



Institute for Empirical Research in Economics
University of Zurich

Working Paper Series
ISSN 1424-0459

Working Paper No. 348

Cultural Transmission and Discrimination

Maria Sáez-Martí and Yves Zenou

Revised version, April 2012

Cultural Transmission and Discrimination*

Maria Sáez-Martí[†] Yves Zenou[‡]

April 12, 2012

Abstract

Workers can have *good* or *bad* work habits. These traits are transmitted from one generation to the next through a learning and imitation process, which depends on parents' investment in the trait and the social environment where children live. We show that if a sufficiently high proportion of employers have taste-based prejudices against minority workers, their prejudices are always self-fulfilled in steady state and minority workers end up having, on average, worse work habits than majority workers. This leads to a ghetto culture. Affirmative Action can improve the welfare of minorities whereas integration can be beneficial to minority workers but detrimental to workers from the majority group.

J.E.L. Classification: J15, J71.

Keywords: ghetto culture, overlapping generations, rational expectations, multiple equilibria, peer effects.

*We are grateful to the editor, William Strange, three anonymous referees, as well as Antoni Calvó-Armengol, Joan Esteban, Anna Sjögren and Fabrizio Zilibotti for their helpful comments.

[†]Institute for Empirical Research in Economics, University of Zurich. E-mail: saez@iew.unizh.ch

[‡]Stockholm University and Research Institute of Industrial Economics. E-mail: yves.zenou@ne.su.se.

1 Introduction

According to a survey conducted in Chicago in 1988, one of the main reasons why employers are not willing to hire inner-city black workers is their lack of basic skills and work ethics. As expressed by a suburban employer in Chicago, “The experiences that I’ve run into with it is that they develop bad habits, I guess is the best way to put it. Not showing up to work on time. Not showing up to work. *Somewhere down the road they didn’t develop good work habits.*”¹

This is consistent with more general evidence from sociology and anthropology² suggesting the existence of a persistent “ghetto culture” which is transmitted across generations. Several scholars have put forward the importance of a low work-ethic culture in inner-city neighborhoods. This culture is in sharp contrast with mainstream American society’s working values rooted in the Protestant tradition. As argued by Wilson (1996), it is the *social*, rather than the *physical distance*, that often separates poor workers from good jobs. This is particularly true for the African American community, which has experienced high levels of segregation for at least a century (Massey and Denton, 1993; Cutler et al., 1999).

“Inner-city social isolation also generates behavior not conducive to good work histories. The patterns of behavior that are associated with a life of casual work (tardiness and absenteeism) are quite different from those that accompany a life of regular or steady work (e.g. the habit of waking up early in the morning to a ringing alarm clock). ... in neighborhoods in which most families do not have a steadily employed breadwinner, the norms and behavior patterns associated with steady work compete with those associated with casual or infrequent work.” (Wilson, 1996)

In the words of a counsellor to a training program aiming at exposing workers to more conventional working values:

“To adopt a regular pattern you have to break with this environment. Your friends laugh at you for going to work, that’s hell, they think you are trying to be better than them! You have to have strong character to resist this pressure. If all your friends and families went to work they would help you adopt a regular schedule.” (cited in Bonney, 1975)³

Why do some groups perform worse in the labor market than others? Several explanations have been put forward in the economics literature. In taste-based models

¹See Wilson (1996) pages 119-120. The italics are ours.

²See, in particular, Hannerz (1969), Lewis, (1969), Wilson (1987), Lemman (1991) and Katz (1993).

³This is related to the idea of “acting white” where the economic success of blacks induces peers’ rejection (Austen-Smith and Fryer, 2005; Battu et al., 2007; Bisin et al., 2011).

(Becker, 1957), prejudiced employers discriminate against some workers who will only be hired at lower salaries. The statistical discrimination theory, on the other hand, stresses the role of employers' beliefs concerning the *average* quality of workers from different groups. A member of the disadvantaged group will be discriminated against if the employer believes her to be less qualified or reliable than a worker from other groups (see, e.g. Phelps, 1972; Arrow, 1973; Lundberg and Startz, 1983); Coate and Loury, 1993; Moro and Norman, 2003; Verdier and Zenou, 2004). In these models, negative stereotypes are self-fulfilling and discriminated workers end up being less productive.

More recently, it has been argued that the existence of neighborhood (or peer) effects can explain the poor performance of some workers. In the absence of interaction between different communities, due to the interaction with poorly performing peers, some groups end up with lower levels of education and adverse labor market outcomes (see, e.g. Arnott and Rowse, 1987; De Bartolome, 1990; Benabou, 1993).⁴

Other aspects, like work habits (i.e., individual's willingness to be unemployed, collect welfare, avoid shirking, or work long hours) can also explain the different performances of workers from different communities. These cultural traits, which affect individual performance in the workplace, are, as documented by the sociological literature, strongly influenced by the family environment and the neighborhood where children grow up.⁵ Employers are reluctant to hire some workers because the prevalent values in their communities serves a proxy of their actual type. Policies promoting integration may, on the contrary, have a positive effect via parents' investments.

We model the formation of a "work-habit" trait using a mechanism that interacts purposeful socialization decisions inside the family (*direct vertical socialization*) and indirect socialization processes via neighborhood effects and social interactions (*oblique* and *horizontal socialization*). Indeed, based on some works on anthropology and sociology (see in particular Boyd and Richerson, 1985 and Cavalli-Sforza and Feldman, 1981), there is a recent literature initiated by the seminal papers of Bisin and Verdier (2000, 2001)⁶ arguing that the transmission of a particular trait (religion, ethnicity, social status, etc.) is the outcome of a socialization inside and outside the family (like e.g. peers and role models). These two types of socialization are cultural substitutes (complements) if parents have less (more) incentives to socialize their children the more widely dominant their values are in the population.⁷ In our model, we assume that

⁴For a general overview of the issue of race in the labor market, see Altonji and Blank (1999) and Lundberg and Startz (2000).

⁵Mulligan (1996) shows that a child of parents who do not work and/or collect government benefits for not working displays a tendency to behave similarly as an adult. Kohn (1969) concludes that parents generalize their experiences on the job and pass them on to their children. More recently, Osborne Groves (2005) suggests that intergenerational transmission of personality may be a channel for explaining intergenerational persistence of income.

⁶Bisin and Verdier (2011) provide a very complete and recent overview of this literature. See also Bisin et al. (2011) who model the child's identity choice in a cultural transmission model.

⁷Bisin et al. (2004), Cohen Zada (2006), Patacchini and Zenou (2011a), Jellal and Wolff (2002),

parents are forward looking and invest resources in order to prepare their children for their future working experiences. Parents' efforts and children's preferences are affected by the neighborhood where children live.

To be more precise, workers belong either to the majority group or the minority group. All individuals are born equal but, depending on the parents' investments and the neighborhood, they acquire either a good (g) or a bad (b) work habit (absenteeism, tardiness, low reliability...). When deciding how much effort to exert on shaping their children's attitudes towards work, parents must form expectations about the working opportunities their children will face in the future.

We assume that each worker is randomly matched to an employer who has to assign the worker to a job. Employers know the group to which a worker belongs but cannot perfectly observe her type. A proportion of employers may be taste-based prejudiced against minority workers and do not want to employ them. All other employers (imperfectly) screen the workers and employ those who seem to have good working habits. This second group of employers are profit maximizers. The different treatments to which the workers are subject create a discrepancy in expected value in the market of the good trait for the two groups.

We first focus on a *segregated society* where minority and majority families live in different neighborhoods. We show that if the fraction of prejudiced employers is sufficiently high and the initial neighborhood quality of the minority sufficiently bad, the employers' beliefs are self-fulfilled. Indeed, in steady-state, all minority workers end up with *bad work habits* and also non prejudiced employers prefer not to employ workers from minority neighborhoods. Due to the adverse opportunities their children are going to face, parents do not find it worthwhile to exert effort to transmit "good" values. As a result, more workers from bad neighbourhoods have bad work habits. This, in turn, influences members of the next generations in the same neighborhood and the initial negative beliefs are confirmed in steady state.

We study the effect of different policies aiming at weakening this "ghetto" culture, which perpetuates bad working habits. First, we analyze two different Affirmative Action policies consisting of imposing (*i*) a quota for the discriminated group in *good* jobs and (*ii*) a quota of workers from the discriminated group who are treated as the workers of the other group. This second policy is just a general case of *anti-discrimination law* where anti-discrimination requires equal treatment of everyone and the Affirmative Action policy we consider here requires equal treatment of a fraction of workers. We show that the first policy has always a negative long-run effect by increasing the fraction of minority individuals with *bad* work habits while the second one always increases the proportion of minority workers with *good* work habits.

Patacchini and Zenou (2011b), and Bisin et al. (2010) provide empirical studies of cultural transmission and socialization of, respectively, religious traits, altruism, preferences for education and ethnic identity.

Second, we study the effect of *integration policies*. In this case, children are, to some extent, influenced by peers from both groups. We show that integration is beneficial for the workers who come from the worst peer group and detrimental for the others, in the short run. This result helps us understand why the latter may have an incentive to resist integration and may be reluctant to accept social mixing.

Our theoretical model is related to the literature on cultural transmission and statistical discrimination since it combines elements of both. However, there are important differences that we would like to highlight.

We assume, unlike most of the literature on cultural transmission (see Bisin and Verdier (2000) and Saez and Sjögren (2008)), that *(i)* all parents, irrespective of their type, agree that one of the traits (good habits) is superior⁸ and *(ii)* the cultural transmission is biased. In particular, the probability of adopting a trait when learning from society is a non-linear function of popularity, so that more popular traits are copied with a probability higher than their population share. Moreover, children may have their own views and favor one trait over the other. In order to allow for these biases we consider a general framework (see Saez-Marti (2011)) where the probability of adopting a certain trait depends on its popularity (frequency dependency) as well as on children's own evaluation of the trait. With unbiased transmission and parents' agreement on the ranking of traits, the only possible steady state is one in which all agents have the desirable trait. With the frequency-dependent bias both traits can co-exist in equilibrium even if all parents invest on the same trait. The children bias affects the actual equilibrium proportions.

Even though our model builds on the statistical discrimination literature (Phelps, 1972; Arrow, 1973; Lundberg and Startz, 1983; Coate and Loury, 1993; Moro and Norman, 2003), there are important differences. First, the decisions are parental rather than made by the worker. Second, the investment decision is the transmission of work habit rather than a generic unspecified investment in productivity. Third, we include taste-based discrimination as well as neighborhood effects. These differences are important but not crucial. We have in fact two major contributions as compared to the statistical discrimination literature. Let us explain them.

As stated above, we consider two types of Affirmative Action policies. The first one, which imposes a quota for the discriminated group in good jobs, is very similar to that of Coate and Loury (1993). Not surprisingly, we obtain similar results in the sense that a patronizing equilibrium emerges in steady state where employers are obliged to lower their standard in order to hire minority workers (those with bad work habits). However, contrary to Coate and Loury (1993), we are able to give the exact conditions under which such an equilibrium exists and is unique. In statistical-discrimination models

⁸Bisin and Verdier (2000) assume that parents prefer their own trait. This is a good assumption for traits like religion since it seems a good approximation to assume that Catholic parents prefer their children to be Catholic rather than belonging to another religion or another Christian denomination, similarly Muslim parents prefer their children to be Muslim, etc.

(such as that of Coate and Loury), it has never been proved that a patronizing interior equilibrium exists. They only showed that one might be created. We also formally show what happens to this intermediate equilibrium in terms of location and existence when the Affirmative Action policy is implemented. This our first main contribution.

Our second main contribution is that, contrary to Coate and Loury, we consider another Affirmative Action policy that imposes a quota for minority workers to be treated like majority workers. As stated above, this is like an anti-discrimination law where anti-discrimination requires equal treatment of everyone while the affirmative action we consider here requires equal treatment of a fraction of workers. Interestingly, we obtain exactly the opposite results than Coate and Loury and our first Affirmative Action policy. In other words, we do not obtain a patronizing equilibrium where firms are obliged to lower their standard in order to higher minority workers. On the contrary, in Proposition 4 below, we show this Affirmative Action policy (anti-discrimination law) always increases the proportion of *good* workers in equilibrium. The intuition is that, following this policy, parents' expectations become more favorable and thus they increase their effort to transmit the good trait to their children. This, in turn, leads to a higher proportion of individuals with good work habits in the minority group. As a result, because of this policy, minority workers are more likely to have good work habits and employers are more likely to maximize their profits and change their negative beliefs about minority workers. This means that the first Affirmative Action policy has to be implemented forever to help minority workers since employers keep their negative beliefs while the second one can be efficient even if implemented once since it changes the perception of employers about the work habits of minority workers. As with the other policy, we are able to give the exact conditions under which such an equilibrium exists and is unique.

To summarize, our main contribution is to propose a dynamic model combining cultural transmission and discrimination. The model can explain why some (inner-city) neighborhoods are populated by people with bad working habits (the "ghetto culture". In our framework, this is the result of a combination of discrimination, low investment in work ethic from parents and bad peers. We are also able to study different policy interventions and study how they affect the long-run outcomes.

Our findings are supported by the recent empirical literature investigating the effects of culture on labor outcomes. Focusing on Switzerland, Brügger et al. (2009) study how unemployment is affected by differences in culturally determined attitudes towards work within a narrowly defined geographic area. Their findings indicate that differences in culture explain differences in unemployment duration in the order of 20 percent. Algan and Cahuc (2005) and Alesina and Giuliano (2010) also investigate the role of "family culture" on labor market outcomes. These studies find that strong family ties reduce labor force participation. Ichino and Maggi (2000) study cultural differences in the propensity to shirk (absenteeism and misconduct) using data from a large Italian

bank. They also find strong effects.⁹ Other empirical studies have looked at the role of culture in explaining how social norms affect unemployment outcomes. Stutzer and Lalive (2004) use a novel measure of the social norm to work: The percentage of citizens in a community who voted in favor of a reduction of unemployment benefits in a Swiss referendum. They find that a one standard deviation increase in the strength of the social work norm on average translates into a reduction of unemployment duration by approximately eleven days. Using British data, Clark (2003) finds similar results: The well-being of the unemployed is shown to be strongly positively correlated with reference group unemployment (at the regional, partner, or household level).

2 The model

There is a continuum of workers who belong either to the majority group (M) or to the minority group (m). Apart from an *observable* trait (for example the place where they come from, the color of their skin or their accent), majority and minority workers are totally identical. There is also an *unobservable* trait that determines their behavior on the job. In particular, we assume that workers have either a “good” (g) or “bad” (b) work habit and are referred to as “good” and “bad” workers.

The employer can observe the group to which the worker belongs (majority or minority) but not the type (“good” or “bad”). Time is discrete and each active worker is randomly matched with an employer. The employer decides whether or not to hire this worker. If a worker of type i ($i = b, g$) is hired, the payoff to the principal is Π_i . The payoff is 0 if the worker is not hired. We assume that $\Pi_g > 0 > \Pi_b$. Irrespective of their type, employed workers earn $w > 0$. Unemployment leads to a payoff which is, without loss of generality, normalized to zero.

As stated in the Introduction, we study the intergenerational transmission of work-habit traits using an overlapping generation model. The way in which this trait is transmitted is through an education and peer-imitation process that depends on parents’ investment in the trait and the social environment where children live. The transmission of the trait is here modeled as a mechanism that interacts socialization *inside* the family (*vertical* socialization) with socialization *outside* the family (*oblique* socialization) via imitation and learning from peers and role models as in Bisin and Verdier (2000, 2001).

Children’s preferences are shaped, via education, by their parents. We assume that teaching good work ethics is costly and that a parent chooses an education effort, τ , possibly zero, so that, with a probability equal to the education effort, education will be successful and the child will have a good work ethic. Otherwise, the child remains naive, without any working trait, and is matched to a group of peers from which she

⁹See Guiso et al. (2006) for an overview of this literature and an interesting discussion of how culture affects outcomes.

learns and adopts the good trait with probability $f(q; r)$. Following Saez Marti (2011) we assume that

$$f(q; r) = \frac{rq^\alpha}{rq^\alpha + (1-q)^\alpha}, \alpha > 0, r > 0. \quad (1)$$

where q is the proportion of agents in the peer group with good habits, α captures the degree of conformism and r reflects the children's own valuation of the traits. It is easily verified that

$$\frac{\partial f(q; r)}{\partial q} > 0, \quad \frac{\partial f(q; r)}{\partial r} > 0 \quad (2)$$

which means that the higher is q , the fraction of individuals with good traits in the peer group or r , the children's own valuation of the trait, the higher is $f(q; r)$. In order to see the role of r , assume that either $\alpha = 0$, so that popularity is irrelevant or $q = 1/2$ so that both traits are equally popular. In these two cases, $f(q; r)$ takes value $r/(1+r)$ and the child would adapt the good trait with a probability which is increasing in r and equal to $1/2$ when $r = 1$, i.e. when the child finds both traits equally attractive and the population distribution of traits has no influence. More generally, if $r > 1$ ($r < 1$) children have an "innate" preference for good (bad) working habits. Note that $f(q; r)$ has three fixed points: 0, $q(r)$ and 1, where

$$q(r) = \frac{r^{\frac{1}{1-\alpha}}}{1 + r^{\frac{1}{1-\alpha}}}$$

is the unique interior q that solves $f(q; r) = q$. It is easy to see that $q(r)$ is decreasing in r and $q(1) = 1/2$.

Since we are interested in conformism, we will hereafter assume that $\alpha \geq 1$. Note that Bisin and Verdier (2000, 2001) is a special case when $r = \alpha = 1$ so that $f(q; r) = q$. Figure 1 displays $f(q; r)$ for $\alpha > 2$ (conformism) and three different values of r : $1/2$ (dotted line), 1 (solid) and 2 (dashed). We have also plotted the 45 degree line, which determines the fixed points. Note that, irrespective of r , 0 and 1 are always fixed points. There is also an interior fixed point $q(r)$ with $q(2) < q(1) = 1/2 < q(1/2)$.

Let p denote the probability that a child acquires the good trait. We have:

$$p = \tau + (1 - \tau)f(q; r) \quad (3)$$

Indeed, the child will have a good work habit if her parents' transmission is successful (with probability τ) or if it is unsuccessful (with probability $1 - \tau$) but learns the good habits from her peers (with probability $f(q; r)$).

Let $C : [0, 1] \rightarrow \mathbb{R}$ be the cost function when parents choose effort τ , with $C(0) = 0$, $C'(0) = 0$, $C'(\tau) > 0$ for all $\tau > 0$, and $C''(\tau) \geq 0$. We assume that all parents, irrespective of their type and group, face the same costs. This assumption can easily be relaxed without affecting the main qualitative results of the paper. In this model,

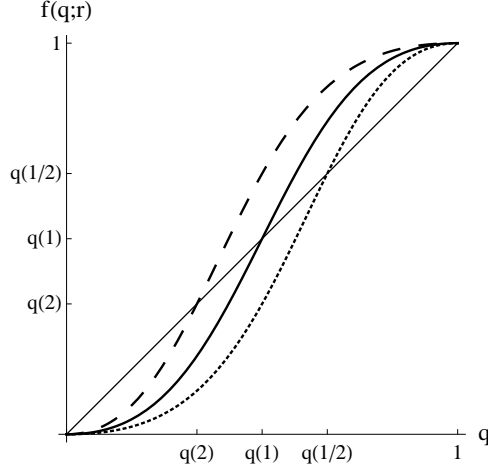


Figure 1: Rest points of $f(q; r)$ for $\alpha = 2$ and $r = 1/2, 1, 2$

the only reason why different parents may be exerting different efforts will be because either they have different expectations concerning their children's future employment prospects and/or their children have different peer groups. Parents choose effort $\tau \in (0, 1]$ that maximizes

$$pV^g + (1 - p)V^b - C(\tau) \quad (4)$$

where V^i is the value a parent attaches to her child being of type i ($i = b, g$) and p is given by (3). We obtain the following first-order condition:

$$C'(\tau) = \frac{dp}{d\tau}V^g - \frac{dp}{d\tau}V^b. \quad (5)$$

By substituting (3) in (5), we easily obtain the optimal education effort:

$$C'(\tau^*) = (1 - f(q; r))U$$

where

$$U \equiv V^g - V^b \geq 0 \quad (6)$$

This implies that:

$$\tau^* = \tau((1 - f(q; r))U) = C'^{-1}((1 - f(q; r))U). \quad (7)$$

Note that the assumptions on C guarantee that $\tau(0) = 0$ and $\tau' > 0$. In order to avoid corner solutions we assume that $C'(1)$ is large enough (relative to the maximum value U can take). Parents will exert no effort either when having the good trait does not give any advantage (i.e., $U = 0$) or when everybody in the neighborhood has trait g (i.e., $q = 1$). Parents' decision depends on society since parents have less incentive to

socialize their children the easier it is for them to learn the good trait from society, namely the larger is $f(q; r)$. Note also that there is cultural substitutability since the lower is q , the fraction of individuals with good traits in the peer group, or r , the innate preference for good trait, the higher effort parents will put in transmitting the good trait.

2.1 Dynamics

We assume a Poisson birth and death process that keeps the population size of active workers constant: with probability λ an active worker will be active in the next period, with probability $1 - \lambda$ and active worker has a child who becomes active the period after. The proportion of *good* workers in the next period is equal to the proportion of *good* workers who survived from the current period (λq) plus all new-born *good* children, i.e. the children born with *good* parents who adopt *good* work habits ($(1 - \lambda)q p$) plus the children of *bad* parents who adopt *good* work habits ($(1 - \lambda)(1 - q) p$), that is

$$\lambda q + (1 - \lambda)q p + (1 - \lambda)(1 - q) p.$$

Note that we have assumed that the probability of adopting a good work habit is independent of the parents' types since all parents value the "good" trait equally. This is an important difference with Bisin and Verdier (2000, 2001). Substituting (3) and (7) and subtracting q we get a difference equation, $\Delta_q(q, U, r)$ describing the change in q :

$$\Delta_q(q, U, r) = (1 - \lambda)(f(q; r) - q) + (1 - \lambda)\tau((1 - f(q; r))U)(1 - f(q; r)). \quad (8)$$

Let $R(U, r)$ denote the set of rest points of (8), i.e., the q 's which solve $\Delta_q(q, U, r) = 0$. It is easy to see that 1 is always a rest point, i.e. $\Delta_q(1, U, r) = 0$ for all U . Moreover for low enough U , (8) has three rest points, two of them interior when $U > 0$. In fact

$$\tau((1 - f(q; r))U)(1 - f(q; r)) = q - f(q; r), \quad (9)$$

has three solutions. In order to see this observe that the left hand side is positive and monotonically decreasing in q (for $U > 0$) while the right hand side is positive and has an inverse U-shape in $[0, q(r)]$. This guarantees that for low enough U , there are three solutions, q^s and q^u ($q^s \leq q^u$) and 1. Moreover, q^s is increasing while q^u is decreasing in U . We state the main results in the following proposition:

PROPOSITION 1. *Assume that $U \geq 0$. For each $r > 0$, there exists an $X(r) > 0$ such that*

(i) *for all $U > X(r)$,*

$$R(U, r) = \{1\},$$

(ii) for all $U \in [0, X(r)]$,

$$R(U, r) = \{q^s(U; r), q^u(U; r), 1\}$$

$$q^s(0; r) = 0, q^u(0; r) = q(r), q_1^s(U; r) > 0, q_1^u(U; r) < 0.$$

$X(r)$ is such that $q^s(X(r); r) = q^u(X(r); r)$. The rest points 1 and $q^s(U; r)$ are stable, $q^u(U; r)$ is unstable.

Proposition 1 shows that for small enough U there two stable equilibria 1 and $q^s(U; r)$, the latter disappears for large enough U . Observe that $q^s(U; r)$ is increasing in U . This explains the main mechanism of the model since the results are driven by parents' valuations of the future prospects of their children. If they expect that these are bad (i.e. U small, i.e, having a good trait does pay little), they will exert little effort and $q^s(U; r)$ is low. If they expect very large benefits of having a good trait, then $q^s(U; r)$ is larger and could eventually cease to exist. In that case the only rest point is 1.

The following figures illustrate these results. Figure 2 shows how the rest points of (8) change with U for a given r . For all parameters in the gray (white) area, $\Delta_q(q, \cdot, r) < 0$ ($\Delta_q(q, \cdot, r) > 0$). In this figure, the solid thick line corresponds to the stable rest point $q = 1$, the solid thick curve to the stable rest point $q^s(U; r)$ and the dotted thick curve to the unstable rest point $q^u(U; r)$. For small enough values of U (i.e. $U < X(r)$), there are three rest points, two stable (1 and $q^s(U; r)$) and one unstable ($q^u(U; r)$). The rest points $q^s(U; r)$ and $q^u(U; r)$ cease to exist when $U > X(r)$.

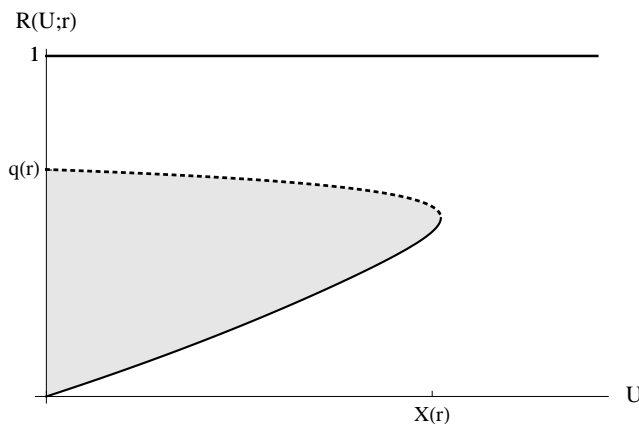


Figure 2: Steady states

Figure 3 shows how the rest points change with r , in the figure $r < r'$. It is easy to see that increasing the children's relative value of the good trait (r), given U , increases the proportion of good workers in the stable rest point, $q = q^s(U; r)$, and reduces $X(r)$,

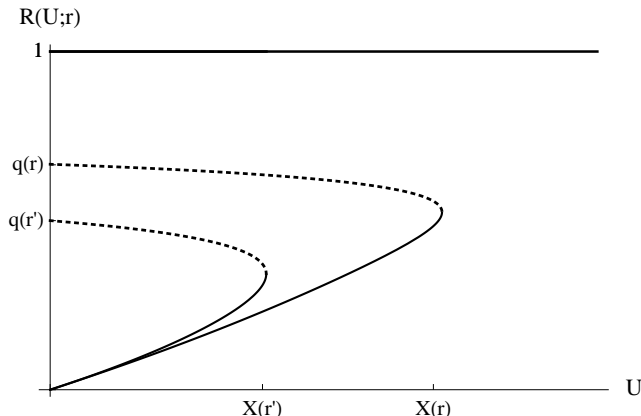


Figure 3: Steady states

which, in turn, implies that a unique stable equilibrium for which $q = 1$ is more likely to emerge.

This figure makes it clear that there are two obvious policies that can reduce the prevalence of the bad trait: changing children attitudes, r , and changing parent's evaluations, U . We next endogenize U . In order to do so we need to analyze the employers' hiring behavior.

2.2 Endogenous evaluations

Assume that each time an employer meets a *bad* worker, she knows her type with probability α . With probability $1 - \alpha$, she (incorrectly) believes that the worker is *good*.¹⁰ A *good* worker is never mistaken for a *bad* one. Our results are robust to the case of a more general screening technology, which allows for the mistaken classification of good workers as long as the probability of making a mistake about bad workers is larger than the probability of making a mistake about good ones. We assume that employers know the actual proportion q of *good* agents.

When the worker and the employer are matched, the employer chooses one of the two following strategies: (i) screening (s) the workers and hire only seemingly *good* workers, i.e. all *good* workers and some *bad* ones who have been (mistakenly) taken for *good* ones and (ii) hire nobody (n).¹¹

Employers prefer strategy s to strategy n when

¹⁰The probability α could be group dependent and written as α_k , for $k = m, M$. This extension would not change any of our results but will unnecessarily complicate the analysis.

¹¹As in a previous version of the paper, we could have assumed, as in Hauk and Saez Marti (2002), that there were two tasks instead of two types of employment statuses. Task 1 is a more complicated and better-paid task while task 2 is a low-paid task. In that case, in the screening policy, the employer

$$q\Pi_g + (1 - q)(1 - \alpha)\Pi_b \geq 0.$$

We can rewrite this inequality as follows:¹²

$$q \geq \frac{-(1 - \alpha)\Pi_b}{\Pi_g - (1 - \alpha)\Pi_b} \equiv \tilde{q} \quad (10)$$

where $\tilde{q} \in (0, 1)$. If the proportion of *good* workers is sufficiently high ($q > \tilde{q}$), then the screening strategy s is optimal. We denote the optimal strategy by μ :

$$\mu(q) = \begin{cases} s & \text{if } q \geq \tilde{q} \\ n & \text{if } q < \tilde{q} \end{cases} \quad (11)$$

and define the following index function:

$$I(q) = \begin{cases} 1 & \text{if } \mu(q) = s \\ 0 & \text{if } \mu(q) = n. \end{cases} \quad (12)$$

In order to compute the child's well-being, a parent needs to form expectations concerning the child's future job opportunities. We assume that parents prefer their children to have preferences leading to decisions with a higher expected income. Let the vector $\boldsymbol{\pi} = \{\pi_z\}_{z=1}^{\infty}$ denote the expectations a parent has concerning his child future employment opportunities, where π_z denotes the probability that the child being educated today is matched z periods later with a screening employer (with probability $(1 - \pi_z)$ the child is expected to meet an employer who will not hire). Let $w > 0$ be the wage when hired, λ the probability that an active worker in a given period will still be active in the next period and $V^i(\boldsymbol{\pi})$ the expected earnings of a worker of type i when the parents' expectation is $\boldsymbol{\pi}$. It is easy to see that

$$U(\boldsymbol{\pi}) = V^g(\boldsymbol{\pi}) - V^b(\boldsymbol{\pi}) = \alpha w \sum_{z=1}^{\infty} \pi_z \lambda^{z-1} \geq 0. \quad (13)$$

When unemployed, *good* and *bad* workers have zero income and when screened *good* workers are always employed and *bad* workers are employed with probability $(1 - \alpha)$. The difference between these payoffs leads to (13); the inequality will be strict whenever $\pi_s > 0$ for some s .

In order to study the effect of discrimination we assume that a proportion θ^p of employers are profit maximizing and act according to $\mu(q)$ (see 11) and a proportion θ^d ($\theta^d = 1 - \theta^p$) discriminate some workers (from the minority) and do not hire them even if this would maximize their profits.

will offer task 1 to seemingly good workers while in the pooling strategy, she will offer task 2 to everyone. The results will obviously be the same.

¹²Remember that $\Pi_b < 0$.

The long run equilibrium will now depend on the initial condition and on parents' expectations which depends on firms' behavior. We focus on stationary states under rational expectations:

DEFINITION 1. *Assume that $q_t = q^*$ and that the proportion of profit maximizing employers is θ^p . The state q^* is a stable steady state under rational expectations and profit maximizing behavior if*

(i) *All non-discriminating employers choose $\mu(q^*)$ for all $t' \geq t$ while all other firms always choose the strategy n , i.e. hire no minority workers.*

(ii)

$$\pi_z = \theta^p I(q^*), \quad z = 1, 2, 3.. \quad (14)$$

where $I(q^*)$ is given by (12), and

(ii) q^* is a stable rest point of $\Delta_q(q, U(\boldsymbol{\pi}), r)$ where $\boldsymbol{\pi}$ is given by (14) and $U(\boldsymbol{\pi})$ by (13).

Let $SS^E(\theta^p)$ be the set of stable rest points under rational expectations when the proportion of profit-maximizing employers is θ^p . We can now state the following proposition:

PROPOSITION 2. *Let $\boldsymbol{\theta}^p = \{\theta^p, \theta^p, \theta^p \dots\}$. Then,*

$$SS^E(\theta^p) = \{0, q^s(U(\boldsymbol{\theta}^p); r), 1\} \quad (15)$$

whenever

$$U(\boldsymbol{\theta}^p) = \theta^p \frac{\alpha w}{(1 - \lambda)} < X(r)$$

and $q^s(U(\boldsymbol{\theta}^p); r) > \tilde{q}$. Otherwise,

$$SS^E(\theta^p) = \{0, 1\}. \quad (16)$$

Figure 4 shows how the set of steady states changes with θ^p . As in the previous figures, the steady-state equilibria are depicted by the thick solid curves. When $\theta^p = \tilde{\theta}^p$, $q^s(U(\tilde{\boldsymbol{\theta}}^p); r) = \tilde{q}$. If the proportion of profit maximizing employers is low (i.e. $\theta^p < \tilde{\theta}^p$), the interior steady-state equilibrium with both types of agents (good and bad) does not exist. The same is true for large enough θ^p 's provided that $U(\mathbf{1}) > X(r)$ (case depicted in Figure 4). If the proportion of discriminating employer is too high, minority parents rationally expect that their children will be unemployed and thus do not put any effort in transmitting the good trait.

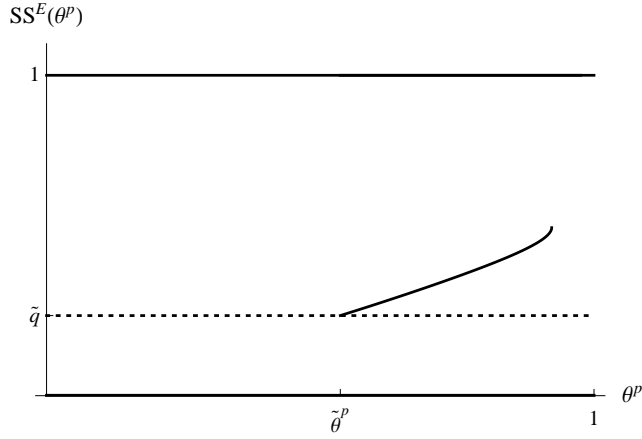


Figure 4: Steady states

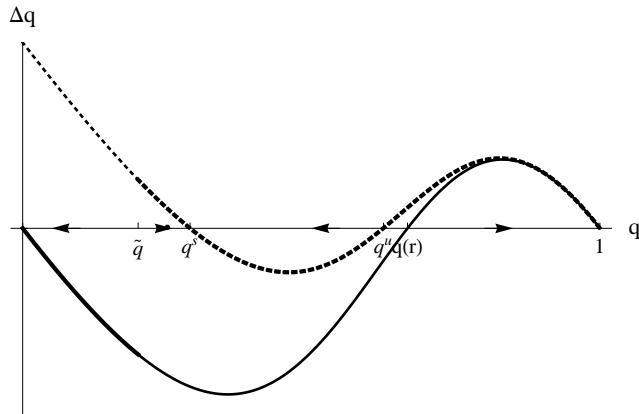


Figure 5: Dynamics

Figure 5 illustrate how initial conditions affect the steady state equilibrium which is actually reached. In this figure, we have taken a θ^p such that $q^s(U(\theta^p); r) > \tilde{q}$ (i.e. $\theta^p > \tilde{\theta}^p$) so that $SS^E(\theta^p) = \{0, q^s(U(\theta^p); r), 1\}$. The solid curve corresponds to $(q - f(q; r))$. i.e $\Delta_q(q, 0, r)$. The dashed curve corresponds to $\Delta_q(q, U(\theta^p), r)$. When $q^s > \tilde{q}$ and $q_t > \tilde{q}$, it is rational to expect θ^p and the dynamics follow the dotted thick curve. Depending on the actual value of q_t , the economy converges either to 1 (when $q_t > q^u$) or to q^s (when $q_t \in [\tilde{q}, q^u]$). If $q_t < \tilde{q}$, it is then profit maximizing to use the pooling assignment and the relevant dynamics are those depicted by the thick solid curve. In that case, the economy converges to 0 since there are no incentives for parents to put effort. Note that, along the convergence path, the non-discriminatory employers are always maximizing their profits.

The following proposition states these claims formally:

PROPOSITION 3. *Let $\theta^p = \{\theta^p, \theta^p, \theta^p \dots\}$ and assume that $U(\theta^p) < X(r)$. There exist (stationary) rational expectations such that*

- (i) $\lim_{t \rightarrow \infty} q_t = 0$ if $q_0 < \tilde{q}$
- (ii) $\lim_{t \rightarrow \infty} q_t = q^s(U(\theta^p); r)$ if $q_0 \in [\tilde{q}, q^u]$, and
- (iii) $\lim_{t \rightarrow \infty} q_t = 1$ if $q_0 > q^u(U(\theta^p); r)$

whenever $q^s(U(\theta^p); r) > \tilde{q}$, otherwise

- (i) $\lim_{t \rightarrow \infty} q_t = 0$ if $q_0 < q^u(U(\theta^p); r)$
- (ii) $\lim_{t \rightarrow \infty} q_t = 1$ if $q_0 > q^u(U(\theta^p); r)$

Note that Proposition 3 describes also the dynamics of the majority. In this case, since workers belonging to this group are not discriminated, $\theta^p = 1$. The only differences between majority and minority workers are given by the location of the interior rest points, i.e. $q^s(U(\mathbf{1}); r) > q^s(U(\theta^p); r)$ and $q^u(U(\mathbf{1}); r) < q^u(U(\theta^p); r)$, and on the initial conditions. To summarize, since $q^s(U(\theta^p); r)$ is increasing in U while $q^u(U(\theta^p); r)$ is decreasing in U , the long run distribution of traits depends not only on the initial conditions (composition of the peer group) but also on parents' expectations and employers prejudices. This could explain why some groups may end up having lower working ethics. Small differences in r and or in U may end up having large long run effects.

In the next sections, we study the effect of policies that can break down this vicious cycle of negative attitudes and behavior that lead to unemployment and bad working habits among minority workers. As stated above, it should be clear that policies can reduce unemployment and the prevalence of the bad trait in the minority population by changing children attitudes, r , and parents evaluations, U .

3 Affirmative Action

Let us start by considering an Affirmative-Action policy that consists of giving a preferential treatment to discriminated groups, for example by imposing minimum hiring quotas to firms.¹³

Let us focus on minority workers and assume first that q is such that the non-discriminatory employers are not willing to hire any minority worker (and neither are the discriminatory ones). In that case $U = 0$, parents exert no effort and the long-run

¹³For an overview and evaluation on Affirmative Action policies in the United States, see Holzer and Neumark (2000, 2006).

equilibrium is $q^* = 0$. Assume now that *all* employers are forced to hire a proportion ϕ of minority workers (*Affirmative Action policy I*). This policy will have an effect on the proportion of good workers only if it changes the parents' evaluation of the traits. Since employment possibilities are independent of type, i.e., both types get the same expected wage, the parents's valuation, U , is zero and the Affirmative Action policy has no effect on the equilibrium.

Consider now an Affirmative Action policy that requires the screening assignment for a quota ϕ of workers (*Affirmative Action policy II*).¹⁴ Observe that Affirmative Action policy II is just a general case of an *anti-discrimination law* where anti-discrimination requires equal treatment of everyone and the affirmative action we consider here requires equal treatment of a fraction of workers. If parents expect this policy to be implemented, they will exert a positive effort since the evaluation of the trait is now:

$$U(\phi) = \phi \frac{\alpha w}{(1 - \lambda)} > 0$$

where $\phi = \{\phi, \phi, \dots\}$. This policy has the potential of being effective since good workers have a higher probability of being employed.

Assume now that the first policy is introduced when all non-discriminatory employers are already hiring some minority workers (those who seem to be good) and parents evaluation is given by $U(\theta^p)$. The introduction of Affirmative Action policy I changes the evaluation of the parents to $U(\phi\theta^p)$, with

$$U(\phi\theta^p) < U(\theta^p),$$

where $\phi\theta^p = \{\phi\theta^p, \phi\theta^p, \dots\}$. This policy has clearly a negative long-run effect since it reduces the equilibrium value of good workers (q^s is increasing in U). The second policy (Affirmative Action policy II) will, instead, have a positive effect because it increases the value of a good child, above the one without intervention, i.e.,

$$U(\theta^p + \phi\theta^d) > U(\theta^p).$$

where $\theta^d = 1 - \theta^p$ and $\theta^p + \phi\theta^d = \{\theta^p + \phi\theta^d, \theta^p + \phi\theta^d, \dots\}$. Without any Affirmative Action policy, all non-prejudiced employers will employ seemingly *good* workers, which gives an advantage to the good workers (since bad workers are detected with positive probability). Under the first Affirmative Action policy, the advantage of being “good” is smaller and, as a result, parents put less effort in transmitting the good work habit trait. This leads to a lower equilibrium value and could possibly lead to the disappearance

¹⁴The main difference between these two Affirmative Action policies is that, in the first one, employers are obliged to hire $\phi\%$ of their workers from the minority group but *cannot* test them. So whether the worker is “good” or “bad” is irrelevant in the employment process and “good” and “bad” workers have the same chance of being hired. In the second Affirmative Action policy, employers are still obliged to hire $\phi\%$ of their workers from the minority group but *can* test them. As a result, only “good” workers and seemingly “good” workers with bad work habits will be hired.

of the interior equilibrium $q^s(U(\phi\theta^p))$ if the economy is sufficiently close to \tilde{q} . Instead, when the second Affirmative Action policy is implemented, a larger share of workers are automatically screened and the return is higher for good workers leading to an increase in the proportion of workers with good habits. We can now summarize these findings in the following proposition.

PROPOSITION 4. *Affirmative Action policy II always increases the equilibrium proportion of good workers. Affirmative action I has a negative effect if introduced when the equilibrium proportion of good workers is positive, otherwise it has no effect on the proportion of good workers. This policy always reduces employers' profits.*

Our results concerning Affirmative Action I are similar to that of Coate and Loury (1993). In their paper, Affirmative Action is modeled as a government-mandated constraint on employers, requiring them to assign workers from each group to more rewarding jobs at the same rate. They show that Affirmative action may sometimes fail because, to comply with the Affirmative-Action policy, employers must lower their standard for assigning the workers that they have negative views about to good jobs. This is referred to as a patronizing equilibrium. Lowering the standard may reduce investment incentives because the favored workers see themselves as likely to succeed without acquiring the relevant skills. Thus, employers' negative stereotypes can continue to be confirmed in equilibrium under Affirmative Action. Coate and Loury show that this equilibrium is more likely to exist if the proportion of these workers is relatively rare in the population. Even if the mechanism is different, this result is close to ours when the first Affirmative Action policy is implemented. In our case, compared to the equilibrium without Affirmative Action, parents put relatively less effort into transmitting the good trait because a fraction of workers, irrespective of their type, are employed with certainty.

Observe, however, that, compared to Coate and Loury (1993), we are able to give the exact condition under which such a patronizing equilibrium exists. In our model, without any policy intervention, if $U(\theta^p) < X(r)$ and $q^s(U(\theta^p); r) > \tilde{q}$ (or equivalently $\theta^p < \tilde{\theta}^p$), there are three stable steady-state equilibria: 0, 1 and $q^s(U(\theta^p); r)$ (see Proposition 3). In this section, we have shown that the Affirmation Action policy I always decreases the proportion of good workers by changing in an unfavorable way parents' expectation and thus by decreasing parents' effort. This means, in particular, that $q^s(U(\theta^p); r) > q^s(U(\phi\theta^p); r)$. We can give the new conditions under which the intermediate patronizing equilibrium exists: $U(\phi\theta^p); r < X(r)$ and $q^s(U(\phi\theta^p); r) > \tilde{q}$.

Observe also that, contrary to Coate and Loury (1993), we are able to analyze another Affirmative Action policy (or anti-discrimination law) that imposes a quota for minority workers to be treated like majority workers. We obtain exactly the opposite results than Coate and Loury and our first Affirmative Action policy in the sense that a patronizing equilibrium does not exist. In Proposition 4, we show this Affirmative Action policy always increases the proportion of *good* workers in equilibrium because

this policy changes parents' expectations so that they increase their effort to transmit the good trait to their children. This, in turn, leads to a higher proportion of individuals with good work habits in the minority group. As a result, because of this policy, minority workers are more likely to have good work habits and employers are more likely to maximize their profits and change their negative beliefs about minority workers. Note that, even if we start at $q_0 = 0$, i.e. all minority workers have bad work habits, introducing this Affirmative Action policy (or anti-discrimination law) will lead to an equilibrium for which $q > 0$. If the quota ϕ is large enough, we may even end up with an equilibrium q for which $q > \tilde{q}$ so that firms do not make losses and maximize their profit. Here also we can give the conditions under which the intermediate equilibrium exists: $U(\theta^p + \phi\theta^d) < X(r)$ and $q^s(U(\theta^p + \phi\theta^d); r) > \tilde{q}$. This implies that the first Affirmative Action policy has to be implemented forever to help minority workers since employers keep their negative beliefs while the second one can be efficient even if implemented temporarily since it may change minority workers' working habits.

4 Integration

Racial integration is a very sensitive and highly debated policy in the United States.¹⁵ For example, the Moving to Opportunity" (MTO) program, which aim is to move very poor households to richer areas, can be considered as an integration policy.¹⁶ Our model can shed some light on this controversial debate. Assume now that children from one group are also influenced by peers from the other group, and let $\sigma_m \in [0, 1/2]$ ($\sigma_M \in [0, 1/2]$) be the parameter capturing the extent of integration of the minority (majority) group, so that if $\sigma_i = 0$ children of group i have no contact with those of the other group, while if $\sigma^i = 1/2$ both groups of peers are equally important for the i -children. We can now write the oblique transmission function of group i as

$$f_i(\mathbf{q}; r) = f(\sigma_i q^j + (1 - \sigma_i) q^i; r), \text{ for } i, j = m, M \quad (17)$$

where $\mathbf{q} = (q_m, q_M)$.

¹⁵For instance, in 1974, federal judge W. Arthur Garrity ordered that Boston's schools be integrated through forced busing (black children were driven by bus to white schools). Twenty-five years later, in June 1999, facing pressure from a lawsuit by white parents and advocates of neighborhood schools, the city's school board voted 5-2 to stop the busing policy and adopt a race-blind admissions policy starting in September 2000 (Education Week, 08/04/99 edition, by Caroline Hendrie).

¹⁶By giving housing assistance (i.e. vouchers and certificates) to low-income families, the MTO programs help them relocate to better and richer neighborhoods. The results of most MTO programs (in particular for Baltimore, Boston, Chicago, Los Angeles and New York) show a clear improvement of the well-being of participants and better labor market outcomes (see, in particular, Ladd and Ludwig, 2001, Katz et al., 2001, Kling et al., 2005, Rosenbaum and Harris, 2001). Observe that the MTO programs do not specifically target minority families (such as blacks) but poor families. But since the two are correlated, this is a good example of an integration policy.

The dynamics of q_n and q_m are jointly determined and can be written as

$$\Delta_{q_m}(\mathbf{q}, U^m) = (1 - \lambda)(f_m(\mathbf{q}; r) - q_m + \tau(U^m, \sigma_m q^M + (1 - \sigma_m)q^M; r))(1 - f_m(\mathbf{q}; r)) \quad (18)$$

and

$$\Delta_{q_M}(\mathbf{q}, U^M) = (1 - \lambda)(f_M(\mathbf{q}; r) - q_M + \tau(U^M, \sigma_M q^m + (1 - \sigma_M)q^m; r))(1 - f_M(\mathbf{q}; r)) \quad (19)$$

Assume that integration occurs when $q^m = 0$ ($U^m = 0$) and $q^M = 1$ ($U^M > 0$). It is easy to see that $(0, 1)$ is not a rest point of (18)-(19) since at $\mathbf{q} = (0, 1)$

$$\Delta_{q_m}((0, 1), U^m) = (1 - \lambda)(f(\sigma_m; r)) > 0 \quad (20)$$

and

$$\Delta_{q_M}((0, 1), U^M) = (1 - \lambda)(f(1 - \sigma_M; r) - 1)(1 - \tau(U^M, 1 - \sigma_M; r)) < 0 \quad (21)$$

and integration has an immediate positive effect on the minority workers and a negative effect on the majority. If integration occurs when $q_m = 0$ ($U^m = 0$) and $q_M = q^s(U^M; r)$ ($U^M > 0$), its immediate effect is again positive for the minority group since

$$\Delta_{q_m}((0, q^s(U^M; r)), U^m) = (1 - \lambda)(f(\sigma_m q^s(U^M; r); r)) > 0. \quad (22)$$

The effect on the majority group is ambiguous. We cannot sign $\Delta_{q_M}((0, q^s(U^M; r)), U^M)$ without knowing to what extent parents' efforts compensate the effect of a worse peer group.

The following figure show the dynamics in the space (q_m, q_M) . We have assumed that if $q_i \geq \tilde{q}$ parents expect the separating assignment and, below that value, they expect the pooling assignment. We also assume that, because of discrimination, under the separating policy, the minority workers are worse off than the majority ones and therefore $q_M^s > q_m^s > 0$ and that $r = 1$ so that children view both traits as similar. The left hand side panel corresponds to $\sigma^m = \sigma^M = 0$. In that case, depending only on its initial value, q_i converges to 0, q_i^s or 1, its dynamics being independent of q_j . On the left hand panel, we make $\sigma^m = \sigma^M > 0$. It is easy to see that integration maybe beneficial for both groups in the long run but not in the short term.

This could explain why the different integration policies implemented in the US and in Europe¹⁷ seem to have failed. Indeed, mixing policies, which include school busing, forced integration of public housing, and laws barring discrimination in housing and

¹⁷For instance, the creations of Zones of Educational Priority (ZEP) and the rehabilitation of bleak housing projects in immigrant neighborhoods under the guise of urban policy ('politique de la ville') in France had very limited effects. See, for example, Benabou et al. (2009) for an evaluation of the ZEP and Brubaker (2001) who compares the different ways of assimilating ethnic minorities in France, Germany and the US.

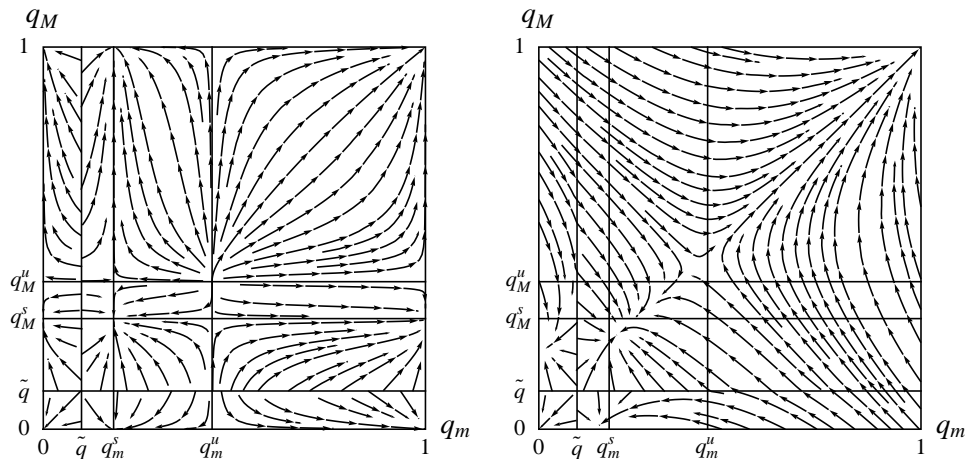


Figure 6: Integration. $\alpha = 2$, $r = 1$

employment,¹⁸ have often had limited effects and are often opposed by the majority groups (see e.g., Jacoby, 1998, and Thernstrom and Thernstrom, 2002).

Interestingly, Chaudhuri and Sethi (2008), who incorporate neighborhood effects into an otherwise standard statistical discrimination model, find a similar result, even though the mechanism is totally different. In their paper, increasing integration tends to lower the costs of human acquisition for B-workers while raising these costs for A-workers. Thus, if integration proceeds sufficiently far, the authors show that negative stereotypes cannot be sustained.

Guryan (2004) shows that the desegregation plans that have been implemented in American schools in the last forty years have mainly benefited the black students by reducing their high school drop out rates while they had no effect on the dropout rates of the other students. Peer effects are shown to be one of the main explanations for this result. Studying the Metco program, a long-running desegregation program that mostly sends Black students out of the Boston public school district to attend schools in more affluent suburban districts, Angrist and Lang (2004) find similar results.¹⁹

¹⁸See Lang (2007) for an overview of these policies in the U.S.

¹⁹There is also a growing literature in the fields of public finance, development and urban economics that shows that investments in public goods, tastes for redistribution, and other forms of civic behavior are less common in racially or ethnically diverse communities (see, in particular, Alesina et al., 1999, Alesina and La Ferrara, 2000, Luttmer, 2001, Vigdor, 2004).

5 Concluding remarks

We have introduced a dynamic model of cultural transmission to explain different outcomes for minority and majority workers. We have shown that if the proportion of taste-based prejudiced employers is sufficiently high, prejudices can be confirmed in equilibrium. Otherwise, multiple equilibria exist, with and without discrimination. We have also studied different policies aiming at reducing discrimination. Both Affirmative Action²⁰ and integration policies may work. The mechanisms through which these two policies affect the quality of the workers are different, though. Affirmative action policies directly affect the expected payoff of the different types of workers and the parents' incentives to invest in those traits. By "improving" the quality of the peers with whom minority children interact, integration policy has a positive effect on those workers with a worse peer group. The opposite happens for the other children since, after integration, they interact with a "worse" quality peer group. From a political economy perspective, it is likely that all workers will support the Affirmative Action policy while only families from bad neighborhoods may favor the integration policy. As far as employers are concerned, it seems plausible that they may object to Affirmative Action policies that impose too small quotas. The reason for this opposition is that they are forced to offer contracts that are suboptimal given the average composition of workers. When the Affirmative Action quotas are sufficiently high, both employers and workers benefit from the policy.

References

- [1] Alesina, A., Baqir, R. and W. Easterly (1999), "Public goods and ethnic divisions," *Quarterly Journal of Economics* 114, 1243-1284.
- [2] Alesina, A. and P. Giuliano (2010), "The power of the family," *Journal of Economic Growth* 15, 93-125.
- [3] Alesina, A. and E. La Ferrara (2000), "Participation in heterogeneous communities," *Quarterly Journal of Economics* 115, 847-904.
- [4] Algan, Y. and P. Cahuc (2005), "The roots of low European employment: Family Culture?" In: J. Frenkel and C.A. Pissarides (Eds.), *NBER International Seminar on Macroeconomics*, Boston: MIT Press.
- [5] Altonji, J.G. and R.M. Blank (1999), "Race and gender in the labor market", in: Ashenfelter, O. and D. Card (Eds.), *Handbook of Labor Economics, Vol.3*, Amsterdam: North Holland, pp. 3143-3259.

²⁰In this discussion, we only focus on the second Affirmative Action policy.

- [6] Angrist, J.D. and K. Lang (2004), “Does school integration generate peer effects? Evidence from Boston’s Metco program,” *American Economic Review* 94, 1613-1634.
- [7] Arnott, R. and R. Rowse (1987), “Peer group effects and educational attainment,” *Journal of Public Economics* 32, 287-305.
- [8] Arrow, K.J. (1973), “The theory of discrimination,” in O. Ashenfelter and A. Rees (Eds.), *Discrimination in Labor Markets*, Princeton: Princeton University Press, pp. 3-33.
- [9] Austen-Smith, D. and R.D. Fryer, Jr (2005), “An economic analysis of ‘acting white’,” *Quarterly Journal of Economics* 120, 551-583.
- [10] Battu, H., McDonald, M. and Y. Zenou (2007), “Oppositional identities and the labor market,” *Journal of Population Economics* 20, 643-667.
- [11] Becker, G.S. (1957), *The Economics of Discrimination*, Chicago: Chicago University Press.
- [12] Benabou, R. (1993), “Workings of a city: Location, education, and production,” *Quarterly Journal of Economics* 108, 619-652.
- [13] Benabou, R., Kramarz, F. and C. Prost (2009), “The French zones d’éducation prioritaire: Much ado about nothing?” *Economics of Education Review* 28, 345-356.
- [14] 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.
- [15] Bisin, A. and T. Verdier (2001), “The economics of cultural transmission and the dynamics of preferences,” *Journal of Economic Theory* 97, 298-319.
- [16] 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, Chap. 9, pp. 339-416.
- [17] Bisin, A., Patacchini, E., Verdier, T. and Y. Zenou (2010), “Bend it like Beckham. Ethnic identity and integration”, CEPR Discussion Paper No. 8054.
- [18] Bisin, A., Patacchini, E., Verdier, T. and Y. Zenou (2011), “Formation and persistence of oppositional identities,” *European Economic Review* 55, 1046-1071.
- [19] Bisin, A., Topa, G. and T. Verdier (2004), “Religious intermarriage and socialization in the United States,” *Journal of Political Economy* 112, 615-664.

- [20] Bonney, N. (1975), "Work and ghetto culture," *British Journal of Sociology* 26, 435-447.
- [21] Borjas, G. J. (1998), "To ghetto or not to ghetto: Ethnicity and residential segregation," *Journal of Urban Economics*, 44, 228-253.
- [22] Boyd, R. and P.J. Richerson (1985), *Culture and the Evolutionary Process*, Chicago: University of Chicago Press.
- [23] Brubaker, R. (2001), "The return of assimilation? Changing perspectives on immigration and its sequels in France, Germany, and the United States," *Ethnic and Racial Studies* 24, 531-548.
- [24] Cavalli-Sforza, L. and M. Feldman (1981), *Cultural Transmission and Evolution. A Quantitative Approach*, Princeton: Princeton University Press.
- [25] Chaudhuri, S. and R. Sethi (2008), "Statistical discrimination with neighborhood effects: Can integration eliminate negative stereotypes?" *Review of Economic Studies* 75, 579-596.
- [26] Clark, A. (2003), "Unemployment as a social norm: Psychological evidence from panel data," *Journal of Labor Economics* 21, 323-351.
- [27] Coate, S. and G.C. Loury (1993), "Will Affirmative-Action policies eliminate negative stereotypes?" *American Economic Review* 83, 1220-1240.
- [28] Cohen Zada, D. (2006), "Preserving religious values through education: Economic analysis and evidence from the U.S.," *Journal of Urban Economics* 60, 372-398.
- [29] Cutler, D.M, Glaeser, E.L. and J.L. Vigdor (1999), "The rise and decline of the American ghetto," *Journal of Political Economy* 107, 455-506.
- [30] De Bartoleme, C. (1990), "Equilibrium and inefficiency in a community model with peer group effects," *Journal of Political Economy* 98, 110-133.
- [31] Glaeser, E.L. Sacerdote, B., and J. Scheinkman (1996), "Crime and social interactions," *Quarterly Journal of Economics* 111, 508-548.
- [32] Guiso, L., Sapienza, P. and L. Zingales (2006), "Does culture affect economic outcomes?" *Journal of Economic Perspectives* 20, 23-48.
- [33] Guryan, J. (2004), "Desegregation and black dropout rates," *American Economic Review* 94, 919-943.
- [34] Hannerz, U. (1969), *Soulside: Inquiries into Ghetto Culture and Community*, New York: Columbia University Press.

- [35] Hauk, E. and M. Sáez-Martí (2002), “On the cultural transmission of corruption,” *Journal of Economic Theory* 107, 311-335.
- [36] Holzer, H. and D. Neumark (2000), “Assessing Affirmative Action,” *Journal of Economic Literature* 38, 483-568.
- [37] Holzer, H. and D. Neumark (2006), “Affirmative Action: What do we know?” *Journal of Policy Analysis and Management* 25, 463-490.
- [38] Ichino, A. and G. Maggi (2000), “Work environment and individual background: Explaining regional shirking differentials in a large italian firm,” *Quarterly Journal of Economics* 115, 1057-1090.
- [39] Jacoby, T. (1999), “Beyond Busing,” *The Wall Street Journal*, July 21st.
- [40] Jellal, M. and F.C. Wolff (2002), “Cultural evolutionary altruism: Theory and evidence,” *European Journal of Political Economy* 18, 241-62.
- [41] Katz, M.B. (1993), *The “Underclass” Debate: Views From History*, Princeton: Princeton University Press.
- [42] Katz, L.F., Kling, J.R. and J.B. Liebman (2001), “Moving to opportunity in Boston: Early results of a randomized mobility experiment,” *Quarterly Journal of Economics* 116, 607-654.
- [43] 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 experience,” *Quarterly Journal of Economics* 120, 87-130.
- [44] Kohn, M. (1969), *Class and Conformity*, Homewood, IL: Dorsey Press.
- [45] Kolm, A.-S. (2005), “Work norms and unemployment,” *Economics Letters* 88, 426-431.
- [46] Ladd, H.F. and J. Ludwig (1997), “Federal housing assistance, residential relocation, and educational opportunities: Evidence from Baltimore,” *American Economic Review* 87, 272-277.
- [47] Lang, K. (2007), *Poverty and Discrimination*, Princeton: Princeton University Press.
- [48] Lemman, N. (1991), *The Promised Land*, New York: Alfred A. Knopf.
- [49] Lewis, O. (1969), “Culture of poverty,” In: D.P. Moynihan (Ed.), *On Understanding Poverty: Perspectives from the Social Sciences*, New York: Basic Books, pp.187-220.

- [50] Lindbeck, A. and S. Nyberg (2006), "Raising children to work hard: Altruism, work norms and social insurance," *Quarterly Journal of Economics* 121, 1473-1503.
- [51] Lindbeck, A, Nyberg, S. and J. Weibull (1999), "Social norms and economic incentives in the welfare state," *Quarterly Journal of Economics* 114, 1-35.
- [52] Lindbeck, A, Nyberg, S. and J. Weibull (2003), "Social norms and welfare state dynamics," *Journal of the European Economic Association* 1, 533-542.
- [53] Lundberg, S.J. and R. Startz (1983), "Private discrimination and social intervention in competitive labor markets," *American Economic Review* 73, 340-347.
- [54] Lundberg, S.J. and R. Startz (2000), "Inequality and race: Models and policy," in K. Arrow, S. Bowles and S. Durlauf (Eds.), *Meritocracy and Inequality*, New York: Princeton University Press, pp. 269-295.
- [55] Luttmer, E. (2001), "Group loyalty and the taste for redistribution," *Journal of Political Economy* 109, 500-528.
- [56] Massey, D.S. and N.A. Denton (1993), *American Apartheid: Segregation and the Making of the Underclass*, Cambridge: Harvard University Press.
- [57] Moro, A. and P. Norman (2003), "Affirmative action in a competitive economy," *Journal of Public Economics* 87, 567-594.
- [58] Mulligan, C.S. (1996), "Work ethic and family background: Some evidence," Unpublished manuscript, University of Chicago.
- [59] Osborne Groves, M.A. (2005), "Personality and the intergenerational transmission of economic status", in: S. Bowles, H. Gintis and M.A. Osborne Groves (Eds.), *Unequal Chances: Family Background and Economic Success*, Princeton: Princeton University Press, 208-231.
- [60] Patacchini, E. and Y. Zenou (2011a), "Social network effects and parental behavior in the intergenerational transmission of religion," CEPR Discussion Paper No. 8443.
- [61] Patacchini, E. and Y. Zenou (2011b), "Intergenerational education transmission: Neighborhood quality and/or parents' involvement?" *Journal of Regional Science* 51, 987-1013.
- [62] Patacchini, E. and Y. Zenou (2012), "Juvenile delinquency and conformism," *Journal of Law, Economic, and Organization* 28, 1-31.

- [63] Phelps, E. (1972), "The statistical theory of racism and sexism," *American Economic Review* 62, 659-661.
- [64] Rosenbaum, E. and L.E. Harris (2001), "Residential mobility and opportunities: Early impacts of the Moving to Opportunity demonstration program in Chicago," *Housing Policy Debate* 12, 321-346.
- [65] Sáez-Martí, M. (2011), "Swimming with the current," mimeo.
- [66] Sáez-Martí, M. and A. Sjögren (2008), "Peers and culture," *Scandinavian Journal of Economics* 110, 73-92.
- [67] Stutzer, A. and Lalive, R. (2004), "The role of social norms in job searching and subjective well-being," *Journal of the European Economic Association* 2, 696-719.
- [68] Thernstrom A. and S. Thernstrom (2002), *Beyond the Color Line: New Perspectives on Race and Ethnicity in America*, New York: Hoover Press and the Manhattan Institute.
- [69] Vigdor, J. (2004), "Community composition and collective action: Analyzing initial mail responses to 2000 Census," *Review of Economics and Statistics* 86, 303-312.
- [70] Wilson, W.J. (1987), *The truly Disadvantaged: The Inner City, the Underclass, and Public Policy*, Chicago: University of Chicago Press.
- [71] Wilson, W.J. (1996), *When Work Disappears. The World of the New Urban Poor*, New York: Alfred A. Knopf Publisher.
- [72] Verdier, T. and Y. Zenou (2004), "Racial beliefs, location and the causes of crime," *International Economic Review* 45, 731-760.