and $q^2_{\min} \in (0,1)$ such that*

[26]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.

(i) When the initial proportion of honest individuals is low, i.e., $q_0 < q^1_{\min}$*, incarceration policies have no effect on long-run honesty.*

(ii) When the initial proportion of honest individuals is sufficiently high, i.e., $q_0 > q^2_{\min}$*, a* ***repressive policy*** $p = R$ *has a negative impact on long-run honesty* q^**.*

When q_0, the initial proportion of honest individuals is sufficiently low and bad transmission by role models (peer effects) is so strong that whatever the incarceration policy, honesty always disappears in the long run, i.e., $q^* = 0$.

When the initial proportion of honest individuals is sufficiently high, incarceration policies matter for the long-run level of honesty q^*. As before, 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. Compared with the previous section, however, the positive impact of crime is reduced since it entails an additional opportunity cost, as it requires a reduction in spending on education. When public resources R are limited (i.e., $R < \frac{\beta - K}{2(\beta+\sigma)}$), 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.

Proposition 5 *Assume that Assumption 2 holds and that*

$$\frac{\beta - K}{2} < \frac{\Delta^B + \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B} K + w\left(\frac{\beta - K}{2(\beta + \sigma)}\right). \tag{13}$$

Consider the two thresholds $q^1_{\min} \in (0,1)$ *and* $q^2_{\min} \in (0,1)$ *in Proposition 4. Then, there exist* $\tilde{R}$*, such that,* $\forall R \in [\tilde{R}, \frac{\beta-K}{2(\beta+\sigma)}]$*:*

(i) When the initial proportion of honest individuals is low, i.e., $q_0 < q^1_{\min}$ *, a* ***repressive policy*** $p = R$ *is efficient.*

(ii) When the initial proportion of honest individuals is high, i.e., $q_0 > q^2_{\min}$ *,* ***permissive policies*** $p \leq \overline{p} < R$ *are efficient.*

When the initial proportion of honest individuals is low, repressive policies are efficient because incarceration does not affect long-run honesty q^* (see Proposition 4). Indeed, whatever the family structure, honesty is initially so rare in the population that it does not persist in the long run. This is equivalent to the Beckerian framework. As explained above, when honesty does not matter, repressive policies are efficient because only the deterrence effect is at work.

When the initial proportion of honest individuals is sufficiently large, however, incarceration affects long-run honesty. In particular, when the budget R takes intermediate values (i.e., $R \in [\tilde{R}, \frac{\beta-K}{2(\beta+\sigma)}]$), incarceration policies may have deep (negative) consequences on community culture because both the crime rate and the incarceration rate are high, which maximizes family disorganization. When condition (13) holds, the long-run crime rate when incarceration is minimal (i.e., $p = 0$) is low compared with the long-run crime rate when incarceration is maximal (i.e., $p_{\max} = \frac{-K+\beta}{2(\beta+\sigma)}$). Indeed, condition (13) imposes that the punishment cost of crime when incarceration is maximal[27] is lower than the expected cost of crime under minimal repression $p = 0$.[28] As a result, when condition (13) holds, the long-run expected incentives to commit crime are higher under maximal repression than under minimal repression and, therefore, any permissive policy with $p < R$ is efficient.[29]

More generally, this proposition shows that the negative effect on the cultural disorganization of incarceration (i.e., the decrease in honesty) outweighs the positive deterrence effect; hence, repressive policies have the unintended effect of increasing long-run crime. As above, this finding implies that even when incarceration policies are costly, raising p may backfire by increasing rather than decreasing crime.

To generate more general results on the efficient repression policy, we now resort to numerical simulations for *any value* of p, namely we solve (12). Figure 7 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 8 provides similar results but for an initially *high* proportion of honest agents, i.e., $q_0 = 0.6$. In both cases, $R = 0.2$.

[27] This is the left-hand side of (13) with $q^* = 0$, which is equal to $p_{\max}(\beta + \sigma) = \frac{\beta-K}{2(\beta+\sigma)}(\beta+\sigma) = \frac{\beta-K}{2}$.

[28] This is when there is a positive long-run proportion $q^* = \overline{q}(0) = \frac{\Delta^B + \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$ of honest individuals in the population, which is the right-hand side of (13)), that is $\overline{q}(0)K + w\left(\frac{\beta-K}{2(\beta+\sigma)}\right)$. This encompasses a pecuniary part, $w\left(\frac{-K+\beta}{2(\beta+\sigma)}\right)$, which is the opportunity wage cost in the legal market when the maximal budget $\frac{\beta-K}{2(\beta+\sigma)}$ is spent on education, and a non-pecuniary moral part, $\overline{q}(0)K$, which is the expected psychological cost from engaging in criminal activities.

[29] Condition (13) holds when socialization rates are low (i.e., Δ^B is large), the psychological cost of crime K is high, the cost of punishment is low, and the proceeds from crime are low.

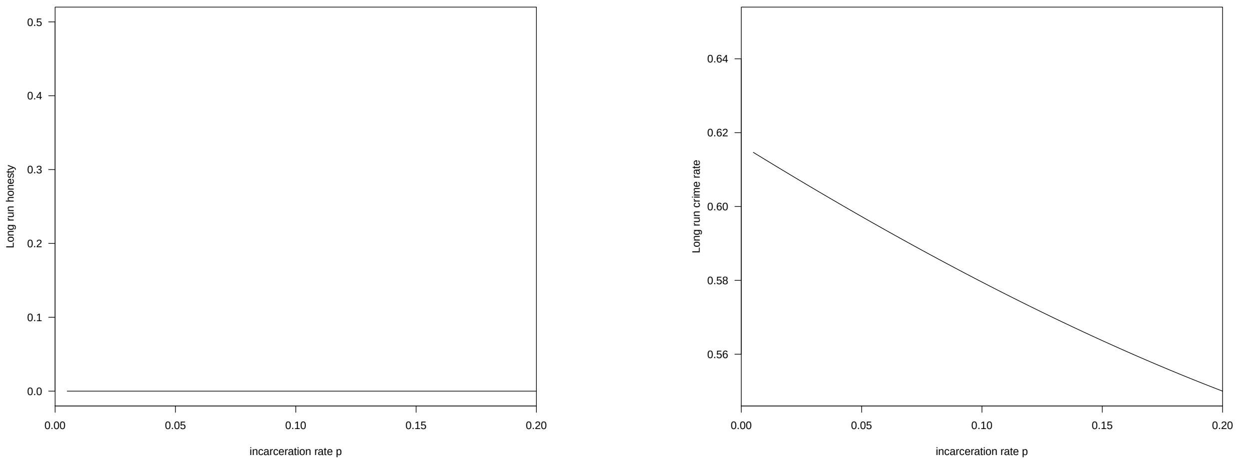

Figure 7: Long-run honesty (left) and the crime rate (right) as a function of the incarceration rate p for $q_0 = 0.1$. We set $w(g) = ag/(1 + g)$ and the parameters are such that $a = 0.5$, $\beta = 0.07$, $K = 1/7$, $\Delta = 0.01$, $c^B = 0.01$, $c^S = 2$, $\sigma = 0.05$, and $R = 0.2$.

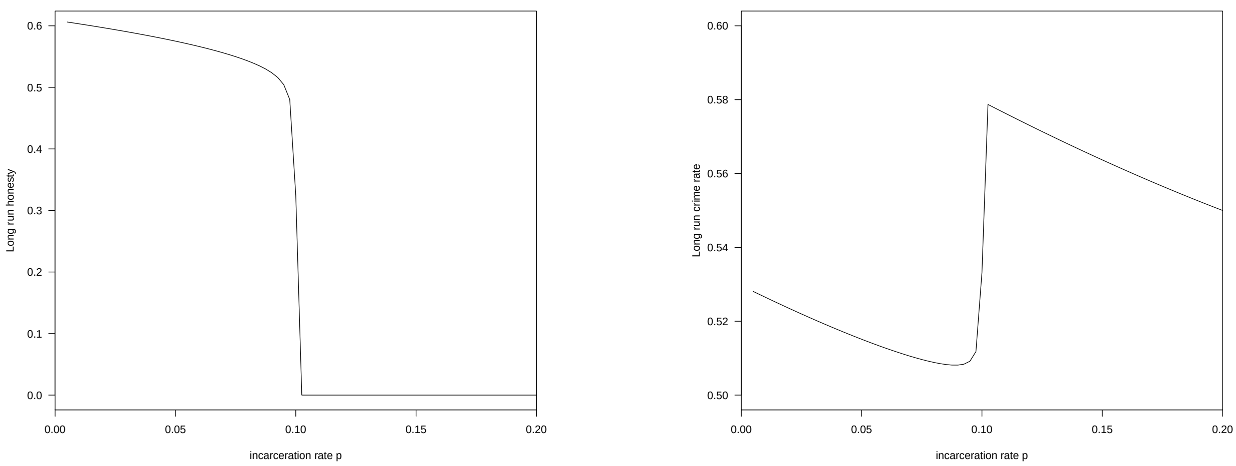

Figure 8: Long-run honesty (left) and the crime rate (right) as a function of the incarceration rate p for $q_0 = 0.6$. We set $w(g) = ag/(1 + g)$ and the parameters are such that $a = 0.5$ $\beta = 0.07$, $K = 1/7$, $\Delta = 0.01$, $c^B = 0.01$, $c^S = 2$, $\sigma = 0.05$, and $R = 0.2$.

As stated in Proposition 5, when the initial proportion of honest individuals is low (Figure 7), 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 8), the effect of an incarceration policy on crime is non-monotonic and may backfire. This example illustrates the result of Proposition 5. 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 reduces because legal income falls, which decreases the opportunity cost of crime. Then, 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 corresponds to an intermediate level of p, given by 0.10 in Figure 8. Figure 6 also shows that when incarceration is costless, the lowest value of long-run crime is obtained when p peaks. In Figure 8, this is not possible because the budget $R = 0.2$ is not sufficiently large for the deterrence effect to outweigh the social disorganization effect.

3.5 Subsidy to single-parent families

Incarceration policies are clearly not the only way in which to reduce crime. In this subsection, we examine a policy that provides financial help in the form of a subsidy to single-parent families.

We assume that the government can grant a subsidy δ to single-mother families, which reduces mothers' socialization cost to $c^S(1-\delta)\frac{(\tau_t)^2}{2}$. In this new framework, the optimal socialization effort of a family of type S is given by

$$\tau_t^S(\delta) = 2q_t(1-q_t)\frac{\Delta^S}{(1-\delta)}.$$

Denote by $C = C(q_t; p, \delta)$ the crime rate at time t as a function of q_t, p but also as a function of the subsidy δ. The government has a budget of R, which it trades off between subsidies for single-mother families, amounting to $pC\delta c^S\frac{\left(\tau_t^S(\delta)\right)^2}{2}$, and spending on education g. The budget constraint can now be written as

$$g(\delta, q_t, C) = R - pC\delta c^S\frac{\left[(\tau_t^S(\delta)\right]^2}{2} = R - 2pC\delta c^S\left[q_t(1-q_t)\frac{\Delta^S}{(1-\delta)}\right]^2. \tag{14}$$

Suppose that $\delta \in [0, 1 - \frac{c^B}{c^S}]$; hence, the maximal subsidy granted by the government is such that the socialization cost of single-mother families is equal to the socialization cost of biparental families. The proportions of honest and dishonest individuals who engage in criminal activities are now given by

$$\begin{aligned}\theta^h &= \max\{\beta - K - p(\beta + \sigma) - w(g(\delta, q_t, C)), 0\}\\ \theta^d &= \max\{\beta - p(\beta + \sigma) - w(g(\delta, q_t, C)), 0\},\end{aligned}$$

and the crime rate $C(q_t; p, \delta)$ is implicitly determined by the following equation:

$$C = -q_t K + \beta - p(\beta + \sigma) - w\left(R - pC\frac{\delta}{(1-\delta)^2}c^S\left[q_t(1-q_t)\Delta^S\right]^2\right). \tag{15}$$

Definition 3 *A subsidy policy δ is efficient if it minimizes long-run crime.*

Formally, an efficient policy δ solves

$$\min_\delta C(q^*(p); p, \delta). \tag{16}$$

Again, we cannot solve this program since we do not know the form of the function $q^*(p)$. Instead, we define two cases: the *no-subsidy policy* (i.e., $\delta = 0$) and the *subsidy policy* (i.e., $\delta > 0$). In particular, for the positive subsidy case, we focus on $\delta = 1 - \frac{c^B}{c^S}$ (because this is convenient for analytical tractability). Whenever the crime rate at $\delta = 1 - \frac{c^B}{c^S}$ is lower than that at $\delta = 0$, we deduce that

$$\text{argmin}_\delta C(q^*(p); p, \delta) \neq 0,$$

suggesting that the subsidy policy is efficient.

Assumption 3

(*i*) $-K + \beta - p(\beta + \sigma) - w(R) > 0.$

(*ii*) $p\frac{(c^S - c^B)}{c^B}\frac{1}{16}\Delta^2 w'(0) < 1.$

(*iii*) $\Delta^B - \frac{(\beta - K - w(1 - \frac{c^B}{c^S}))^2}{4(\beta+\sigma)}(\Delta^B - \Delta^S) - 1 < 1.$

If R and p are such that part (*i*) of Assumption 3 does not hold, then the repression policy removes all incentives to engage in criminal activities. In that case, the subsidy case is inefficient because whatever the long-run honesty, the crime rate is equal to zero. Part (i) excludes this trivial case. Moreover, parts (i) and (ii) together ensure the uniqueness of the crime rate implicitly given by (15). Finally, part (iii) of Assumption 3 is the counterpart of Assumption 1. By considering such situations when social disorganizing forces are strong, if p takes intermediate values (i.e., the proportion of single-parent families is high), honesty cannot be maintained whatever the initial proportion of honest agents.

Proposition 6 *Suppose that Assumption 3 holds and that*

$$R > \frac{\beta^2}{4(\beta+\sigma)}(c^S - c^B). \tag{17}$$

(i) When the initial proportion of honest individuals is low, i.e., $q_0 < \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$*, subsidy policies have no effect on long-run honesty.*

(ii) When the initial proportion of honest individuals is high, i.e., $q_0 > \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$*, the subsidy policy (*$\delta > 0$*) maximizes long-run honesty.*

When R satisfies condition (17), the government budget is sufficiently large to provide a subsidy that cancels out the negative impact of family disruption on the transmission of honesty. In other words, in that case, monoparental and biparental families both have the same cost of transmitting the honesty trait. As a result, when part (i) of Assumption 3 holds, that is $-K + \beta - p(\beta+\sigma) - w(R) > 0$, meaning that the crime rate is positive whatever the value of the subsidy and strong subsidies can be implemented to ensure that (17) holds, the incentives to commit crime are low. Therefore, the proportion of single-parent families is low and the socialization cost of single-parent families c^S is low, and thus the positive effect on socialization outweighs the negative impact on crime.[30] Therefore, if the initial proportion of honest agents is sufficiently high (so that bad peer effects are low), a sufficiently strong subsidy has a positive impact on long-run honesty.

[30]Indeed, subsidies affect long-run honesty in two ways. On the one hand, they have a positive impact on the dynamics of the honesty trait by increasing the socialization efforts of single-parent families. On the other hand, they have a negative short-run (static) impact by increasing incentives to commit crime (i.e., an opportunity cost in terms of education).

Proposition 7 *Suppose that Assumption 3 holds,* $p \in]\hat{p}^1(R), \min\{\hat{p}^2(R), 1\}[$*, and*

$$R \in \left] \frac{\beta^2(c^S - c^B)}{4(\beta+\sigma)}, w^{-1}\left(K\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B}\right)\right[. \tag{18}$$

Then,

(i) If $q_0 < \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$*, the subsidy policy has no impact on long-term crime.*

(ii) If $q_0 > \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$*, the subsidy case is efficient, i.e.,* $\delta > 0$*.*

When the proportion of honest agents is initially low (i.e., society is culturally disorganized), transmission by bad role models is so strong that whatever the effort of parents and, therefore, whatever the subsidy, the culture of honesty cannot expand. Since long-run honesty converges to zero, owing to cultural complementarity, the socialization efforts of both types of parents decrease to zero. This fact implies that spending on the subsidy policy is zero (since this depends on the socialization effort; see (14)). As a result, the subsidy has no impact on long-term crime. Hence, in this case, a subsidy policy should not be implemented.

When the initial proportion of honest individuals is sufficiently high, socialization by parents affects long-run honesty and thus subsidies matter. In particular, when p takes intermediate values (i.e., family disorganization is maximal), some subsidies have deep (positive) consequences on community culture. Indeed, without any policy, socialization rates of honesty are too low to overcome bad peer transmission and the honesty trait disappears in the long run. When R is sufficiently high (see (18)), although family disorganization is high, some subsidies have a sufficiently positive impact on socialization rates to overcome bad peer transmission and thus allow honesty to persist in the long run. The subsidy reduces the negative impact of family disorganization on honesty and long-run crime. Finally, whenever R is not too high (see (18)), the opportunity cost of crime is high, and thus the strong positive impact on long-run honesty outweighs the negative short-run impact on crime (in the short run, raising the subsidy increases crime because it decreases the legal wage).

Again, to provide more general results on the efficient subsidy value, namely the subsidy that solves (16), we run some numerical simulations. Figure 9 displays long-run honesty and the crime rate as a function of the subsidy rate δ for an initially low proportion of honest agents (i.e., $q_0 = 0.1$). Figure 10 depicts long-run honesty and the crime rate as a function of the subsidy rate δ for an initially high proportion of honest agents (i.e., $q_0 = 0.6$).

Figure 9 shows that when q_0 is low, a subsidy policy has no impact on long-run crime. This is because independent of the value of the subsidy, long-run honesty will converge to zero (see Proposition 6). In this case, the socialization efforts of both types of parents are zero and so is the spending on the subsidy policy. Therefore, this policy affects neither community culture nor the opportunity cost of committing crime (through the decrease in legal income). In other words, it is neutral for long-run crime.

By contrast, when q_0 is sufficiently high (Figure 10), the effect of the subsidy δ on long-run crime is negative. Above a certain threshold ($\delta = 0.8$ in the figure), the effect on long-run honesty starts to be different from zero and this leads to a sharp increase in honesty. As a result, long-run crime also strongly decreases when δ is above the threshold value of 0.8 and reaches its lower value when $\delta = 1$, i.e., when the cost of the socialization of single-mother families is totally subsidized and thus equal to that of biparental families. This is because the short-run effect of the subsidy, which increases crime through a reduction in the legal wage, is outweighed by the long-run effect of the subsidy, which reduces crime through an increase in honesty.

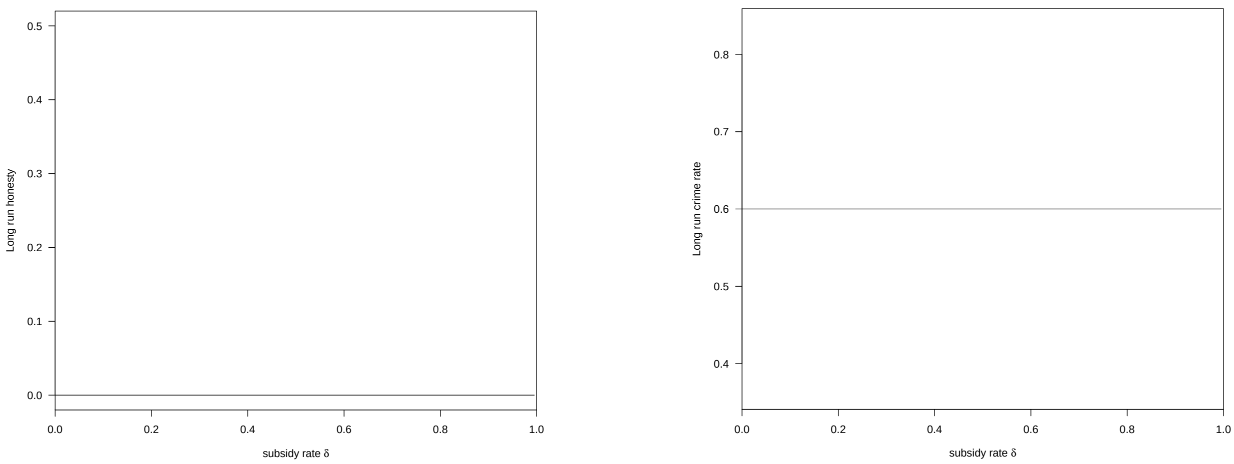

Figure 9: Long-run honesty (left) and the crime rate (right) as a function of the subsidy rate δ for $q_0 = 0.1$. We set $w(g) = ag/(1+g)$ and the parameters such that $a = 0,02$ $\beta = 0.07$, $K = 1/7$, $\Delta = 0.01$, $c^B = 0.01$, $c^S = 2$, $\sigma = 0.05$, $p = 0.15$, and $R = 0.7$.

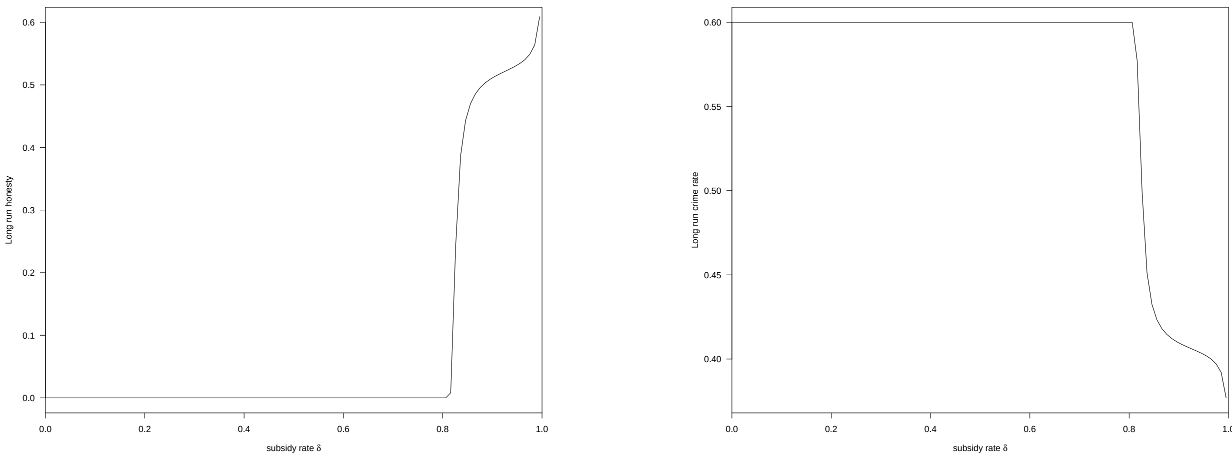

Figure 10: Long-run honesty (left) and the crime rate (right) as a function of the subsidy rate δ for $q_0 = 0.6$. We set $w(g) = ag/(1+g)$ and the parameters such that $a = 0,02$ $\beta = 0.07$, $K = 1/7$, $\Delta = 0.01$, $c^B = 0.01$, $c^S = 2$, $\sigma = 0.05$, $p = 0.15$, and $R = 0.7$.

4 Broken families, segregation, and crime

In Section 2.4, we presented evidence that crime as well as the structure of families and communities are spatially differentiated. In particular, for understanding the variation in neighborhood crime rates, the spatial characteristics of urban environments have been central to criminological theory since the early Chicago School development of the social disorganization approach (Shaw and McKay, 1942). Indeed, under social disorganization theory, family structure is a good predictor of a child's future criminal behavior, which, together with the criminal behavior of other children, leads to the demise of the community in two ways. First, a broken family creates conditions that predispose children to criminal activities. Fatherless families with mothers unable to provide adequate child supervision and discipline are common characteristics of broken families that contribute to criminal activity. Second, children who come from these broken families tend to have negative community experiences (peer effects) that further encourage their criminal participation. Ultimately, these conditions lead to the collapse of the local community, often leaving criminal youths living in high-crime neighborhoods. Each reinforces the other in a destructive relationship, spiraling downward into violence and social chaos. The results of empirical research rooted in social disorganization theory have offered robust evidence in support of the basic expectations of the theory over many decades (Kubrin and Weitzer, 2003).

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 to show the importance of broken families in the formation and persistence of ghettos.

4.1 The model

Let us return to the benchmark model without a costly policy. 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. This assumption is made to simplify the analysis and relaxing it does not alter the main results of this study. 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.

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,

$$u^k_n = P^{hk}_{n,t} V^h + P^{dk}_{n,t} V^d - c^k \frac{\left(\tau^k_{n,t}\right)^2}{2}. \tag{19}$$

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

$$U_n^i = \int_0^{\theta^i} \left[(1-p)\,\beta - p\sigma - K\,\mathbf{1}_{i=h} - \theta\right] d\theta + \int_{\theta^i}^1 w d\theta - \rho_{n,t}^i + \int_0^{\theta^i} \left[pu_n^S + (1-p)\,u_n^B\right] d\theta + \int_{\theta^i}^1 u_n^B d\theta,$$

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 is the expected utility of individual i, who does not know his or her θ, in terms of criminal behavior. If θ, the (inverse) ability to commit crime, is below θ_i, then individual i will commit crime. By contrast, if $\theta > \theta_i$, he or she will not commit crime and will instead receive a wage of w. In both cases, the individual will pay a land rent (or bid for a rent) of $\rho_{n,t}^i$ to reside in neighborhood n. The second part of this utility function corresponds to the expected utility from being a parent without knowing the type of family. If $\theta < \theta_i$, individual i will become a criminal and with a probability p will be arrested and thus give rise to a single-parent family. On the contrary, if not arrested, he or she will form a biparental family. If $\theta > \theta_i$, the individual will never be a criminal and, thus, will always form a biparental family.

In Online Appendix B, we determine $\rho_{n,t}^i$, the bid rent for a parent of type $i = h, d$ residing in neighborhood $n = 1, 2$ (see (B.3)).

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; Benabou, 1993), the urban equilibrium is defined as follows.

Definition 4 *At any date t and given Q_t, the urban configuration, characterized by $\rho_{1,t}^{i*}$, $q_{1,t}^*$, $\tau_{1,t}^{S*}$, $\tau_{1,t}^{B*}$, $\tau_{2,t}^{S*}$, $\tau_{2,t}^{B*}$, 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.*

In any urban equilibrium, both neighborhoods need to be equally attractive to each type$-i$ parent, for $i = h, d$. In Online Appendix B, we determine the bid rent differential,

$\Delta\rho \equiv \rho^h - \rho^d$ between honest and dishonest families, which is given by (B.5). Two urban equilibria may exist.

Definition 5

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

Since this is a key result, let us provide its intuition. 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 characterizes the long-run spatial pattern of crime and honesty culture in the urban area.

Proposition 9 *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 that 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.[31] 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 9, 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 equilibrium for which $q_1^* = q_2^* = 0$. The intuition is the same as in Proposition 1. In the steady state, all agents in both neighborhoods are endowed with a dishonest culture and

[31]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).

the crime rate is very high. Similar but opposite results are obtained in case (*iii*) of Proposition 9. 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 9). Consider first the case when $\underline{q} < \frac{1}{2}$, meaning that $Q_0 \in \left[\underline{q}, 1\right[$. Since $Q_0 < 1$, we have shown in Proposition 8 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.

This proposition 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. Indeed, in neighborhood 1, honest families have a higher chance of transmitting the honesty trait, which implies that individuals are less likely to be criminals and therefore families are more likely to be biparental. This fact, in turn, implies that the honesty trait is more likely to be transmitted to children born in neighborhood 1. The opposite is true in neighborhood 2. This result thus shows 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 9 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 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. Our microfoundation stems from the fact that families decide their socialization effort depending on the proportion of honest individuals residing in their neighborhood and cost of socialization, which itself depends on the structure of the family. Importantly, when this initial proportion of honest families is sufficiently low, because of the complementarity between the socialization effort and this proportion, parents with the honesty trait residing in neighborhood 1 increase their efforts and can bid away parents with the dishonesty trait to the other neighborhood. This leads to a segregated equilibrium at which crime and the structure of the family are spatially differentiated.

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.

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

Definition 6 (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.*

Our framework also allows us to discuss equity issues, where equity is defined as follows.

Definition 7 (Equity) *For a given Q_0, an equitable allocation of agents is the one that minimizes the difference in crime between the two neighborhoods.*

We have the following results.[32]

Proposition 10 *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 $q_{max} = \max\{\underline{q}, \frac{1}{2}\}$.*

[32]Given Proposition 9 and its proof, the proofs of Propositions 10 and 11 are straightforward and thus omitted.

(*i*) $\forall Q_0 \in \left[0, 2\underline{q}\right[\cup [2q_{max}, 2]$, *the market solution is always efficient. In particular, when* $Q_0 \in [\underline{q}, 2\underline{q}[$, *segregation is efficient but conflicts with equity. When* $Q_0 \in [0, \underline{q}[\cup]2q_{max}, 2[$, *there is spatial integration and efficiency matches with equity.*

(*ii*) $\forall Q_0 \in \left[2\underline{q}, 2q_{max}\right[$, *the market solution (i.e., segregation) is not efficient because the planner would choose integration (first best), i.e.,* $q^1 = q^2 = \overline{q}$. *Efficiency matches with equity.*

In case (*i*), for $Q_0 \in \left[0, \underline{q}\right[$ (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. For $Q_0 \in [\underline{q}, 2\underline{q}[$, 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 the crime rates across neighborhoods are not minimized and efficiency conflicts with equity because each neighborhood can only encompass a mass 1 of families and $Q_0 < min\left\{2\underline{q}, 1\right\}$.

In case (*ii*), when $Q_0 \in \left[2\underline{q}, 2q_{max}\right[$ is equivalent to $Q_0 \in \left[2\underline{q}, 1\right[$,[33] 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.

Figure 11 illustrates case (*i*) of Proposition 10. Each rectangle represents a neighborhood. The upper-panel rectangles show the spatial equilibrium in the city (case (*i*) of Proposition 9). As shown in Proposition 9, the unique steady-state equilibrium is such that the city is segregated with $q_1^* = \overline{q}$ and $q_2^* = 0$. The lower-panel rectangles depict the spatial distribution

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

of agents by type if the planner chooses to reallocate people. When such a policy is implemented, the outcome is worse than the equilibrium one since the proportion of honest agents falls below $\underline{q}$ in both neighborhoods. This leads to a long-run outcome where $q_1^* = q_2^* = 0$. As a result, it is better that the planner does not try to integrate agents in the urban space, implying that, in this case, *segregation is efficient.*

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-parent families (through a non-costly subsidy).[34] From Proposition 7, we can deduce that whenever $Q_0 > \underline{q}^\delta$, where $\underline{q}^\delta := \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$, this policy is efficient. Indeed, a sufficiently high subsidy (so that the cost of transmitting the honesty trait is the same for single and biparental families) reduces $\underline{q}$ to $\underline{q}^\delta < \underline{q}$ and increases the value of honesty in the steady state to $\overline{q}^\delta > \overline{q}$. We now compare the efficiency (in terms of a reduction in total crime) of a reallocation policy (as above) with a subsidy policy.

Proposition 11 *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 $q^\delta_{max} = \max\{\underline{q}^\delta, \frac{1}{2}\}$.*

(*i*) $\forall Q_0 \in \left[0, \underline{q}^\delta\right]$, *neither the subsidy policy nor the reallocation policy is efficient.*

(*ii*) $\forall Q_0 \in \left]\underline{q}^\delta, 2\underline{q}^\delta\right[\cup [2q^\delta_{max}, 2]$, *the subsidy policy is more efficient than the reallocation policy.*

(*iii*) $\forall Q_0 \in \left[2\underline{q}^\delta, 2q^\delta_{max}\right[$, *the reallocation policy is more efficient than the subsidy policy.*

When Q_0 is very low (case (*i*)), 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.

Part (*ii*) of the proposition uses Proposition 10, which shows that the reallocation policy is not efficient. Indeed, in that case, there is segregation and this is efficient because reallocating families would only lead to a worse outcome. On the contrary, a subsidy policy has a positive

[34] 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.

Figure 11: Spatial equilibrium under segregation where $q_1^* = \overline{q}$ and $q_2^* = 0$ (upper panel) and under integration where $q_1^* = q_2^* = 0$ (lower panel).

impact on honesty because it increases the proportion of honest agents in neighborhood 1 from $\overline{q}$ to $\overline{q}^{\delta}$.

Finally, when $Q_0 \in \left[2\underline{q}^{\delta}, 2q^{\delta}_{max}\right[$ (part (*iii*) of the proposition), the reallocation policy is more efficient than the subsidy policy. Indeed, consider that $\underline{q}^{\delta} < \frac{1}{2}$ so that $q^{\delta}_{max} = \frac{1}{2}$. Thus, $Q_0 \in \left[2\underline{q}^{\delta}, 1\right[$.[35] 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 than the subsidy policy.

Figure 12 illustrates this last result. In this figure, the upper-panel rectangles display the long-run spatial equilibrium when a subsidy policy is implemented. As stated above, the spatial equilibrium after the subsidy policy is implemented is segregated because $Q_0 < 1$. Compared with the long-run equilibrium without a policy, the only change is that the proportion of honest families in neighborhood 1 increases from $\overline{q}$ to $\overline{q}^{\delta}$. As a result, the subsidy policy is efficient because it reduces total crime.

The lower-panel rectangles in Figure 12 show the spatial distribution of agents when the planner chooses to optimally reallocate the agents to force integration. In such a case, given that $Q_0 > 2\underline{q}$, the proportion of honest families in both neighborhoods is higher than $\underline{q}$, meaning that it converges to $\bar{q}$. Since $2\overline{q} > \overline{q}^{\delta}$ (because $2\overline{q} > 1$), the reallocation policy is more efficient than the subsidy policy because, in the latter, the proportion of honest families is $\overline{q}^{\delta}$ compared with $2\overline{q}$ in the former.

5 Concluding remarks and policy implications

Policymakers are beginning to recognize 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

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

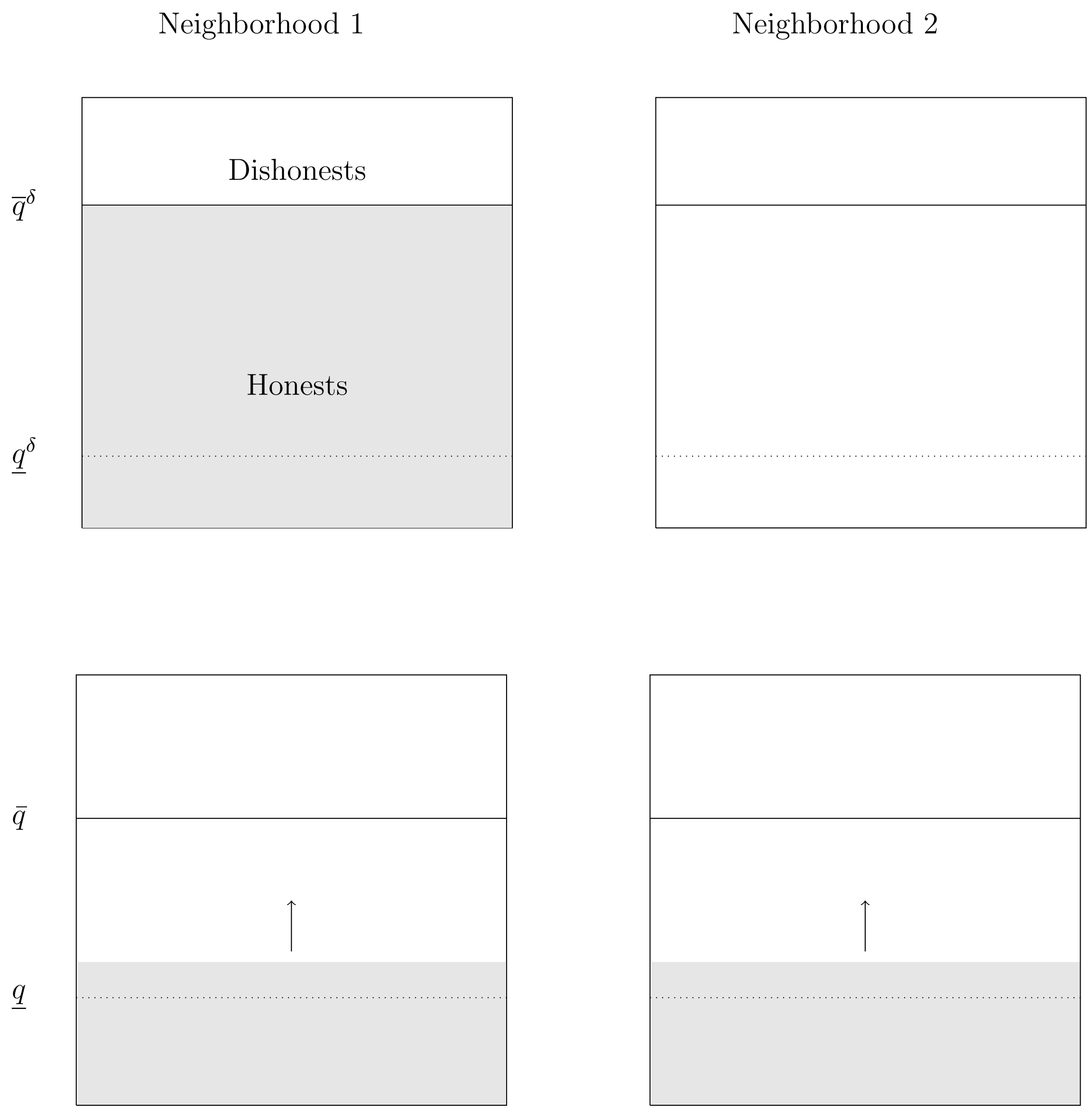

Figure 12: Spatial equilibrium under a subsidy policy (upper panel) and under a reallocation policy (lower panel) when $Q_0 \in \left[2\underline{q}, 2q_{max}\right[$.

dependence themselves in later years. These children thus face a daunting array of problems.

While this link between illegitimacy and chronic welfare dependency is now better understood, policymakers also need to appreciate another strong and disturbing pattern evident in scholarly studies: the links between illegitimacy and crime and between the lack of parental attachment and crime. 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. Indeed, without that knowledge, sound policymaking is impossible.

To capture these issues, we developed a dynamic model of the cultural transmission of crime ethics in a framework where parents (vertical transmission) and peers (horizontal transmission) affect the evolution over time of individuals with the honesty trait. The new aspect of our model is to include and endogenize the structure of the family in this transmission process by differentiating between monoparental (or single-mother) families and biparental families. In an overlapping generations model, we studied the criminal behavior of individuals who live for two periods and determined the long-run crime rate in the population. Because of time constraints (and other costs such as the forgone income from the father), it is costlier for single mothers to transmit the honesty trait to their offspring and, therefore, the latter are more likely to become criminals since less honesty leads to more crime.

We first show that the probability of being arrested (incarceration rate) is crucial to determining the long-run crime rate in the economy. In particular, under some conditions, we show that increasing the incarceration rate can backfire by raising rather than decreasing long-run crime rates. Indeed, increasing incarceration has two main effects. First, there is the standard *deterrence effect* for which an increase in the probability of being caught reduces crime. Second, and this is new, there is the *social disorganization effect* for which an increase in incarceration disrupts the family structure by raising the number of single-mother families in the population, which has a negative impact on the transmission of the honesty trait. If the second effect is sufficiently strong, then more incarceration leads to more crime. This is even truer if incarceration is costly since the higher spending on crime, the less the government spends on education, which reduces wages and thus increases the decision to become a criminal.

We also consider another policy that subsidizes the socialization cost of single-mother families. We show that there is a non-monotonic relationship between the subsidy rate and long-run crime, suggesting that increasing the subsidy can, in fact, increase rather than decrease long-run crime.

We extended our model to include two neighborhoods in which agents with different honesty traits freely decide where to live to explain the emergence and persistence of urban ghettos characterized by a large proportion of single-mother families and high crime rates. We also showed that segregation (i.e., the development of urban ghettos) can be efficient from a welfare viewpoint but not equitable.

There is a widespread belief that race is a major explanatory cause of crime. This belief is anchored in the large disparity in crime rates between whites and blacks.[36] However, in this study, we argue that the real variable is not race but family structure. This is supported by the data. Indeed, the incidence of broken families is much higher in the black community. Smith and Jarjoura (1988), in a major study of 11,000 individuals, find that "the percentage of single-parent households with children between the ages of 12 and 20 is significantly associated with rates of violent crime and burglary." The same study makes it clear that the popular assumption of an association between race and crime is false. Illegitimacy is the key factor. In other words, the absence of marriage and failure to form and maintain intact families explain the high crime rates in neighborhoods among whites as well as blacks.

In terms of policy implications, unless we try to understand the importance of stable families to reducing crime, particularly youth crime, we risk making young people the target. It is the child who grows up in a broken home with an absent father involved in crime who is most likely to commit crime him- or herself and become an absent parent him- or herself. Unchecked, the cycle looks set to continue and to multiply in its effects. Simply threatening to lock young people up will not break the cycle. Although criminals need to face penalties for their actions, we desperately need to deal with the reasons why they are committing crime in the first place.

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, although reductions in stealing

[36] Approximately 12-13% of the American population is African-American, but they make up 37% of prison inmates of the 2.2 million male inmates as of 2014 (U.S. Department of Justice, 2014).

and other delinquent activities have in some instances proved short-lived. We also believe that family preservation programs should be funded. This is because the family may be able to resist the deleterious effects of social disorganization on their children, and since strong families may also work together to reduce social disorganization in their communities. Family preservation programs are short-term, intensive, empowerment model programs that focus not on an individual client but rather on the needs of the entire family. For example, Olds et al. (1998) investigate the effects of a home visiting program on pregnant women. The home visitors (nurses) gave the women advice about childrearing, infant development, nutrition, and the need to avoid alcohol and drugs. A 15-year follow-up showed that the children of visited mothers were arrested at a significantly (54%) lower rate than the children of non-visited mothers.[37]

Our model also suggests that 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.[38] Examples include neighborhood regeneration policies implemented in the United States and Europe through Enterprise Zone Programs (Papke, 1994; Boarnet and Bogart, 1996; Bondonio and Engberg, 2000; Bondonio and Greenbaum, 2007; Givord et al., 2013; Briant et al., 2015; Mayneris et al., 2017) and Empowerment Zone Programs (Busso et al., 2013).[39] 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. Such an approach would also simultaneously strengthen residential ties and interconnections within neighborhoods.

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

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

[39] The Enterprise Zone Program designates a specific urban (or rural) area that is depressed and targets it for economic development through government-provided subsidies to labor and capital. It aims to revitalize distressed urban communities and represents a nexus between social welfare policy and economic development efforts.

References

[1] Ballester, C., Calvó-Armengol, A. and Y. Zenou (2006), "Who's who in networks. Wanted: The key player," *Econometrica* 74, 1403-1417.

[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.

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

[8] Bernasco, W., de Graaff, T., Rouwendal, J. and W. Steenbeek (2017), "Social interactions and crime revisited: An investigation using individual offender data in Dutch neighborhoods," *Review of Economics and Statistics* 99, 622-636.

[9] 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.

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

[11] Bisin, A. and T. Verdier (2011), "The economics of cultural transmission and socialization," In J. Benhabib, A. Bisin, and M.O. Jackson (Eds.), *Handbook of Social Economics*, Amsterdam: Elsevier Science, pp. 339-416.

[12] Boarnet, M.G. and W.T. Bogart (1996), "Enterprise zones and employment: Evidence from New Jersey," *Journal of Urban Economics* 40, 198-215.

[13] Bondonio, D. and J. Engberg (2000), "Enterprise zones and local employment: Evidence from the states? program," *Regional Science and Urban Economics* 30, 519-549.

[14] Bondonio, D. and R.T. Greenbaum (2007), "Do local tax incentives affect economic growth? What mean impact miss in the analysis of enterprise zone policies," *Regional Science and Urban Economics* 37, 121-136.

[15] Boyd, R. and P.J. Richerson (1985), *Culture and the Evolutionary Process*, Chicago: University of Chicago Press.

[16] Bhuller, M., Dahl, G.B., Løken, K.V. and M. Mogstad (2018), "Intergenerational effects of incarceration," NBER Working Paper No. 24227.

[17] Briant, A., Lafourcade, M. and B. Schmutz (2015), "Can tax breaks beat geography? Lessons from the French EZ experience," *American Economic Journal: Economic Policy* 7, 88-124.

[18] Bumpass, L.L., Raley, R.K. and J.A. Sweet (1995), "The changing character of step-families: Implications of cohabitation and nonmarital childbearing," *Demography* 32, 425-436.

[19] Bureau of the Census (1994), "Marriage and family statistics," Washington, DC: U.S. Department of Commerce.

[20] Bursik, R.J., Jr. and H.G. Grasmick (1993), *Neighborhoods and Crime: The Dimensions of Effective Community Control*, Lanham, MD: Lexington Books.

[21] Busso, M., Gregory, J. and P. Kline (2013), "Assessing the incidence and efficiency of a prominent place based policy," *American Economic Review* 103, 897-947.

[22] 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.

[23] Case, A. and L.F. Katz (1991), "The company you keep: The effects of family and neighborhood on disadvantaged youths," NBER Working Paper No. 3705.

[24] Cavalli-Sforza, L. and M. Feldman (1981), *Cultural Transmission and Evolution. A Quantitative Approach*, Princeton: Princeton University Press.

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

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

[27] Corno, L. (2017), "Homelessness and crime: Do your friends matter?" *Economic Journal* 127, 959-995.

[28] 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.

[29] 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.

[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] Duncan, G., Kalil, A., Mayer, S., Tepper, R. and M. Payne (2005), "The apple does not fall far from the tree," In: Bowles, S., Gintis, H., Osborne Groves, M. (Eds.), *Unequal Chances: Family Background and Economic Success*, Chapter 1, Princeton: Princeton University Press.

[33] 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.

[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] Garoupa, N. (1997), "The theory of optimal law enforcement," *Journal of Economic Surveys* 11, 267–295.

[39] Geller, A. (2013), "Paternal incarceration and father-child contact in fragile families," *Journal of Marriage and Family* 75(5), 1288-1303.

[40] Geller, A., Cooper, C.E., Garfinkel, I., Schwartz-Soicher, O. and R.B. Mincy (2012), "Beyond absenteeism: Father incarceration and child development," *Demography* 49(1), 49-76.

[41] Geller, A., Garfinkel, I. and B. Western (2011), "Paternal incarceration and support for children in fragile families," *Demography* 48, 25-47.

[42] Givord, P., Rathelot, R. and P. Sillard (2013), "Place-based tax exemptions and displacement effects: An evaluation of the Zones Franches Urbaines program," *Regional Science and Urban Economics* 43, 151-163.

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

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

[45] Glaze, L.E. and L.M. Maruschak (2010), "Parents in prison and their minor children," Bureau of Justice Statistics Special Report.

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

[47] Grogger, J. and M. Willis (2000), "The emergence of crack cocaine and the rise in urban crime rates," *Review of Economics and Statistics* 82, 519-529.

[48] Hall, S.G. (1904), *Adolescence: Its Psychology and Its Relations to Physiology, Anthropology, Sociology, Sex, Crime, Religion, and Education*, New York: Appleton.

[49] Hjalmarsson, R. and M.J. Lindquist (2010), "Driving under the influence of our fathers," *The B.E. Journal of Economic Analysis & Policy* 10, 1-15.

[50] Hjalmarsson, R. and M.J. Lindquist (2012), "Like godfather, like son: Exploring the intergenerational nature of crime," *Journal of Human Resources* 47, 550-582.

[51] Hjalmarsson, R. and M.J. Lindquist (2013), "The origins of intergenerational associations in crime: Lessons from Swedish adoption data," *Labour Economics* 20, 68-81.

[52] Kellam, S.G., Ensminger, M.E. and R.J. Turner (1977), "Family structure and the mental health of children," *Archives of General Psychiatry* 34, 1012-22.

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

[54] 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.

[55] Kubrin, C.E. and R. Weitzer (2003), "New directions in social disorganization research," *Journal of Research in Crime and Delinquency* 40, 374-402.

[56] Liu, X., Patacchini, E., Zenou, Y. and L.-F. Lee (2012), "Criminal networks: Who is the key player?" CEPR Discussion Paper No. 8772.

[57] 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.

[58] 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.

[59] Mayneris, F., Mayer, T. and L. Py (2017), "The impact of urban Enterprise Zones on establishments' location decisions: Evidence from French ZFUs," *Journal of Economic Geography* 17, 709-752

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

[61] 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.

[62] 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.

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

[64] Papke, L. (1994), "Tax policy and urban development: Evidence from the Indiana enterprise zone program," *Journal of Public Economics* 54, 37-49.

[65] Patacchini, E. and Y. Zenou (2012), "Juvenile delinquency and conformism," *Journal of Law, Economic, and Organization* 28, 1-31.

[66] Philbrick, K., Ayre, L. and H. Lynn (2014), *Children of Imprisoned Parents: European Perspectives on Good Practice*, 2nd Edition.

[67] Polinsky, A.M. and S. Shavell (2000), "The economic theory of public enforcement of law," *Journal of Economic Literature* 38, 45-?76.

[68] 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.

[69] Reeves, R.V. and K. Howard (2013), "The parenting gap," Center on Children and Families, Brookings Institution.

[70] Richters, J.E. and P. Martinez (1993), "The NIMH community violence project: I. Children as victims and witnesses to violence," *Psychiatry* 56, 7-21.

[71] Rowe, D.C. and D.P. Farrington (1997), "The familial transmission of criminal convictions," *Criminology* 35, 177-201.

[72] Sampson, R.J. (2012), *Great American City: Chicago and the Enduring Neighborhood Effect*, Chicago, IL: University of Chicago Press.

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

[74] Smith, D.A. and G.R. Jarjoura (1988), "Social structure and criminal victimization," *Journal of Research in Crime and Delinquency* 25, 27-52.

[75] South, S.J. and K.D. Crowder (1997), "Residential mobility between cities and suburbs: Race, suburbanization, and back-to-the-city moves," *Demography* 34, 525-538.

[76] 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.

[77] Stevenson, M. (2017), "Breaking bad: Mechanisms of social influence and the path to criminality in juvenile jails," *Review of Economics and Statistics* 99, 824-838.

[78] Swisher, R. and M.R. Waller (2008), "Confining fatherhood: Incarceration and paternal involvement among nonresident White, African-American, and Latino fathers," *Journal of Family Issues* 29, 1067-1088.

[79] Tach, L., Mincy, R.B. and K. Edin (2010), "Parenting as a "package deal": Relationships, fertility, and nonresident father involvement among unmarried parents," *Demography* 47, 181-204.

[80] 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.

[81] Thornberry, T.P. (2009), "The apple doesn't fall far from the tree (or does it?): Intergenerational patterns of anti-social behavior," *Criminology* 47, 297-325.

[82] U.S. Census Bureau (2010), "Current Population Survey 2009," Annual Social and Economic Supplement, Release date January 2010, Available at: www.census.gov.

[83] U.S. Department of Justice (2014), Bureau of Justice Statistics, Annual Survey of Jails, midyear 2000 and midyear 2006?2014.

[84] Ventura, S.J. (2009), "Changing patterns of nonmarital childbearing in the United States," NCHS Data Brief No. 18, May, Available at: www.ndc.gov.

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

[86] Wildeman, C. (2010), "Paternal incarceration and children's physically aggressive behaviors: Evidence from the Fragile Families and Child Wellbeing Study," *Social Forces* 89, 285-309.

[87] 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.

[88] Williams, J. and R. Sickles (2002), "An analysis of the crime as work model: Evidence from the 1958 Philadelphia Birth Cohort Study," *Journal of Human Resources* 37, 479-509.

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

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

Online Appendix
Crime, Broken Families, and Punishment

By Emeline Bezin[1], Thierry Verdier[2] and Yves Zenou[3]

A Proofs

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

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

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

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

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

We have

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

[1]Paris School of Economics and IRES (UCL). France. E-mail: emeline.bezin@gmail.com.
[2]Paris School of Economics and Ecole des Ponts ParisTech, PUC-Rio and CEPR. Email: thierry.verdier@ens.fr.
[3]Monash University and CEPR. Email: yves.zenou@monash.edu.

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 $h(q_t)$ 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 A1 and A2 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

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

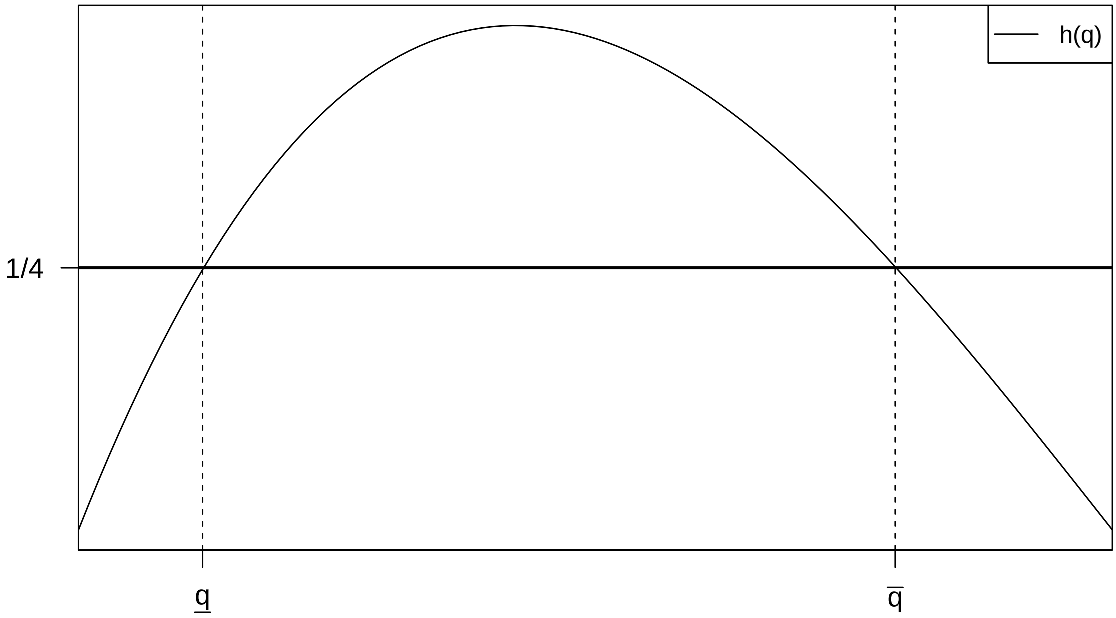

Figure A1: Case $h(q_m) > \frac{1}{4}$.

Let us differentiate u with respect to p. We obtain

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

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

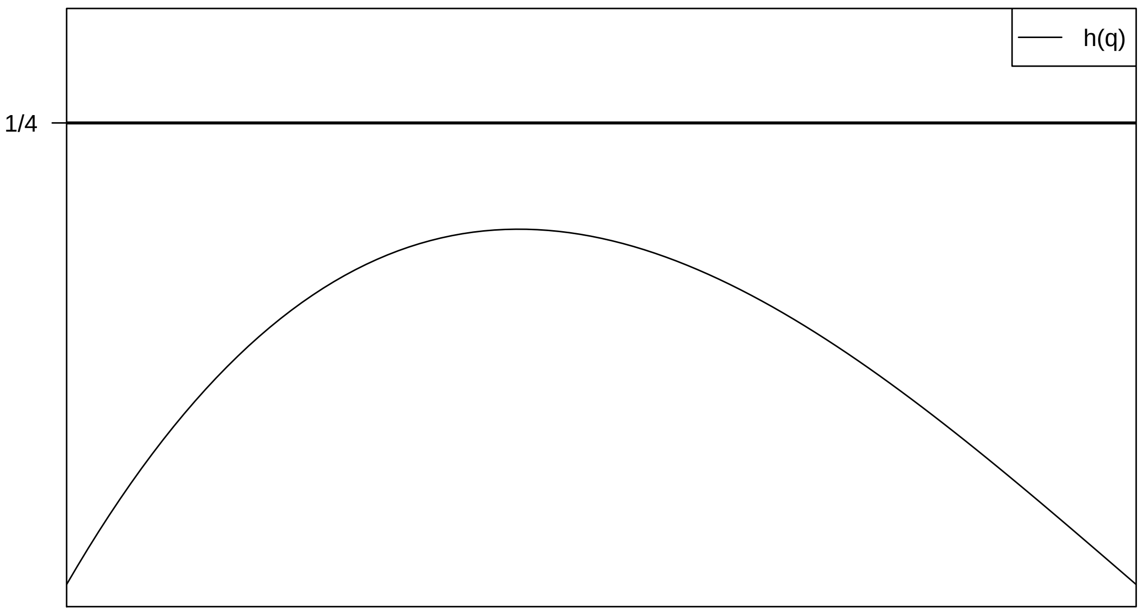

Figure A2: Case $h(q_m) < \frac{1}{4}$.

study this condition. We have

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

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. Furthermore, we know that at $p = (\beta - w - K) / \left[2\left(\beta+\sigma\right)\right]$, $\widetilde{q} = 1$; hence, $q_m(p) < \widetilde{q}$ and the polynomial is negative. Further, at $p = (\beta - w) / \left[2\left(\beta+\sigma\right)\right]$, $\widetilde{q} = 0$, 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

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

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

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

At $p = 0$, we have $h\left(\widetilde{q}(0); 0\right) = \frac{1}{4}\Delta^B$, 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 A3.

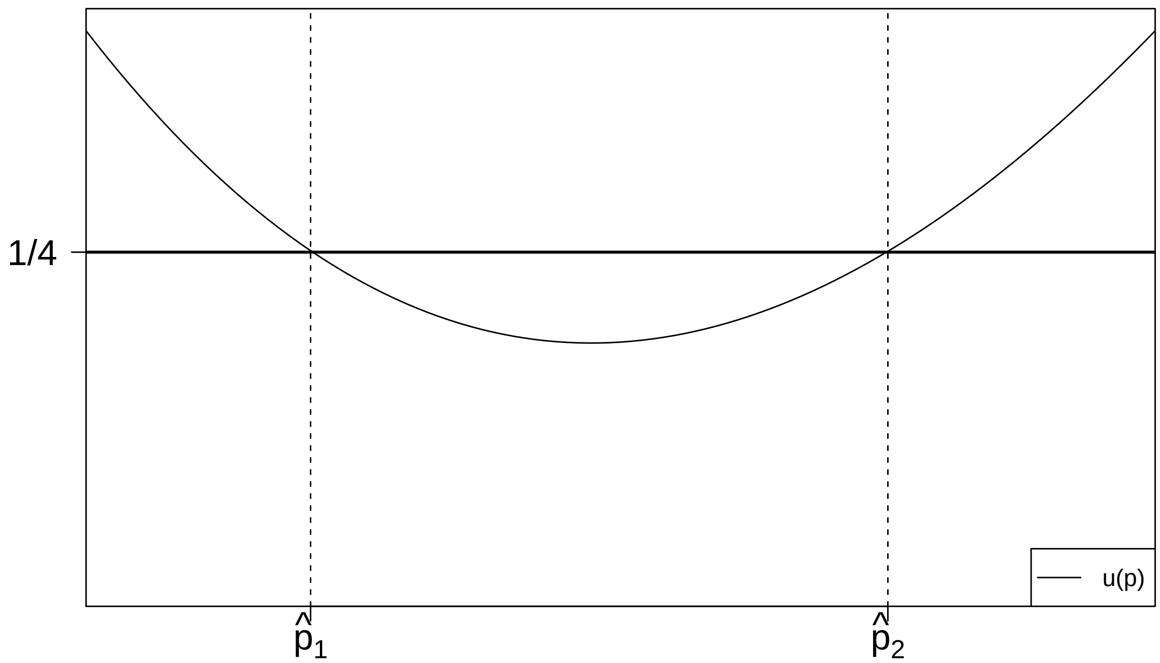

Figure A3: The function $u \equiv h(q_m(p), p)$.

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 $\overline{q}$. ∎

Proof of Proposition 2: Let us focus on the interior equilibrium $\overline{q}$ of Proposition 1 and, therefore, assume that $p \in [0, \widehat{p}_1] \cup [\widehat{p}_2, \overline{p}_2]$. In the proof of Proposition 1, we showed that the steady-state proportion of honest individuals in the population, $\overline{q}$, is implicitly given by

$$f(\overline{q}) = 0,$$

where $f(.)$ is defined by (A.1). Hence,

$$\frac{\partial \overline{q}}{\partial w} = -\frac{\partial f(.)/\partial w}{\partial f(.)/\partial q}.$$

In the proof of Proposition 1, we showed that $\partial f(.)/\partial q < 0$, which means that the sign of $d\overline{q}/dw$ is the same as the sign of $\partial f(.)/\partial w$. Totally differentiating $f(.)$ in (A.1) and using (4) lead to

$$\frac{\partial f(q_t)}{\partial w} |_{q_t = \overline{q}} = \overline{q}(1 - \overline{q}) 4\overline{q} \,(1 - \overline{q})\, p \left(\Delta^B - \Delta^S\right) > 0.$$

As a result,

$$\frac{\partial \overline{q}}{\partial w} > 0.$$

By using the same approach, we can show that $\frac{\partial \overline{q}}{\partial \beta} < 0$ and $\frac{\partial \overline{q}}{\partial \sigma} > 0$, $\frac{\partial \overline{q}}{\partial c^S} < 0$, $\frac{\partial \overline{q}}{\partial c^B} < 0$. ∎

Proof of Proposition 4: Following the proof of Proposition 1, the stationary equilibria of the economy are 0, 1, and q such that

$$h^*(q_t) = \frac{1}{4},$$

where $h^* : [0,1] \to [0,1]$ is given by

$$h^*(q_t) = q_t(1-q_t)\left[\Delta^B - pC(q_t)\left(\Delta^B - \Delta^S\right)\right],$$
$$\text{with} \quad C(q_t) = -q_t K + \beta - p(\beta+\sigma) - w(R-p).$$

Step 1. From the proof of Proposition 1, we deduce that h^* admits a unique maximum q^*_m.

Step 2. Define

$$u^*(p) \equiv h^*(q^*_m(p), p).$$

We show that for any $R < \frac{-K+\beta}{2(\beta+\sigma)}$, $\frac{du^*}{dp} < 0$.

Let us perform the derivative of the function u^*,

$$\begin{aligned}\frac{du^*}{dp} &= -\left(\Delta^B - \Delta^S\right)(-q_t K + \beta - 2p(\beta+\sigma) - w(R-p)) - pw'(R-p) \\ &< -\left(\Delta^B - \Delta^S\right)(-K + \beta - 2p(\beta+\sigma) - w(R-p)) \\ &< -\left(\Delta^B - \Delta^S\right)(-K + \beta - 2R(\beta+\sigma)).\end{aligned}$$

Note that $\Xi(p) = -(-K+\beta-2p(\beta+\sigma)-w(R-p))$ is increasing in p. Indeed, $\Xi'(p) = 2(\beta+\sigma) - w'(R-p)$ and $\Xi''(p) = w''(R-p) < 0$. Hence, $\Xi'(p)$ is decreasing in p and positive as $\Xi'(R) = 2(\beta+\sigma) - w'(0) > (\beta+\sigma) - w'(0) > 0$ from Assumption 2. Hence, $\Xi(p)$ is increasing in p and

$$\begin{aligned}\frac{du^*}{dp} &< -\left(\Delta^B - \Delta^S\right)(-K+\beta-2p(\beta+\sigma)-w(R-p)) = \left(\Delta^B - \Delta^S\right)\Xi(p) \\ &< \left(\Delta^B - \Delta^S\right)\Xi(R) = -\left(\Delta^B - \Delta^S\right)(-K+\beta-2R(\beta+\sigma)).\end{aligned}$$

Consequently, for any $R < \frac{-K+\beta}{2(\beta+\sigma)}$, $\frac{du^*}{dp} < 0$.

Step 3. We show that there exists $\tilde{R}$ such that
- for any $R \in [0, \tilde{R}]$, $u^*(p) > \frac{1}{4}$ $\forall p \in [0, R]$ (equivalent to $h^*(q_t) = \frac{1}{4}$ admits two solutions $\underline{q}(p)$ and $\bar{q}(p)$),
- for any $R \in]\tilde{R}, \frac{-K+\beta}{2(\beta+\sigma)}]$, there exists $\hat{p}^{1*}$ such that $u^*(p) > \frac{1}{4}$ $\forall p \in [0, \hat{p}^{1*}]$ (equivalent to

$h^*(q_t) = \frac{1}{4}$ admits two solutions), and $u^*(p) < \frac{1}{4}\ \forall p \in]\hat{p}^{1*}, R]$ (equivalent to $h^*(q_t) = \frac{1}{4}$ admits no solution).

First, we study $u^*(p)$ at $p = R$. Define

$$v(R) \equiv u^*(R) = q_m^*(R)(1-q_m^*(R))\left[\Delta^B - R(-q_m^*(R)K + \beta - R(\beta+\sigma))\left(\Delta^B - \Delta^S\right)\right],$$
$$\text{and}\quad v^1(R) \equiv q_m^*(R)(1-q_m^*(R))\left[\Delta^B - R(-K + \beta - R(\beta+\sigma))\left(\Delta^B - \Delta^S\right)\right],$$

with $v^1(R) > v(R)\ \forall R$.

Note that $v^1(\frac{-K+\beta}{2(\beta+\sigma)}) < 1/4$. Indeed, since

$$v^1(R) < \frac{1}{4}\left[\Delta^B - R(-K+\beta - R(\beta+\sigma))\left(\Delta^B - \Delta^S\right)\right],\quad \forall R,$$

we have

$$v^1(\frac{-K+\beta}{2(\beta+\sigma)}) < \frac{1}{4}\left[\Delta^B - \frac{(-K+\beta)^2}{4(\beta+\sigma)}\left(\Delta^B - \Delta^S\right)\right] < \frac{1}{4},$$

where the latter inequality comes from Assumption 2. We deduce

$$v(\frac{-K+\beta}{2(\beta+\sigma)}) < v^1(\frac{-K+\beta}{2(\beta+\sigma)}) < \frac{1}{4}.$$

Following Step 2, we have $\frac{\partial v}{\partial R} < 0$. Further, we have $v(0) = q_m^*(0)(1-q_m^*(0))\Delta^B = (1/4)\Delta^B > 1/4$ (recall that from Assumption 1, $\Delta^B > 1$). Hence, we deduce that there exists $\tilde{R}$ such that

$$(i)\ v(R) > \frac{1}{4}\quad \forall R < \tilde{R},$$
$$(ii)\ v(R) \leq \frac{1}{4}\quad \forall R \in [\tilde{R}, \frac{-K+\beta}{2(\beta+\sigma)}].$$

Returning to the function $u^*(p)$. We have $u^*(0) > 1/4$, $\partial u^*/\partial p < 0$. When item (i) holds, that is when $R < \tilde{R}$, $u^*(R) > 1/4$. We deduce that $u^*(p) > 1/4\ \ \forall p \in [0, R]$. When item (ii) holds, that is when $R \in [\tilde{R}, \frac{-K+\beta}{2(\beta+\sigma)}]$, we have $u^*(R) < 1/4$. We deduce that there exists $\hat{p}^{1*} \geq \tilde{R}$ such that $u^*(p) \geq \frac{1}{4}\ \forall p \in [0, \hat{p}^{1*}]$, and $u^*(p) < \frac{1}{4}\ \forall p \in]\hat{p}^{1*}, R]$.

The function u^* is represented in Figures A4 (for $R = \tilde{R}$) and A5 (for $R > \tilde{R}$).

Step 4. We show that $\forall q_0 < \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$, q_t converges to 0. Indeed, $u^*(p) < u^*(0)$

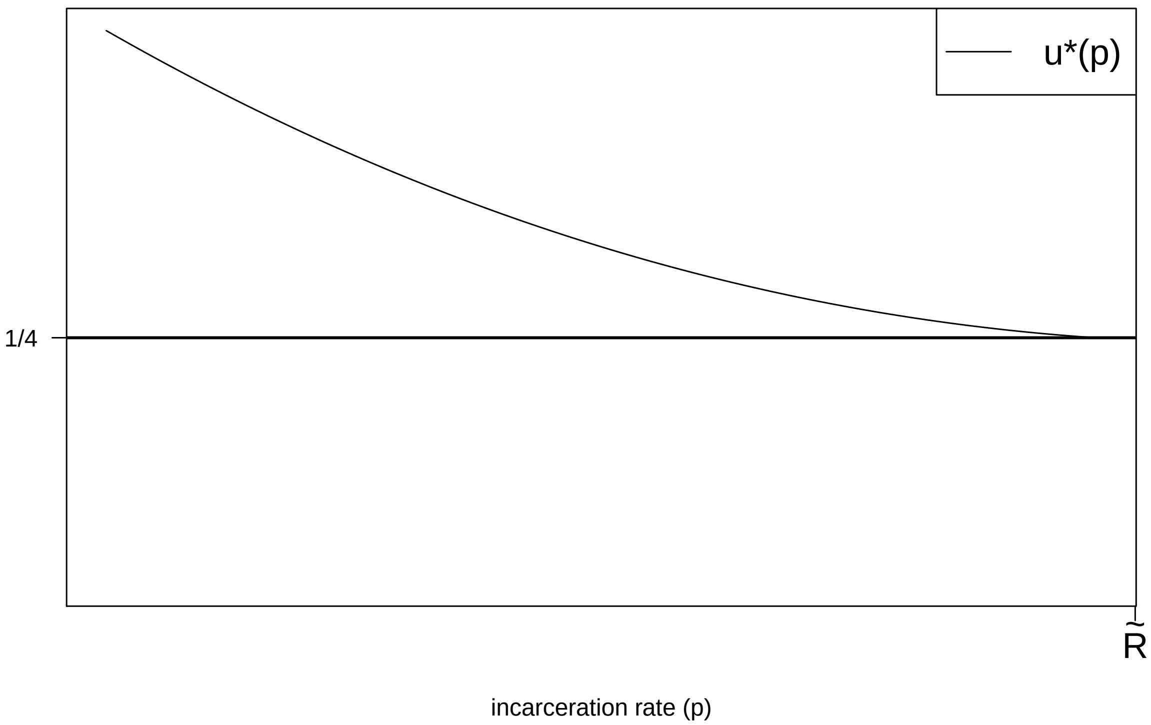

Figure A4: The function u when $R = \tilde{R}$. Whenever $R < \tilde{R}$, then $u(p) > \frac{1}{4}$.

$\forall p$ implies $\underline{q}(p) > \underline{q}(0)$. We know that $\forall q_0 < \underline{q}(p)$, the sequence q_t converges to 0. We deduce that $\forall q_0 < \underline{q}(0) = \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B}$, the sequence q_t converges to 0.

Step 5. We then show that

(i) When $R < \tilde{R}$, $\forall q_0 > \bar{q}(\tilde{R})$, the population converges to $\bar{q}(p)$, with $\partial \bar{q}/\partial p < 0$.

(ii) When $R \in [\tilde{R}, \frac{-K+\beta}{2(\beta+\sigma)}]$, $\forall q_0 > \underline{q}(\hat{p}^{1*})$, the population converges to $\bar{q}(p)$, with $\partial \bar{q}/\partial p < 0$ (if $p \leq \hat{p}^{1*}$), or to zero (if $p > \hat{p}^{1*}$).

First, from proof of Proposition 1, we know that when $h^*(q)$ admits two solutions $\underline{q}(p)$ and $\bar{q}(p)$, for any $q_0 > \underline{q}(p)$, the sequence q_t converges to $\bar{q}(p)$.

Hence, when $R < \tilde{R}$, $\forall p$, for any $q_0 > \underline{q}^*(p)$, the sequence q_t converges to $\bar{q}(p)$. When $R \in [\tilde{R}, \frac{-K+\beta}{2(\beta+\sigma)}]$, $\forall p < \hat{p}^{1*}$, for any $q_0 > \underline{q}^*(p)$, the sequence q_t converges to $\bar{q}^*(p)$ and $\forall p \in \hat{p}^{1*}, R]$, for any $q_0 \in [0, 1]$, the sequence q_t converges to 0.

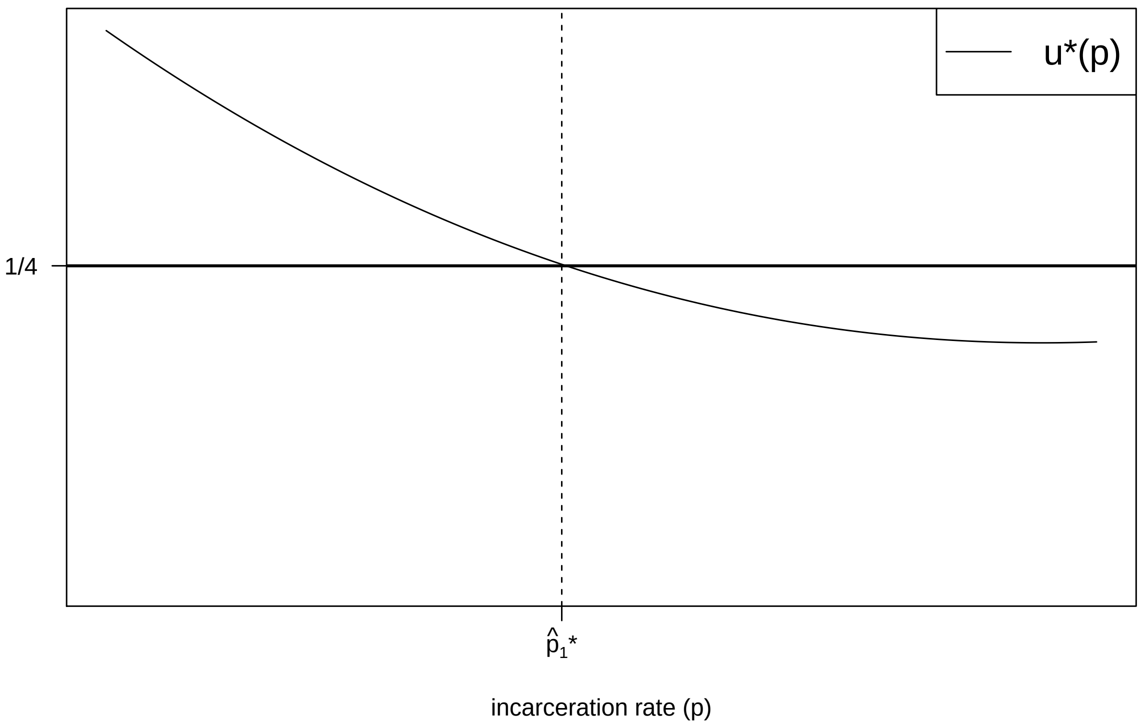

Figure A5: The function u when $R > \tilde{R}$. Here, we have $u(p) > \frac{1}{4} \Leftrightarrow p > \hat{p}_1^*$.

The convergence condition above depends on p. We must find a condition that holds for any p to compare convergence under different incarceration rates. To do so, we find an upper bound for $\underline{q}^*(p)$. When $R < \tilde{R}$, the upper bound for $\underline{q}^*(p)$ is $\underline{q}(\tilde{R})$ (this can be shown by using the fact that $\partial u^*/\partial p < 0$ and $u^*(\tilde{R}) = 1/4$). When $R \in [\tilde{R}, \frac{-K+\beta}{2(\beta+\sigma)}]$, the upper bound is $\underline{q}(\hat{p}^{1*})$ (this is shown by using similar arguments).

We deduce that when $R < \tilde{R}$, and $\forall q_0 > \underline{q}(\tilde{R})$, the sequence q_t converges to $\bar{q}(p)$. Then, we easily show that $\partial \bar{q}/\partial p < 0$ (i.e., the higher the incarceration rate p, the lower is the steady-state level of honesty).

When $R \in [\tilde{R}, \frac{-K+\beta}{2(\beta+\sigma)}]$, and $\forall q_0 > \underline{q}(\hat{p}^{1*})$, the sequence q_t converges to $\bar{q}(p)$, with $\partial \bar{q}/\partial p < 0$ for $p < \hat{p}^{1*}$ and to zero for p higher than $\hat{p}^{1*}$. Again, an increase in incarceration has a negative impact on long-run honesty.

Figures A6 and A7 depict cases (i) and (ii) of Step 5. Figure A6 shows that when $R < \tilde{R}$, for any $q_0 > \underline{q}(\tilde{R})$, $\forall p$, the sequence q_t converges to $\bar{q}(p)$. Figure A7 shows that for

$R \in [\tilde{R}, \frac{-K+\beta}{2(\beta+\sigma)}]$, $\forall q_0 > \underline{q}(\hat{p}^{1*})$, the sequence q_t converges to $\bar{q}(p)$.

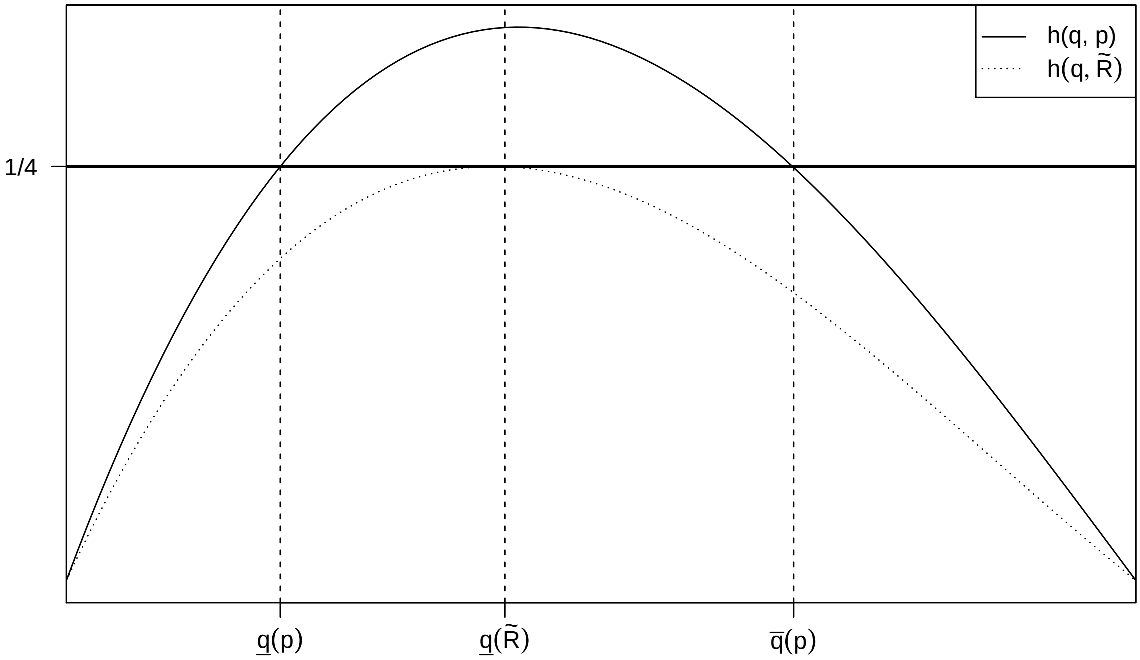

Figure A6: Case $R = \tilde{R}$: functions $h^*(q,p)$ and $h^*(q,\tilde{R})$, with $h^*(q,p) > h^*(q,\tilde{R})$.

Finally, part (i) of Proposition 4 follows from Step 4, posing $q^1_{\min} = \underline{q}(0) = \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$.

Part (ii) of Proposition 4 follows from Step 5, posing $q^2_{\min} = \underline{q}(\tilde{R})$ for $R < \tilde{R}$, and $q^2_{\min} = \underline{q}(\hat{p}^{1*})$ for $R \in [\tilde{R}, \frac{-K+\beta}{2(\beta+\sigma)}]$.

Proof of Proposition 5: In this proof, we compare the long-run crime rate (denoted by $\bar{C}$) for $p = R$ with the long-run crime rate for $p < R$.

To show case (i), note first that $\forall q_0 < q^1_{\min} = \underline{q}(0) = \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$, q_t converges to 0. We deduce that for any incarceration policy, the crime rate is given by $\bar{C} = \beta - w(R-p) - p(\beta+\sigma)$,

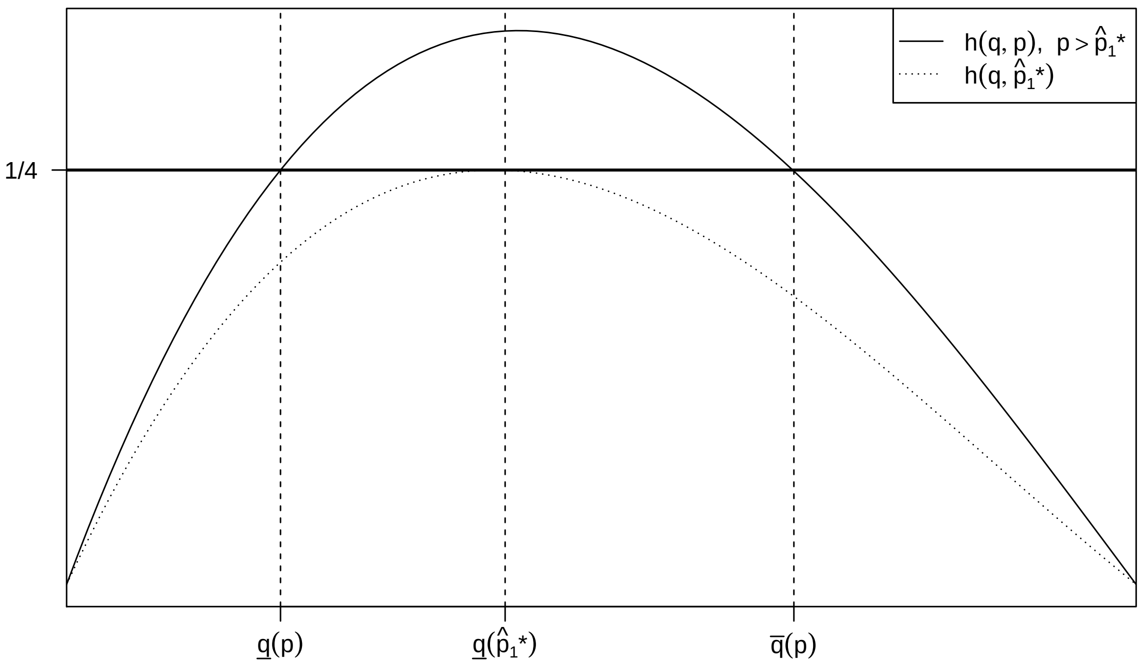

Figure A7: Case $R > \tilde{R}$: functions $h^*(q,p)$ and $h^*(q,\hat{p}_1^*)$, with $h^*(q,p) > h^*(q,\hat{p}_1^*)$.

which is a decreasing function of p. Hence, repressive policies are those where $p = R$ minimizes long-run crime.

(ii) We know from the proof of Proposition 4 that when $R \in [\tilde{R}, \frac{-K+\beta}{2(\beta+\sigma)}]$ and $q_0 > q^2_{\min} = \underline{q}(\hat{p}^{1*})$, for $p = R$, q_t converges to 0, while for $p < \hat{p}^{1*}$, q_t converges to $\bar{q}(p) > 0$. In particular, for $p = 0$, q_t converges to $\bar{q}(0) = \frac{\Delta^B + \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B}$.

We deduce that at $p = R$, the long-run crime rate $\bar{C}$ is given by $\bar{C}(R) = \beta - R(\beta+\sigma)$. At $p = 0$, we have $\bar{C}(0) = -\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B} K + \beta - w(R)$.
A repressive policy $p = R$ does not minimize $\bar{C}$ and is dominated by continuity by a

permissive policy $p \leq \overline{p} < R$ if $\bar{C}(0) < \bar{C}(R)$ or

$$- \frac{\Delta^B + \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B} K + \beta - w(R) < \beta - R(\beta + \sigma),$$
$$\Leftrightarrow - \frac{\Delta^B + \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B} K - w(R) + R(\beta + \sigma) < 0.$$

Since $-w(R) + R(\beta + \sigma)$ is increasing in R by assumption, the above inequality holds if

$$-\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B} K - w(\frac{-K + \beta}{2(\beta + \sigma)}) + \frac{\beta - K}{2} < 0$$

or

$$\frac{\beta - K}{2} < \frac{\Delta^B + \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B} K + w(\frac{-K + \beta}{2(\beta + \sigma)}),$$

which is condition (13) of Proposition 5

Proof of Proposition 6:

Let us first state the following lemma.

Lemma 1 *There exist $\widehat{p}^1(R)$ and $\widehat{p}^2(R)$ such that:*

(i) For any $p \in]\widehat{p}^1(R), \min\{\widehat{p}^2(R), 1\}[$, $\forall q_0 \in [0, 1]$, the sequence q_t converges to zero.

(ii) For any $p \leq \widehat{p}^1(R)]$ or $p \geq \min\{\widehat{p}^2(R), 1\}$, there exists $\underline{q}(p)$, $\bar{q}(p)$ such that $\forall q_0 \in [0, \underline{q}(p)[$, the sequence q_t converges to zero, $\forall q_0 \in [\underline{q}(p), 1]$, the sequence q_t converges to $\bar{q}(p)$.

Proof of Lemma 1: The proof uses the same arguments as in the proof of Proposition 1.

Let us now prove Proposition 6.

Step 1. We show that when R is sufficiently high, the government can grant a subsidy $\delta = 1 - \frac{c^B}{c^S}$.

The government can implement a subsidy $\delta = 1 - \frac{c^B}{c^S}$ if and only if

$$R \geq pC(q_t; p, 1 - \frac{c^B}{c^S})(1 - \frac{c^B}{c^S})c^S(\tau^S)^2/2.$$

The left-hand side is bounded above by[4]

$$\frac{\beta^2}{4(\beta+\sigma)}(c^S - c^B).$$

A sufficient condition for the above inequality is

$$R > \frac{\beta^2}{4(\beta+\sigma)}(c^S - c^B).$$

Step 2. We study the dynamics when $\delta = 1 - \frac{c^B}{c^S}$.

If the government implements the subsidy $\delta = 1 - \frac{c^B}{c^S}$, following the proof of Proposition 1, the stationary equilibria are 0, 1, and q such that

$$h^{**}(q_t) = q_t(1-q_t)\Delta^B = \frac{1}{4}.$$

From the proof of Proposition 1, we know that the above equation admits two solutions: $\frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$ and $\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$. We also know that for any $q_0 > \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$, q_t converges to $\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$.

Step 3. We study the dynamics when $\delta = 0$.

When $\delta = 0$, the stationary equilibria are 0, 1, and q such that

$$h(q_t, R) = q_t(1-q_t)\left(\Delta^B - p\left(-q_t K + \beta - p(\beta+\sigma) - w(R)\right)(\Delta^B - \Delta^S)\right) = \frac{1}{4}.$$

Depending on p, this equation admits zero or two solutions $\underline{q}(p, R)$ $\bar{q}(p, R)$.

[4]This inequality comes from the fact that $pC(q_t; p, 1 - \frac{c^B}{c^S}) < p(\beta - p(\beta+\sigma)) < \frac{\beta^2}{4(\beta+\sigma)}$.

Step 4. Let us compare the dynamics under $\delta = 1 - \frac{c^B}{c^S}$ and $\delta = 0$. We have

$$q_t(1-q_t)\Delta^B > q_t(1-q_t)\left(\Delta^B - p\left(-q_tK + \beta - p(\beta+\sigma) - w(R)\right)(\Delta^B - \Delta^S)\right),$$

which implies $\frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B} < \underline{q}(p,R)$ and $\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B} > \bar{q}(p,R)$. For any $q_0 > \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$, when a subsidy $\delta = 1 - \frac{c^B}{c^S}$ implemented, the sequence q_t converges to $\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$. For any $q_0 > \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$, when $\delta = 0$, the sequence q_t converges to $\bar{q}(p,R) < \frac{\Delta^B + \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$ or $0 < \frac{\Delta^B + \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$. Hence, $\forall q_0 > \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$ strong subsidies have a positive impact on long-run honesty.

Proof of Proposition 7 :

Step 1. Express the long-run crime rate under $\delta = 0$.

Under $p \in]\hat{p}^1(R), \min\{\hat{p}^2(R), 1\}[$, the long-run crime rate under $\delta = 0$, which we denote by $\bar{C}_0$, is given by $\bar{C}_0 = \beta - p(\beta+\sigma) - w(R)$.

Step 2. Let us express the long-run crime rate under $\delta = 1 - \frac{c^B}{c^S}$ and $q_0 > \frac{\Delta^B - \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B}$. This is implicitly given by

$$\bar{C}_\delta - \left(-K\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B} + \beta - p(\beta+\sigma) - w(R - p\bar{C}_\delta \frac{\delta}{(1-\delta)^2} c^S \left[q_t(1-q_t)\Delta^S\right]^2)\right) = 0.$$

Step 3. Let us compare the crime rate without a subsidy with the crime rate with a strong subsidy. Note first that

$$\bar{C}_\delta < -K\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B-1)}}{2\Delta^B} + \beta - p(\beta+\sigma).$$

A sufficient condition for $\bar{C}_\delta < \bar{C}_0$ is

$$\begin{aligned}
&-K\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B} + \beta - p(\beta+\sigma) < \beta - p(\beta+\sigma) - w(R) \\
&\Leftrightarrow K\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B} > w(R), \\
&\Leftrightarrow w^{-1}(K\frac{\Delta^B + \sqrt{\Delta^B(\Delta^B - 1)}}{2\Delta^B}) > R,
\end{aligned}$$

whenever the inverse function w^{-1} exists.

Proof of Proposition 8: 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 9:

(i) Suppose that $Q_0 \in \left[0, \underline{q}\right[$, i.e., the initial proportion of honest families is low. First, from Proposition 8, 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 $q_1^* = q_2^* = \overline{q}$. ∎

B Urban equilibrium

Let us first calculate the bid rent of each parent of type $i = h, d$ residing in neighborhood $n = 1, 2$. Consider utility (19). Then, by using (5) and (7), we have

$$u_n^k = 4q_{n,t}^2(1 - q_{n,t})^2 \frac{(\Delta V)^2}{c^k} + q_{n,t}^2 \Delta V + V^d - c^k \frac{\left(\tau_t^k\right)^2}{2}, \tag{B.1}$$

where $\Delta V = V^h - V^d$.

To calculate each bid rent in the land market, we need to compute the expected utility of a worker of type (i, n) before the revelation of θ. To simplify the presentation, we skip the time index. We have

$$U_n^i = \int_0^{\theta^i} \left[(1-p)\,\beta - p\sigma - K\,\mathbf{1}_{i=h} - \theta\right] d\theta + \int_{\theta^i}^1 w d\theta - \rho_{n,t}^i + \int_0^{\theta^i} \left[p u_n^S + (1-p)\, u_n^B\right] d\theta + \int_{\theta^i}^1 u_n^B d\theta,$$

where $\mathbf{1}_{i=h}$ is an indicator function equal to one if the parent is of type d and zero otherwise. This utility can be written as

$$U_n^i = \left[(1-p)\,\beta - p\sigma - K\,\mathbf{1}_{i=h}\right]\theta^i - \frac{\theta^2}{2} + \left(1-\theta^i\right) w - \rho_{n,t}^i + p\theta^i u_n^S + \left(1 - p\theta^i\right) u_n^B. \tag{B.2}$$

We can now define the bid rent for a parent of type $i = h, d$ residing in neighborhood $n = 1, 2$. We have

$$\rho_{n,t}^i = \left[(1-p)\,\beta - p\sigma - K\,\mathbf{1}_{i=h}\right]\theta^i - \frac{\theta^2}{2} + \left(1-\theta^i\right) w + p\theta^i u_n^S + \left(1 - p\theta^i\right) u_n^S - U_n^i, \tag{B.3}$$

where u_n^S and u_n^B are defined in (B.1).

To determine the urban equilibrium, we need to determine the bid rent differential $\Delta\rho \equiv \rho^h - \rho^d$ between honest and dishonest families. Again, we skip the time index. The bid rent

$\rho^i = \rho_1^i$ that makes both neighborhoods equally attractive to a trait$-i$ parent is such that $U_1^i = U_2^i$ for $i = h, d$. Given that $\rho_2^i = 0$, we obtain

$$\rho^i = p\theta^i \left(u_1^S - u_2^S\right) + \left(1 - p\theta^i\right)\left(u_1^B - u_2^B\right).$$

The bid rent differential, $\Delta\rho \equiv \rho^h - \rho^d$, is then given by

$$\Delta\rho = p\left(\theta^d - \theta^h\right)\left[\left(u_1^B - u_2^B\right) - \left(u_1^S - u_2^S\right)\right].$$

By using (B.1), it is easily verified that

$$\left(u_1^B - u_2^B\right) - \left(u_1^S - u_2^S\right) = 4\left(\Delta V\right)^2\left(\frac{1}{c^B} - \frac{1}{c^S}\right)\left[q_1^2(1-q_1)^2 - q_2^2(1-q_2)^2\right].$$

As a result, by using (1) and (2), we obtain

$$\Delta\rho = 4pK\left(\Delta V\right)^2\left(\frac{1}{c^B} - \frac{1}{c^S}\right)\left[q_1^2(1-q_1)^2 - q_2^2(1-q_2)^2\right]. \tag{B.4}$$

Since $q_2 = Q - q_1$, this equation can be written as

$$\Delta\rho\left(q_1\right) = 4pK\left(\Delta V\right)^2\left(\frac{1}{c^B} - \frac{1}{c^S}\right)\left[q_1^2(1-q_1)^2 - \left(Q - q_1\right)^2\left(1 - Q + q_1\right)^2\right]. \tag{B.5}$$