"ECONOMICS OF CULTURE, INSTITUTIONS, AND CRIME"

Hosted by Fondazione Eni Enrico Mattei (FEEM)

Supported by
FP6 Priority 7 "Citizens and governance in a knowledge-based society" Project: Sustainable Development in a Diverse World (SUS.DIV) (Contract No. CIT3-CT-2005-513438)

University of Padua Research Project “Economic analysis of crime and social interactions” (grant CPDA071899)

Fondazione Eni Enrico Mattei (FEEM)

Organized by
SUS.DIV, FEEM, University of Padua and CEPR

Milan; 20-22 January 2010

Crime and Education in a Model of Information Transmission

Darwin Cortes, Guido Friebel and Dario Maldonado

We are grateful to the following institutions for their financial and organizational support: SUS.DIV, FEEM, University of Padua and CEPR.

The views expressed in this paper are those of the author(s) and not those of the funding organization(s) or of CEPR, which takes no institutional policy positions.
Crime and Education in a Model of Information Transmission\textsuperscript{1}

Darwin Cortes\textsuperscript{2} Guido Friebel\textsuperscript{3} Dario Maldonado\textsuperscript{4}

September 30, 2009

\textsuperscript{1}We are grateful to Giorgio Brunello, Georges Cassamatta and Gerhard Glomm, and participants at LACEA07 conference and Padua seminar for their comments. The usual disclaimer applies. A previous version of this paper appears as the third chapter of Darwin Cortes’ PhD dissertation at TSE. Cortés gratefully acknowledges financial support from the Colombian Central Bank and the University of Padua (Research Grant CPDA071899), as well as the hospitality of the Economics Department of the University of Padua.

\textsuperscript{2}Department of Economics, Universidad del Rosario, Bogotá, Colombia

\textsuperscript{3}Goethe-Universität, Frakfurt, Germany

\textsuperscript{4}Department of Economics and CeiBA-Complejidad, Universidad del Rosario, Bogotá, Colombia
Abstract

We model the decisions of young individuals to stay in school or drop-out and engage in criminal activities. We build on the literature on human capital and crime engagement and use the framework of Banerjee (1993) that assumes that the information needed to engage in crime arrives in the form of a rumor and that individuals update their beliefs about the profitability of crime relative to education. These assumptions allow us to study the effect of social interactions on crime. We first show that a society with fully rational students is less vulnerable to crime than an otherwise identical society with boundedly rational students. We also investigate the spillovers from the actions of talented students to less talented students and show that policies that decrease the cost of education for talented students may increase the vulnerability of less talented students to crime. This is always the case when the heterogeneity of students with respect to talent is sufficiently small.

Keywords: human capital, crime, rumors, education policies

JEL codes: D82, D83, I28
1 Introduction

Many developing countries and poor areas in developed countries are plagued by high crime rates and low levels of education. Young people seem to be particularly vulnerable to crime engagement. Oftentimes, once crime has started it spreads in an epidemiological way through a community. We here suggest a theory of juvenile crime that is motivated by the idea that the farther people are from receiving a return on educational investments they have made, and the more likely they are to be surrounded by other young criminals, the more they will be willing to engage in crime. It allows us to investigate the effect that educational policies have on the diffusion of crime engagement among young people.

Our theory is motivated by the fact that crime is a social phenomenon. Following Becker (1968), economic theory sees crime as an occupational choice or investment opportunity. A person compares the streams of payoffs from crime versus other occupations or investments in human capital such as going to school to obtain a good job later. Particularly interesting is Lochner (2004) who builds a dynamic model of education and crime engagement and explains the decreasing age-petty crime pattern. The more individuals have invested in education, the larger the opportunity cost of crime. Hence, older people who have accumulated more human capital or are closer to graduation, will be less prone to engaging in crime.

There is evidence supporting the ideas of Becker and Lochner (see Levitt (1998) and Mocan and Rees (2005)). But there is also evidence showing that crime engagement decisions are not completely described by the traditional neoclassical model and that a model with that aim should introduce new features. In particular, social interactions are important determinants of crime engagement. Ludwig et al. (2001) and Kling et al (2005) show that neighborhood’s wealth has an incidence in youth crimes. Particularly important to our paper is the evidence found by Case and Katz (1991) who show that in low-income Boston neighborhoods the behavior of peers appears to affect youth behaviors in a manner suggestive of contagion models. Another important piece of evidence is Luallen (2006) where it is shown that reducing school incapacitation increases crime rates among youngsters.

Taken together, the previous literature shows that the causal link between crime and low levels of human capital is quite complex. However, there seems to be agreement that fostering education is a good way to fight crime.

We investigate the interaction between educational policies and juvenile crime. We assume that everybody is rational, but that information on the opportunity to become a criminal is not readily available. Rather it is transmitted through a matching process in society: people who have become criminals meet students and students learn about the possibility to become a criminal rather than going to school. Our assumption is in line with the
evidence cited above. We investigate the nature of the information transmission process between criminals and students and carry out an investigation on the policies that reduce the cost of education such as scholarships, meals or transport subsidies, better teachers and materials.

We consider social interactions using a model of a rumor process à la Banerjee (1993). People are rational, they are young and go to school. Going to school costs some effort or money. Some of the students are more talented, thus have lower costs while others are less talented, and have higher costs of going to school. Talent (or ability) is private information. There is aggregate uncertainty: crime may pay or not and, because of differences in the opportunity costs of crime engagement, the payoff of engaging in crime depends on whether you are talented or not. Information on crime is not common knowledge but travels as a rumor. Upon hearing the rumor, a student updates the likelihood of crime being profitable and decides whether to stay in school or become a criminal. The time that passes before a given student meets a criminal for the first time provides crucial information about the probability that crime is profitable. This is so because the speed of the rumor transmission depends on the number of criminals, which in turn depends on profitability of crime.

We show that there is a point in time after which talented students will not be tempted anymore to become criminals. There is also a point in time for the less talented, but it occurs later. Hence, the less talented are more vulnerable to crime engagement. Indeed if students are fully rational and take into account the time passed before the hear the crime rumor, they will be less likely to engage in crime. This holds for both the talented and the less talented. Hence, social interaction need not increase crime, provided that people understand the diffusion process.

More important for policy issues, we show that social interactions play a role in fully rational students’ decisions. The behavior of the talented students affects the behavior of the less talented ones, but not the other way round. Consider a policy reducing the cost of schooling for talented students (for instance, a meritocratic scholarship program). This policy directly reduces the vulnerability to crime of talented students. To understand the effect on less talented students, the way the rumor about crime spreads at any time afterwards is crucial. This rumor diffusion depends on both the number of individuals who have already engaged in criminal activities and the speed with which the rumor spreads. A reduction in the costs of education for talented students results in these two forces going into opposite directions. Consequently, a policy reducing the cost of education of talented students may increase the vulnerability of less talented students. Moreover, when the heterogeneity of students is sufficiently small, such a policy always increases crime among less talented students.

The paper is related to a broader literature on information diffusion, such as Banerjee (1992) and Scharfstein and Jeremy (1990), who develop
models of herd behavior. In those models information goes through a pro-
cess of word-of-mouth learning and they are thought to explain financial
runs, behavior facing new products, etc. In the context of social economics,
Jackson and Yariv (2008) have recently reviewed the literature on the in-
fluence of social networks on diffusion processes in different realms, such as
disease contagion, technology adoption, vote decisions, etc. Previously, Ak-
erlof (1997) developed a model that shows how social position may affect
decisions such as education attainment or childbearing. Economic models
of social interactions and crime were firstly developed by Sah (1991) and
Glaeser et al. (1996). The former develops a model in which the decision
of a person to commit crime reduces the probability of other offenders to
be arrested. The latter develop a model in which the individuals decision
about crime depends on their neighbors’ decisions about criminal activities.
The model helps them to explain the cross-city variance of crime rates.

The rest of the paper is divided as follows: in the next section, we present
our model and a benchmark that brings across any extension of a Becker-
type. In the benchmark the probabilities associated to the profitability
of crime are exogenous. In our model this is a result of assuming that
students are boundedly rational - they do not understand the rumor process.
Alternatively, the same results are obtained assuming that the information
set available to students is more limited. We maintain the bounded / full
rationality terminology across the whole paper. In the third section, we
present our main results when we consider a society conformed by fully
rational students. In the fourth section, we analyze the social interactions
among student types and how this may affect education policies. In the last
section we conclude with some final remarks.

2 Model Setting

We consider population of students given by the interval \([0, 1]\) with equal
life length \(T\). We denote \(s\) the length of schooling of a student. After
graduation, students earn an income of \(W\) in each moment of the rest of
their lives. Education is costly; the instantaneous cost of education (in
terms of effort, tuition etc) is \(C(e) = e\). There are two types of students: a
proportion \(q\) of the students have high costs, \(e\), and a proportion \(1 - q\) of the
population have low costs, \(e\). Leaving problems of access to credit markets
asides (a topic that is beyond the scope of this paper), notice we can refer
for simplicity to high-cost students as “less talented” and low-cost students
as “talented”.

To simplify the model, we assume that the discount rate is equal to zero
and that:

**Assumption A0** At \(t = 0\) the entire population is attending school.
Education is a riskless project\(^1\). Its value depends on the moment of life of a person. At any moment \(t < s\) the instantaneous continuation value of education is:

\[
R(t) = \frac{(T - s)W - (s - t)e}{T - t} \tag{1}
\]

We have then \(R(t)\) and \(R(t)\) for \(e_i\) and \(e_j\), respectively. The idea of \(R(t)\) is that students must study a proportion \(s/T\) of life in order to obtain a degree and to earn \(W\) in each moment of the rest of their lives. Students hence first have to invest the cost of education to obtain its benefits afterwards. Clearly, the value of education increases in \(t\). The sunk-cost nature of education will be a crucial feature in our model. We will simply refer to \(R(t)\) as the value of education.

Assume for a first benchmark that becoming a criminal were a riskless project with instantaneous returns \(a_0\). Then if \(a_0 < R(0)\) there is no crime. If \(R(0) < a_0 < R(0)\), the high-cost (less talented) students commit crime during their entire life and low-cost (talented) students commit no crime and the total number of criminals is \(q\). And, if \(R(0) < a_0\), all students commit crime during all life, in this case the total number of criminals is 1.

It makes much more sense, however, to consider that being a criminal is a risky project. Consider that its returns are \(a\) with probability \(p\), \(b\) with probability \(r\) and \(d\) with probability \(1 - p - r\), where \(a > b > d\).\(^2\)

We make the following assumptions about the interaction between education, crime and different types of students:

**Assumption A1** \( (T - s)W - se > 0 \)

**Assumption A2** \( pa + rb + (1 - p - r)d < 0 \)

**Assumption A3** \( W \geq a \)

**Assumption A4** \( R(0) < b < R(0) < a \)

Assumption A1 says that education pays for the less talented students and hence also for the talented one. Assumption A2 is a somewhat stronger assumption: it says that the *ex ante* expected value of crime is negative such that without further information neither type of student would engage in crime. Assumptions A3 and A4 are about the crime-education decisions

---

\(^1\)One can argue that education may also be a risky project. However, the existence of institutions like minimum wages, that are common in both developed and developing countries, make the education project less risky than the crime project. Moreover, in those contexts in which education is riskier than crime, rumors about criminal projects may be more pervasive. Hansen and Machin (2001) present empirical evidence showing that the establishment of minimum wage actually *causes* a decrease in crime rates.

\(^2\)At the cost of further complication but without much benefit in terms of economic insights one could assume that each of the states \(a, b\) and \(d\) were lotteries themselves, with \(a\) the best lottery and \(d\) the worst.
during schooling time. Assumption A3 says that the riskless reward of education is larger or equal to the largest payoff of crime. As a consequence, nobody becomes a criminal after $s$, when all educational efforts are sunk. Assumption A4 says that, at $t = 0$, crime is profitable for the talented type only if the true state of the world is $a$. Crime is profitable for the less talented type if the true state of the world is either $a$ or $b$.

The previous assumptions deserve further discussion. Assumptions A1 and A2 together say that, ex-ante, education pays more than crime. This is in line with previous evidence on gang earnings showing that risks of criminal activities more than offset its wage premium with respect to legal earnings (Levitt and Venkatesh (2000)). Assumption A3 gathers the findings of Lochner and Moretti (2004), that high school graduation significantly reduces engagement in crime.

At $t = 0$, all the population is attending school. A proportion $x$ of the population learns the true state of the world, which is either $a$, $b$ or $d$. These students then choose whether to drop out of school (and commit crime) or to attend school (and exert effort).

**Assumption A5** If the student commits crime once, he stays a criminal for the rest of life, that is there is no way back to school once it is interrupted.

From $t = 0$ on, each student matches another agent in each instant. The matched agent may be either a criminal or a student. The student learns whether the matched agent is a criminal or not, but he does not learn the true state of the world or the agent’s type (talented or less talented as a student). We will let $m$ denote the event in which the student effectively matches an individual who had previously engaged in crime. When a student matches a criminal for the first time, he can choose whether to commit crime or not. This reflects the idea that crime is an occupational choice that becomes available only through social interaction. In order to adopt crime, one needs to have contact with other people who are criminals, because there are no

---

3. Recent research on brain development has shown that the adolescent brain is less developed than previously believed. During the teenage, it takes place a process similar to cutting down the branches of a tree that stimulates health and growth. The teenage brain discards grey matter – the tissue in charge of “thinking”. The brain arrives to maturity by the age of 22. These findings mean that teenagers are less able to assess consequences of acts, planning for the future, etc. In terms of our model these findings affect how teenagers assess both the education and the crime project. On the one hand, since the education project takes time and requires ability to plan the future properly, a teenage brain is more likely to undervalue education. On the other hand, since teenagers are less able to forward looking, they may not care for many of the consequences of engaging in criminal activities. Therefore, a teenage brain is more likely to overvalue crime. Taking together, teenagers will engage in crime more oftenly than predicted by our model.

4. This is so because wages after graduation are much higher than wages with no graduation.
formal channels through which one can take this type of career.⁵ Upon
meeting a criminal, a student is then confronted with the choice of staying
in school or engaging in a very different type of career.

Assume now, for a second benchmark, that updating is boundedly ra-
tional in the following sense:

**Assumption A6** A student who matches a criminal at time \( t \) updates
his belief about the state of the world, but he does not take into account
the point of time \( t = 0 \) at which the crime diffusion process has started.

Assumption A6 is meant to capture people’s limited knowledge or un-
derstanding of the diffusion process of crime. Assumption A6 implies that
the only information used by students to update their beliefs about the prof-
itability of crime when they meet a criminal is the distribution of types of
students and the *ex-ante* distribution of the profitability of crime. Students
who have never matched a criminal act as if nobody had committed crime,
and a student who meets a criminal at any time \( t' > t \) assigns the same in-
formational value to meeting a criminal as a student who meets the criminal
at time \( t \).

Formally, suppose the true state of the world is either \( a \) or \( b \) (we do not
have to consider the true state of the world \( d \) because then nobody commits
crime). When a student matches a criminal he learns that crime may be
profitable for him and updates the probability of each state of the world.
Let \( \pi = \frac{p}{p+q} \) be the updated probability of state \( a \). Students will commit
crime if and only if

\[
EC \equiv \frac{\pi}{\pi + q(1-\pi)} a + (1 - \frac{\pi}{\pi + q(1-\pi)}) b > R(t)
\]

**Result 1** Under boundedly rational updating, there is a time \( \tau \) at which a
student with cost of education \( e \) will not be tempted to adopt crime, but will
stay in school. \( \tau \) is given by:

\[
\tau = T \cdot E[C] + se - (T - s)W
\]

\[
\frac{E[C]}{E[C] + e}
\]

If we let \( \tau \) and \( \bar{\tau} \) represent the time at which, respectively, talented and less
talented students are not tempted to engage in crime we have that \( \tau < \bar{\tau} \).

This benchmark brings across the direct intuition from any extension of
a Becker-type model of crime. The lower the opportunity costs of crime,
the more likely people will adopt it. In our model, the opportunity cost
modeled explicitly is not the risk of being detected and punished (this is

---

⁵We exclude that one can become a criminal without having any contact with other
criminals, as we are interested in crime, education and social interaction through infor-
mation diffusion and not in the isolated decision of an individual to commit crime which
has been thoroughly studied by Becker and other scholars building on his work.
contained in reduced form in the parameters \( a, b, d \). Rather, we look at the process through which education is acquired. Over time, education becomes relatively less costly, because effort has already been sunk. Hence the opportunity costs of crime increase over time, independently whether one is talented or not. Further, the opportunity costs for the talented are higher, which explains why their temptation to engage in crime ends earlier.

3 Fully Rational Students

We maintain Assumptions A1 to A5 and now assume that students take the time dimension into account when updating:

**Assumption A6’** Students know the distribution of types and the date \((t = 0)\) in which the rumor started.

Under assumption A6’, the process of information transmission about crime becomes a rumor process in the sense of Banerjee (1993). Criminals become a source of the rumor on crime and, thus, the probability of hearing the rumor (matching a criminal) increases with the number of criminals. Anybody’s decision whether or not to engage in crime thus creates information externalities. Under our assumptions, especially assumptions A1 and A2, nobody will invest in crime unless somebody learns that someone else has already committed crime. Rumor begins if the true state of the world is either \( a \) or \( b \). If \( a \), a proportion \( x \) of people will become criminals at \( t = 0 \). If \( b \), a proportion \( qx \) will do so. If the true state is \( d \), nobody will.

In our analysis there are three critical points in time for understanding the decision of individuals regarding their education/crime decisions; these are \( t^* \), \( \tau \) and \( \tau^* \). The first is the moment in which less talented students start finding attractive to engage in crime only when the true state is \( a \). The second is the moment when talented students stop engaging in crime and the third is when the less talented students stop engaging in crime.

We know from assumption A4 that at \( t = 0 \) crime is profitable for less talented students if the true state is either \( a \) or \( b \). Since the rumor on crime only begins if the true state of the world is either \( a \) or \( b \), by continuity, a less talented student that matches a criminal at a moment near zero, say at \( t = 0 + \varepsilon \), will engage in crime. We also know that the value of education is increasing with time since a portion of total cost of education is sunk at each moment. Formally, from equation (1) we obtain that

\[
\hat{R}'(t) = \frac{\hat{R}(t) + \varepsilon}{1 - t} > 0.
\]

Therefore, there exists a \( t^* \) such that \( \hat{R}(t^*) - b = 0 \). Then, less talented students who match a criminal before \( t^* \) will commit crime.

A talented student at any \( t \) or a less talented student at any \( t > t^* \) who match a criminal will commit crime if:
\[ EC(t) = p(t)a + (1 - p(t))b > R(t). \]  

(2)

where \( p(t) \) is the probability of the true state being \( a \) given that the student matches a criminal for the first time at time \( t \). It is estimated using Bayes’ rule:

\[ p(t) = \frac{\pi}{\pi + \frac{\text{Prob}[m|b,t]}{\text{Prob}[m|a,t]}(1 - \pi)} \]

with \( \pi = \frac{p}{p+r} \). \( \text{Prob}[m|s,t] \) is the probability that in state \( s \in \{a, b\} \) a student matches a criminal for the first time at moment \( t \). This ratio will be important for our analysis and we hence define it formally.

**Definition 1** \( z(t) \equiv \frac{\text{Prob}[m|b,t]}{\text{Prob}[m|a,t]} \).

Let us also also define, \( z^*(t) \), the relative net gain of engaging in crime when it is profitable relative to the relative net loss when it is not profitable:

**Definition 2** \( z^*(t) \equiv \frac{\pi(a-R(t))}{(1-\pi)(R(t)-b)} \).

Notice that \( z^*(t) \) depends on the type of student which affects \( R(t) \). Consequently, we let

\[ z^*(t) = \frac{\pi(a - R(t))}{(1 - \pi)(R(t) - b)} \quad \text{and} \quad z^*_*(t) = \frac{\pi(a - \bar{R}(t))}{(1 - \pi)(\bar{R}(t) - b)}. \]

Notice also that \( z^*_*(t) \) is defined in the interval \((t^*, T]\) while \( z^*(t) \) is defined in \([0, T]\).

It can then readily be shown that inequality (2) holds for talented students if \( z(t) < z^*(t) \) and for less talented students if \( z(t) < z^*_*(t) \). This analysis is summarized in the following result:

**Result 2** Under fully rational updating, the behavior of a less talented student at a time \( t \) is as follows:

1. If a less talented student matches a criminal for the first time at \( t \leq t^* \), he will engage in crime.

2. For \( t > t^* \), if \( z(t) \leq z^*_*(t) \), a less talented student engages in crime, while otherwise a less talented student stays in school.

3. If \( z(t) \leq z^*(t) \), the talented student engages in crime, while if \( z(t) > z^*_*(t) \), the talented student stays in school.
The previous analysis shows that the crucial element that appears in this analysis is that the decision to become a criminal or not when hearing the rumor depends on the relative probability of matching a criminal in each of the two states of the world. Indeed, note that the rumor on crime only begins if the condition in Equation (2) holds at \( t = 0 \), that is, both types must be vulnerable to crime at \( t = 0 \). Otherwise the student that matches a criminal will know that the criminal is high-cost type. Since the probability of matching a criminal depends on the number of criminals in each state of the world (which depends on the age of rumor), he will use the age of the rumor to calculate the number of criminals in each state of the world and accordingly update his beliefs about the probabilities of matching a criminal.

The decision rule we just established shows that there is a critical level for the relative probability of matching a criminal in state \( a \) with respect to the same probability in state \( b \) after which students keep on in school; the critical level depends on the costs that each individual faces to complete school.

From Result 2, it becomes clear that both types of students are vulnerable to crime. High-cost students are more vulnerable than low-cost students since they are likely to become criminals during a longer period of life. To understand the dynamics of the diffusion process and, in the next step, the effect of different policies, \( z(t) \), \( z^*(t) \) and \( \overline{z}^*(t) \) are crucial. We will analyze each of these functions in turns. The properties of \( z^*(t) \) and \( \overline{z}^*(t) \) are directly stated in Result 3 which follows; we then show the properties \( z(t) \) which is a more challenging task.

**Result 3** The functions \( z^*(t) \) and \( \overline{z}^*(t) \) are both monotonically decreasing in \( t \) and convex.

Proofs are in Appendix 2. According to Result 3 the profitability of crime decreases with time for both types, in particular because the education cost is continuously sunk at each moment of time. The result holds in the respective domain of each function; that is, for \( z^*(t) \) in \( t \in [0, T] \), and for \( \overline{z}^*(t) \) in \( t \in (t^*, T] \).

The analysis of \( z(t) \) is more challenging and it requires the use of additional notation and definitions. It is important to note here that while in the model where students are boundedly rational they believe that \( z(t) = q \), in the model where they are unboundedly rational they believe \( z(t) \) changes over time. Our analysis characterizes \( z(t) \).

The following is borrowed from Banerjee (1993).

**Definition 3** For \( i = a, b \), we define:

1. \( N(i,t) \equiv \text{the proportion of the population that has committed crime until time } t \text{ in state } i \).
2. \( P(i, t) \equiv \text{the proportion of the population that has not matched a criminal until time } t \text{ in state } i. \)

Using Definition 3 and the fact that \( \text{Prob}[m|b, t] = N(b, t)P(b, t) \) and \( \text{Prob}[m|a, t] = N(a, t)P(a, t) \), we can then write:

\[
 z(t) = \frac{N(b, t)P(b, t)}{N(a, t)P(a, t)}
\]  

(3)

Notice that \( P(a, 0) = 1 - x \), \( P(b, 0) = 1 - x \), \( N(a, 0) = x \), \( N(b, 0) = xq \), and that \( z(0) = q \).

From now on we will need to distinguish “regimes” in which \( z(t) \leq z^*(t) \) (called Regime 1), and \( z(t) > z^*(t) \) (Regime 2). Assume for the time being that \( \tau \) is the moment in which \( z(\tau) = z^*(\tau) \). Below we will show that this moment exists.

Notice that the process must start off in Regime 1 (i.e. the condition in Equation 2 must hold at \( t = 0 \)), otherwise, there will be no uncertainty about crime. If the process began in Regime 2, everybody who matches a criminal will know the criminal is a less talented student. Formally, the process must start off when \( z(t) \leq z^*(t) \), which at \( t = 0 \) boils down to 

\[
 q \leq \frac{\pi(a - R(0))}{(1 - \pi)(R(0) - b)}.
\]

In Regime 1, the dynamics of \( N(i, t) \) and \( P(i, t) \) are given by

\[
 \frac{dP(i, t)}{dt} = -N(i, t)P(i, t),
\]  

(4)

\[
 \frac{dN(i, t)}{dt} = N(i, t)P(i, t).
\]  

(5)

In Regime 2, the dynamics of \( N(i, t) \) and \( P(i, t) \) are given by

\[
 \frac{dP(i, t)}{dt} = -N(i, t)P(i, t),
\]  

(6)

\[
 \frac{dN(i, t)}{dt} = qN(i, t)P(i, t).
\]  

(7)

The intuition for the difference is of course that in Regime 2 only less talented students (a proportion \( q \) of the total population) may become criminals at \( t > \tau \).

Furthermore, for an economy that has always been in Regime 1 holds:

\[
 P(a, t) = 1 - N(a, t),
\]  

(8)

\[
 P(b, t) = 1 - x(1 - q) - N(b, t).
\]  

(9)

For an economy that has made its first transition to Regime 2 at moment \( \tau \) holds:

\[
 q[P(a, \tau) - P(a, t)] = N(a, t) - N(a, \tau),
\]  

(10)

\[
 q[P(b, \tau) - P(b, t)] = N(b, t) - N(b, \tau).
\]  

(11)
Equations (8) and (9) evaluated in \( t = 0 \) together with equations (10) and (11) evaluated in \( t = \tau \) provide the initial conditions for the differential equations (4)-(7), respectively. With these things established the following Result can be shown readily (see the formal proof in Banerjee (1993)).

**Lemma 1 (The Banerjee effect)** (a) In both Regime 1 and Regime 2, we have that \( P(a, t) < P(b, t) \) and \( N(a, t) > N(b, t) \). (b) The term \( z(t) \) increases monotonically in \( t \) and is unbounded.

Result 1 is quite intuitive. The first part says that, in a given \( t \), the proportion of individuals that engage in crime is higher in state \( a \) than in state \( b \); consequently, in a given \( t \), the number of individuals that have not heard the rumor is smaller in state \( a \) than in state \( b \). The second part says that the older the rumor, that is the larger \( t \), students believe that the more likely state \( b \) is. The later one meets a criminal for the first time, students believe that it is more likely that the benefits of crime are low.

From equations (4)-(11) one can see that since the dynamics in both regimes differ \( z(t) \) will have different forms in each regime. For \( t < \tau \), in Regime 1, \( z(t) \) is defined by (4), (5), (8) and (9), we will let \( z^r_1(t) \) represent this part of the function. For \( t \geq \tau \), in Regime 2, \( z(t) \) is defined by (6), (7), (10) and (11), we will let \( z^r_2(t) \) represent this part of the function. \( z^r_1 \) depends on \( x \) and \( q \); \( z^r_2 \) depends on \( x \), \( q \) and the parameters that determine \( z^* \) (since \( \tau \) is determined by the equality \( z^* = z^r_1 \)). Indeed \( \tau \) defines the initial conditions for (and determines the actual path followed by) \( z^r_2 \). We can then define formally the function \( z(t) \) as follows:

\[
z(t) = \begin{cases} z^r_1, & \text{if } t \leq \tau; \\ z^r_2, & \text{if } t > \tau. \end{cases}
\]  

Explicit expressions for \( z^r_1(t) \) and \( z^r_2(t) \) can be easily obtained using (3), (4)-(11). For our analysis we will need the explicit expression for \( z^r_2 \) which can be expressed in terms of \( P(i, \tau), N(i, \tau) \) and \( (t - \tau) \), as follows

\[
z^r_2 = \frac{dN(b,t)}{dN(a,t)} = \frac{qN(b,t)P(b,t)}{qN(a,t)P(a,t)} \frac{dN(b,\tau)}{dN(a,\tau)} g(P(b, \tau), t - \tau) \frac{dN(a,\tau)}{dN(a,\tau)} f(P(a, \tau), t - \tau)
\]

The specific forms \( z^r_2 \) and \( z^r_1 \) appear in the Appendix 1 where it will be clear that \( g \) and \( f \) are two specific functions.

The following Proposition shows that \( \tau \) exists.

**Lemma 2** Provided that \( q \leq \frac{\pi(a-R(0))}{(1-\pi)(R(0)-b)} \), there will be an instant \( \tau \) at which there will be a transition from Regime 1 to Regime 2.

We have so far established that the beliefs on the true state of the world converge to \( b \) (Result 3), that the value of education relative to crime is
increasing over time (Result 3), and that \( \tau \) exists (Result 2). It is important to note that the rumor on crime goes beyond \( \tau \). Consider an student that matches a criminal in \( t > \tau \). Although he knows that talented students are not vulnerable anymore, he also knows there are criminals who had been talented students, but matched a criminal before \( \tau \). At \( \tau \) non talented students are still vulnerable to crime.

Following a similar reasoning we can show how the less talented students behave. Indeed, from Results 3 and 1 we have that \( z(t) - z^*(t) \) is monotonically increasing. Therefore there exists a \( \tau \) such that \( z(\tau) = z^*(\tau) \). Since an increasing amount of education cost is sunk over time, there is a moment in which no student becomes a criminal anymore. After \( \tau \), education is more valuable than crime for both types. The total number of criminals thus reaches its maximum at \( \tau \). After \( \tau \), some students will still match criminals (who can be of either type) but they will not engage in crime. As we have argued before, the time at which talented students stop engaging in crime, \( \tau \), is strictly shorter than the time at which less talented students do so, \( \overline{\tau} \). In a nutshell, less talented students are more vulnerable to crime than talented students since their cost of schooling is larger \( (s\overline{\tau} > s\tau) \). Figure 1 depicts the solutions we presented above.

The main lesson of this section can be summarized in the following proposition.

**Proposition 1** A society with fully rational students will be less vulnerable to crime than an otherwise identical society with boundedly rational students.
Figure 2: Decisions of boundedly and fully rational students

When boundedly rational students match a criminal they learn that crime is profitable (contrary to what was signalled by its ex-ante value) and thus update their belief on the true state of the world. Since they do not fully understand the diffusion process of crime, they believe the relative probability of \( b \) to be the true state of the world is equal to \( q \) regardless of the moment of matching a criminal. Fully rational students that match criminals also update their beliefs about the true state of the world. Unlike boundedly rational students, they do so by estimating the number of people that may have heard the rumor in each state of the world. The age of rumor is crucial in this estimation. Fully rational students consider that the older the rumor the less valuable it is. Therefore, fully rational students are less vulnerable to crime than boundedly rational students. Figure 2 illustrates Proposition 1. Each type of boundedly rational students are no longer vulnerable to crime at \( \tau \) (low-cost) and \( \bar{\tau} \) (high-cost). Each type of fully rational students are no longer vulnerable to crime at \( \tau' \) (low-cost) and \( \bar{\tau}' \) (high-cost). Fully rational students are less vulnerable than boundedly rational students since \( \tau' \) and \( \bar{\tau}' \) smaller than \( \tau \) and \( \bar{\tau} \), respectively.

Before addressing the effect of policy on criminal engagement of young people there is one more characteristic of \( z(t) \) that we must consider, this is stated in the following result.

**Lemma 3** At \( t = \tau \), there is a downward kink in \( z(t) \), that is the slope of \( z(t) \) is larger for \( t \leq \tau \) than for \( t > \tau \).  
There is a kink at \( \tau \) because beyond this point, talented students who have not yet met a criminal will never engage in crime. At the kink the speed of the rumor decreases, which has important consequences for the policy effects we present below. Interestingly, changes in speed at a given \( t \) affect the relative likelihoods of the states \( a \) and \( b \); the likelihood of state \( a \) is actually higher if the speed is low. Policies may move the point at which the slope changes left or right. Suppose that there is a change that moves \( \tau \) to the right. Then, there will be a range of time between the old and the new \( \tau \) in which the rumor is dispersing faster than before the policy shock. This introduces some subtleties into the policy analysis, which will be used in the following section.

4 Policy: the effect of changing the costs of education

In the previous section we have investigated how crime spreads in a society and how it affects education. We here show that policies reducing the cost of education may have surprising effects. Examples of such policies are reductions of tuition fees, food for school programmes, improvements in school infrastructure or teachers. These measures can be given depending on performance of a student, which makes them contingent on their type (talented vs less talented).

Consider policies such that \( e \), \( \tau \) or both are reduced. These reductions have a direct effect on the vulnerability to crime of the targeted type of student, but there is also an indirect effect through the transmission of information about crime profitability among students of different types. The direct effect of reducing \( e \) (\( \tau \)) shifts \( \tau \) (\( \tau \)) to the left. That is, the point in time at which no more students of a given type will engage in crime occurs earlier when their cost of attending and succeeding at school decreases. Put differently, students become less vulnerable to crime.

The indirect interaction, i.e. effects between different types, are much more subtle, but they only go from talented students to less talented students. To understand this statement notice first that both the talented and the less talented students carry out the same type of comparison between costs and benefits of engaging in crime. More precisely, they both have the same uncertain benefit of engaging in crime and they assign the same probability distribution to the states \( a, b, c \). Also, wages upon graduation are assumed to be the same for both types. The opportunity costs of engaging in crime, however, are type-dependent: talented students have lower costs of going to school than less talented ones, which explains why \( \tau \) is to the left of \( \tau \).

For any \( t \) smaller than \( \tau \), any change in the parameters affects equally the expected benefits both types of students assign to crime. This implies
that among criminals, the proportion of talented and less talented types is constant, reflecting the respective proportions in the entire population. There are interaction effects here, but they do not depend on the types. This changes at \( t = \tau \). Here, no further talented student engages in crime, while less talented students who meet a criminal continue to do so, implying that the proportion of less talented types among criminals increases. Hence, when one wants to understand the interaction effects between different types, it suffices to investigate how a shift in \( \tau \) will affect the behavior of less talented students. This is stated in the following proposition.

**Proposition 2** Effects of a decrease in \( e \) (the costs of education for talented students): (i) \( \tau \) shifts to the left i.e. talented students become less vulnerable to crime. (ii) The effect of a reduction in \( e \) on \( \tau \) is ambiguous, in particular a reduction of \( e \) may result in an increase of \( \tau \) i.e., less talented students may become more vulnerable to crime.

We have shown in section 3 that the diffusion of crime depends on the probability of meeting a criminal in state \( b \) relative to the probability of meeting a criminal in state \( a \), at a given instant \( t \). This means that whether the vulnerability of less talented students increases or decreases when \( e \) decreases depends on whether this relative probability \( (z(t)) \) decreases or increases. Proposition 2 results from the interplay of two forces that affect this relative probability: First, the decrease of \( e \) generates an instantaneous reduction in the number of criminals, thus reducing the overall number of criminals in each state of the world. As shown by Lemma 3 this means that it is more likely to match a criminal when the true state of the world is \( a \) instead of \( b \). Second, the Banerjee effect holds (Lemma 1), namely, with time the belief of \( b \) being the true state of the world increases. When the first effect dominates, decreasing \( e \) increases \( \tau \); when the second effect dominates, \( \tau \) decreases.

To see these two effects analytically, consider two levels of effort for low-cost students, \( e' \) and \( e'' \), such that \( e'' < e' \); these levels correspond to the initial one \( (e') \), before the policy shock and the final one, after the policy shock \( (e'') \). For these two levels of effort we have two values for \( \tau \), \( \tau' \) and \( \tau'' \), and two different functions for \( z(t) \), \( z'(t) \) and \( z''(t) \). \( \tau' \) and \( \tau'' \) are such that \( \tau'' < \tau' \). At each value of \( \tau \), there is a proportion of the population that has not heard the rumor. These proportions are \( P(a, \tau) \) and \( P(a, \tau'') \) in state \( a \) and \( P(b, \tau) \) and \( P(b, \tau'') \) in state \( b \). To know which is the result of changing \( e \) on crime engagement of untalented students we must compare \( z'(t) \) and \( z''(t) \).

Two different systems that differ in \( e \) only differ in their dynamics after the lower \( \tau \). The differences are consequence of changing the initial conditions for \( z^{\tau} \) and on changes in the dynamics of \( z(t) \) after \( \tau \). Precisely, what one wants to know is whether, for \( t > \tau \), \( z'' \) is to the left or to the right of \( z' \).
or whether they cross. If they cross one also wants to know if $z''$ crosses $z'$ from above or from below. If $z''$ is to the right (left) of $z'$ this would mean that reducing the cost of talented students increases (reduces) the vulnerability of less talented students. In other words if $z''$ is to the right (left) of $z'$ a reduction in $e$ would bring the undesirable effect of increasing $\tau$; in the other case reducing $e$ would have a positive externality since it would also reduce $\tau$. A full comparison of $z'(t)$ and $z''(t)$ is in general not possible. However, the following example shows that for two sets of parameters in which the only difference is the value of $a$, a change in $e$ of the same size induces changes in $\tau$ of different signs.

Consider the following values for the parameters of our model: $x = 0.1$, $q = 0.6$, $\tau = 10$, $W = 69$, $a = 55$, $b = 15$, $\pi = 0.6$ and $T = 34$. The effort of the less talented students is $e = 153$ and the effort of the talented students is $e = 100$. With this information, all less talented students that hear the rumor about crime before $\tau \approx 5.44$ choose to become criminals. Similarly, all talented students that hear the rumor before $\tau \approx 4.07$ will do so. Let us consider a policy that reduces the costly effort of education for talented students. It reduces $e$ to 35. This policy reduces the vulnerability to crime of talented students to 1.58 and makes less talented students more vulnerable to crime increasing $\tau$ to 5.51. This example is depicted in Figure 3.

Now, consider an alternative situation in which crime pays more, such
that $a = 64$. In this case the initial $\tau$ is 5.39; the initial $\tau$ is 5.96. When the policy that reduces $e$ from 100 to 35 is implemented, the vulnerability of talented students decreases to 2.68 but less talented students become less vulnerable to crime; $\tau$ decreases from 5.96 to 5.93. The example is in figure 4. The next proposition identifies a sufficient condition under which there is no ambiguity.

**Proposition 3** Comparative statics with respect to student heterogeneity $(\bar{e} - e)$. For sufficiently low levels of student heterogeneity, a decrease in the cost of education of talented students makes less talented students more vulnerable to crime.

**Proof.** Consider a situation in which $\bar{e} - e$ is small. Consequently $\bar{\tau} - \tau$ is also small. Consider two levels of cost of education for talented students $\bar{e}'$ and $\bar{e}''$, such that $\bar{e}' > \bar{e}''$. The corresponding moments in which talented students stop engaging in crime are $\tau'$ and $\tau''$, and they satisfy $\tau' > \tau''$. We also have two functions for $z(t)$; let these functions be $z'(t)$ and $z''(t)$. These two functions are exactly the same for any $t \leq \tau''$ and differ for $t > \tau''$. Consider a $t$ such that $\tau' > t > \tau''$. Since $\tau' > t > \tau''$, $t$ belongs to Regime 1 when $e = \bar{e}'$ and to Regime 2 in the second case. From Lemma 3, we have that there is downward kink at $\tau$. Therefore, since functions $z(t)$, $z'(t)$ and $z''(t)$ are continuous, for $t$ near enough to $\tau''$ we have that $z'(t) > z''(t)$ for $t > \tau'$. Since $\tau'$ is near $\tau', \tau''$ is near $\tau''$ and $z'(t)$ is downward sloping we then have that $\tau' < \tau''$. □
5 Concluding remarks

We have suggested a theory of education and crime of young individuals. Information needed to engage in crime is not available to everybody; to become a criminal an individual has to hear information that affects the profitability of crime from someone who is already a criminal. Crime thus spreads in an epidemic fashion, as in the literature on rumors. We have shown that social interactions among fully rational students reduce crime engagement. Second, we have studied how informational externalities between different types of individuals affect crime engagement. The informational externality is such that policies aiming to decrease the costs of education of talented students have effects also on the education success and on crime engagement of less talented students. In particular, they may increase the vulnerability of less talented students to crime, which is always the case when heterogeneity of students with respect to talent is sufficiently low.
Appendix

A The form of $z^{r_1}$ and of $z^{r_2}$

The specific forms of the functions $z^{r_1}$ and $z^{r_2}$ are

$$z^{r_1} = \frac{dN(b,t)}{dt}$$

where $\frac{dN(b,t)}{dt} = \frac{(1-x)(1-x(1-q))^2xqe^{(1-x(1-q))t}}{(1-x+xe^{(1-x(1-q))t})^2}$ and $\frac{dN(a,t)}{dt} = \frac{x(1-x)e^t}{(1-x+xe^t)^2}$.

$$z^{r_2} = \frac{dN(b,\tau)}{dt} \cdot g(P(b, \tau), t-\tau) \quad \frac{dN(a,\tau)}{dt} \cdot f(P(a, \tau), t-\tau)$$

where $\frac{dN(b,\tau)}{dt} = q(1-x)(1-x(1-q))^2xq^{(1-x(1-q))\tau} \quad \frac{dN(a,\tau)}{dt} = qx(1-x)e^{\tau}(1-x+xe^{\tau})^2$.

Besides,

$$g(P(b, \tau), t-\tau) = \frac{(1-x)(1-q) - (1-q)P(b, \tau))^2e^{(1-x(1-q)-(1-q)P(b, \tau))(t-\tau)}}{[qP(b, \tau) + (1-x(1-q) - P(b, \tau))e^{(1-x(1-q)-(1-q)P(b, \tau))(t-\tau)}]^2}$$

and

$$f(P(a, \tau), t-\tau) = \frac{(1-(1-q)P(a, \tau))^2e^{(1-(1-q)P(a, \tau))(t-\tau)}}{[qP(a, \tau) + (1-P(a, \tau))e^{(1-(1-q)P(a, \tau))(t-\tau)}]^2}$$

B Proofs

1. Proof of Result 3:

Since $z^*(t)$ is continuous and differentiable then to show the result it is enough to analyze the signs of the first two derivatives of $z^*(t)$.

We first need to know that

$$R'(t) = \frac{R(t) + e}{1-t} > 0$$

$$R''(t) = 2 \frac{R'(t)}{(1-t)} > 0$$

Taking the first derivative of $z^*(t)$, we obtain that

$$z^{**}(t) = -\frac{R'(t)}{R(t) - b} \left( \frac{\pi}{1 - \pi} + z^*(t) \right) < 0.$$
Taking the second derivative
\[ z''(t) = \left( \frac{\pi}{1 - \pi} + z^*(t) \right) \left[ 2 \left( \frac{R'(t)}{(R(t) - b)} \right)^2 - \frac{R''(t)}{(R(t) - b)} \right] > 0. \]

2. Proof of Lemma 2:
Since \( q \leq \frac{\pi(a - R(0))}{(1 - \pi)(R(0) - b)} \), the rumor on crime starts off. Let us first consider the case of strict inequality. Once a talented student matches a criminal, he updates beliefs on the state of the world and takes decisions following the rule in Result 2. Since, \( z(0) < z^*(0) \), the crime is profitable for talented students. Those students matching criminals at \( t = 0 \) will become criminals. This will be the behavior of talented students for all \( t > 0 \) provided that \( z(t) < z^*(t) \). From Result 3, \( z^*(t) \) is monotonically decreasing with time and from Result 1, \( z(t) \) is monotonically increasing with time. Therefore, the difference \( z(t) - z^*(t) \) monotonically increases with time. At \( t = 0 \), it is negative, then it becomes zero and finally it becomes positive. Let \( \tau \) be the moment at which \( z(\tau) - z^*(\tau) = 0 \). For all \( t > \tau \) the talented student that match a criminal will stay in school. At \( \tau \) there will be a transition from Regime 1 and Regime 2. Now let us consider the case of strict equality. In this case the talented students that know the true state of the world are indifferent between staying at school and becoming criminals. Those talented students that hear the rumor at \( t = 0 \) are also indifferent between school and crime. However, since \( z(t) - z^*(t) \) is monotonically increasing, all talented students that hear the rumor at all \( t > 0 \) will stay at school. In this case, \( \tau = 0 \).

3. Proof of Lemma 3:
To show that there is a downward kink at \( \hat{t} \), we have to show that
\[ \lim_{t \to \hat{t}}^{-} z'(t) > \lim_{t \to \hat{t}}^{+} z'(t). \]
Indeed,
\[ \lim_{t \to \hat{t}}^{-} z'(t) = \lim_{t \to \hat{t}}^{+} [z(t) [2(P(b, t) - P(a, t)) + x(1 - q)] \]
\[ = z(\hat{t}) [2(P(b, \hat{t}) - P(a, \hat{t})) + x(1 - q)] \]
and
\[ \lim_{t \to \hat{t}}^{+} z'(t) = \lim_{t \to \hat{t}}^{+} \left[ 2q(P(b, t) - P(a, t)) + (1 - q)(P(b, \hat{t}) - P(a, \hat{t})) + x(1 - q) \right] \]
\[ = z(\hat{t}) [(1 + q)(P(b, \hat{t}) - P(a, \hat{t})) + x(1 - q)] \]
From Lemma 1, \( P(b, \tau) > P(a, \tau) \). Since \( 2 > 1 + q \), then there is a downward kink in \( \tau \).

4. Proof of Proposition 1:

The idea of the proof is to compare the moment at which both types of students stop to become criminals under two models: one with time-independent updating, the other with time-dependent updating. Let us denote \( \tau \) and \( \tau' \) the moments in which low-cost students and high-cost students stop to become criminals, respectively, assuming time-independent updating (Benchmark 2).

Let us denote \( \tau' \) and \( \tau'' \) the moments in which low-cost students and high-cost students stop to become criminals, respectively, assuming time-dependent updating (our model). Then, we need to show that \( \tau > \tau' \) and \( \tau > \tau'' \).

On the one hand, with time-independent updating we have that

\[
\tau = \frac{T \cdot E[C] + s\bar{e} - (T - s)W}{E[C] + \bar{e}} \quad \text{and} \quad \tau' = \frac{T \cdot E[C] + s\bar{e} - (T - s)W}{E[C] + \bar{e}},
\]

where

\[
E[C] = \frac{\pi}{\pi + q(1 - \pi)} a + (1 - \frac{\pi}{\pi + q(1 - \pi)}) b.
\]

On the other hand, with time-dependent updating we have that

\[
\tau' = \frac{T \cdot EC(t) + s\bar{e} - (T - s)W}{EC(t) + \bar{e}} \quad \text{and} \quad \tau'' = \frac{T \cdot EC(t) + s\bar{e} - (T - s)W}{EC(t) + \bar{e}},
\]

where \( EC(t) = p(t) a + (1 - p(t)) b \) and \( p(t) = \frac{\pi}{\pi + z(t)(1 - \pi)} \).

We also have that at \( t = 0 \), \( z(0) = q \). Therefore \( E[C] = EC(0) \).

Finally,

\[
\frac{\partial \tau}{\partial z(t)} = \frac{\partial \tau}{\partial EC(t)} \frac{\partial EC(t)}{\partial z(t)},
\]

\[
\frac{\partial \tau}{\partial EC(t)} = \frac{(T - s)(\bar{e} + W)}{(EC(t) + \bar{e})^2} > 0
\]

and

\[
\frac{\partial EC(t)}{\partial z(t)} < 0.
\]

Then

\[
\frac{\partial \tau}{\partial z(t)} < 0.
\]
From Lemma 1, $z(t)$ is always increasing. Then

$$\frac{T \cdot E[C] + s \bar{\varepsilon} - (T - s)W}{E[C] + \bar{\varepsilon}} = \frac{T \cdot EC(0) + s \bar{\varepsilon} - (T - s)W}{EC(0) + \bar{\varepsilon}} = \tau > \tau'$$

for all $t > 0$.

Following the same procedure we obtain that $\tau > \tau'$ for $\bar{\varepsilon}$. 
References


