UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for New Economic Thinking

Cambridge-INET Institute

Cambridge-INET Working Paper Series No: 2014/17

Cambridge Working Paper in Economics: 1466

DISCRIMINATION WITHOUT TASTE –

HOW DISCRIMINATION CAN SPILLOVER AND PERSIST

Rajesh Ramachandran (Goethe University Frankfurt)
Christopher Rauh (University of Cambridge)

ABSTRACT

We introduce coordination failures driven by beliefs regarding the presence of taste discriminators as a channel of discrimination in activities requiring the input of more than one individual for production to occur. We show discrimination can persist forever under perfectly observable ability, when taste for discrimination has died out, and under absence of discriminatory social norms. Empirically, we analyze the market for self-employment in the US, a market requiring input from multiple sources. Consistent with the theoretical prediction, we find beliefs about discrimination to be a significant negative correlate of self-employment rates of blacks in the US.

Discrimination Without Taste - How Discrimination Can Spillover and Persist

By Rajesh Ramachandran and Christopher Rauh*

We introduce coordination failures driven by beliefs regarding the presence of taste discriminators as a channel of discrimination in activities requiring the input of more than one individual for production to occur. We show discrimination can persist forever under perfectly observable ability, when taste for discrimination has died out, and under absence of discriminatory social norms. Empirically, we analyze the market for self-employment in the US, a market requiring input from multiple sources. Consistent with the theoretical prediction, we find beliefs about discrimination to be a significant negative correlate of self-employment rates of blacks in the US.

JEL: C73, D83, J71

Keywords: Discrimination, Coordination Failure, Beliefs

Discrimination against certain social groups over long time periods has been a historical feature of many societies. For instance, in the US discrimination in the form of slavery officially ended in 1865 after more than

* Ramachandran: Department of Microeconomics and Management, Goethe University Frankfurt, Germany. (email: ramachandran@econ.uni-frankfurt.de). Rauh: Department of Economics, University of Cambridge, UK. (email: cr542@cam.ac.uk).
We would like to thank Miguel-Angel Ballester, Steven Durlauf, Tore Ellingsen, Joan Esteban, Itzhak Gilboa, Jim Heckman, Aviad Heifetz, Hannes Mueller, Jan Stuhler, Jack Willis, and participants of the seminar at Toulouse School of Economics, the SSSI 2012 at the Becker Friedman Institute of University of Chicago, the 2012 Asia Meeting of the Econometric Society, New Delhi, and the 2013 annual meeting of the Royal Economic Society, London, for their valuable feedback. All errors are ours.

two centuries, though racial segregation was maintained in the form of Jim Crow laws until 1965.[1] Starting with the civil rights movements in the early 1960's, one has seen significant advances in the rights and outcomes of the black population. However, today the black population still lag behind whites in a range of socio-economic characteristics. In India, caste, which is inherited by birth, was a marker for social discrimination for centuries. At independence in 1947, the practice of untouchability was made illegal and affirmative action was enshrined in the constitution for disadvantaged groups. However, the lower castes continue to trail significantly behind other social groups in terms of most socio-economic indicators. What contributes to the gap between groups that faced discrimination over long time periods and those that did not? In what outcomes and why might we observe persistent gaps? Could discrimination persist due to more subtle channels than the traditionally assumed channels of taste based discrimination, statistical discrimination, and discriminatory social norms?

In this paper, we posit a channel of discrimination, where even under perfect observability of individual ability, the absence of discriminatory social norms, and when taste for discrimination has already died out, to discriminate can be the optimal response. The theoretical mechanism put forth rests on the existence of beliefs about discrimination by others in society, and on distinguishing between activities characterized by the need for *interlinkages* versus *no* need for interlinkages. In our model, activities with interlinkages require coordinated actions. If an individual decides to establish interlinkages, she requires the input of two principals to form a productive unit. The

[1]Note that slavery had existed in colonial America since the 17th century but the United States as an independent nation state came into existence in 1774 and slavery ended 91 years later.

success and return for all, the individual and the two principals, is contingent on the participation of all three in the venture. The coordination failure results from the belief that somebody else might discriminate and refuse to participate in the venture, which imposes losses due to the complementarity of inputs in the production process.

The classic example would be the case of entrepreneurs who need to establish multiple interlinkages (productive relations) to be able to start and operate a venture (Basu, 2010). In the theoretical model, individuals choose between entering activities which require establishing productive relations and those that do not. Individuals intending to enter activities involving interlinkages are randomly matched with a pair of "principals", for instance a lender and a distributor, with whom they need to establish interlinkages to form a productive unit. The individual cannot produce without capital and cannot sell without a distributor. In case one of the principals agrees to participate and the other does not, the investment of the first principal is held up and imposes a fixed cost. We show how in the presence of beliefs about discrimination against a certain group, principals without a taste for discrimination in equilibrium also discriminate against that group.

To clarify the theoretical mechanism, picture the following situation: Both the lender and the distributor have no preference for discrimination against the individual and know she has the requisite ability to be an entrepreneur. However, the lender does not know whether the distributor is a discriminator because historically principals had been discriminating against individuals of her type. The lender has a belief about the presence of taste discriminators, which he has been updating through Bayes rule based on past observations. If the lender believes with a sufficiently high probability that the distribu-

tor has a preference for discrimination, he will reject the loan application. Now, if alternatively the distributor had signed the contract to distribute the goods but the individual has not obtained a loan to produce, the distributor loses out, because by setting up the productive relation he has foregone the chance for an alternative investment. Therefore, in the future the distributor will account for the possibility of the individual being matched with a taste discriminator. The individual also faces a cost because she invested time and effort to become self-employed but did not manage to do so. Consequently, people from her type might also refrain from attempting to become self-employed. This leads us to a persistent equilibrium in which able people are not becoming self-employed due to past discrimination and the resulting coordination failure based on beliefs, leaving everybody worse off.[2]

Under certain conditions, the model predicts lower participation rates and higher cost of establishing interlinkages for the discriminated group relative to the non-discriminated group in equilibrium, leading to an overall welfare loss for society.[3] The model also establishes conditions under which the steady-state equilibrium is characterized by the existence of discrimination due to beliefs about the existence of taste discriminators, although there are no taste discriminators left in society. The persistence of beliefs regarding discrimination in the steady state are rationalized by presenting evidence that these can be interpreted as intergenerational transmission of beliefs in the sense of *collective memories*, consistent with utility maximizing or cultural trait preserving strategies.

The theoretical framework identifies occupational choice, such as self-

[2]The example of the distributor believing the lender will discriminate is analog and would exhibit the same outcome.

[3]Later we show that this result is robust to allowing for communication between the two principals as long as there are no reputation effects.

employment, as markets characterized by interlinkages, making it a suitable candidate for empirical scrutiny. We examine the market for self-employment of blacks and whites in the United States. Using data from the General Social Survey (GSS) from the years 1972-2012 (Smith et al., 2012), we create proxies of beliefs about and tastes for discrimination against blacks for every region and year. We provide evidence that our constructed measures of taste for discrimination are a reasonable proxy for the actual trend of taste for discrimination, despite relying on self-reported preferences.

The time trends of taste for discrimination and beliefs about discrimination from the GSS and the self-employment rates for blacks and whites from the Current Population Survey (CPS) for the time period 1972-2012 are shown in Figure 1.[4] Taste for discrimination against blacks linearly declines over the period, whereas beliefs about discrimination against blacks as well as the gap between the self-employment rates for blacks and whites remain remarkably constant over the same time period. Figure 1 captures the mechanism and the role of *sticky or unchanging* beliefs highlighted by the theoretical model in a snapshot. The unchanging beliefs perfectly correspond to the invariant gap in self-employment rates over the period analyzed, as predicted by the theoretical framework. Using a logit model, we find our proxy for beliefs about presence of discrimination to be a significant and negative correlate of the probability of becoming self-employed for blacks in the US. The results are robust to the inclusion of a race dummy to account for other unobservable characteristics of racial groups, as well as year and region fixed effects. The presented statistical associations are persistent across a variety of specifications and present suggestive evidence

[4]See Section II.A for how the measures of taste and beliefs about discrimination are constructed.

in favor of the theoretical framework, though no causal claims can be made on the basis of the available data.

The literature of the economics of discrimination was pioneered by the

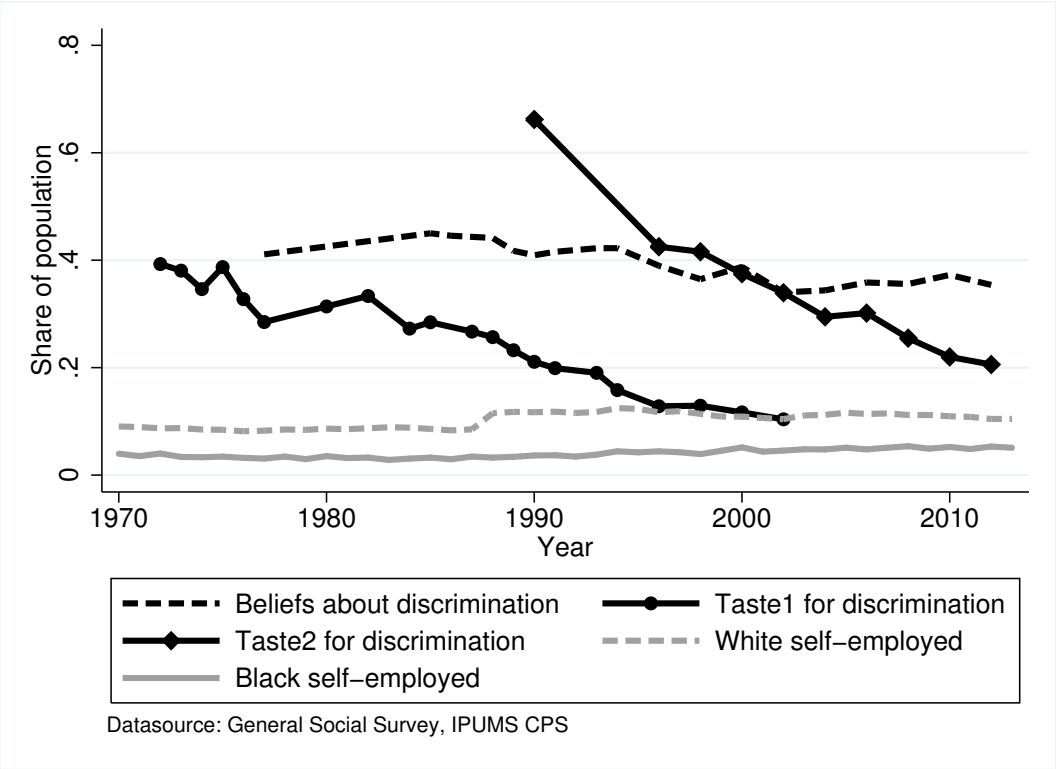


Figure 1. : Self-employment rates by race and beliefs and taste regarding discrimination in the US

seminal work of Becker (1957). In the setting envisaged, employers hold a taste for discrimination, such that working with members of a particular group imposes a cost on them, and hence these workers have to compensate the employer by either being more productive or accepting lower wages. The class of models of statistical discrimination (Phelps 1972; Arrow 1973; Aigner and Cain 1977; Lundberg and Startz 1983, 1998; Coate

and Loury 1993; Rosén 1997) and categorical thinking (Fryer and Jackson 2008) rely on the imperfect observability of worker productivity. In absence of complete information employers base their decision on easily observable characteristics, such as race or gender, to infer the expected productivity of the worker. Mailath, Samuelson and Shaked (2000) present a model of endogenous discrimination arising from the search decision of firms. The asymmetric discriminatory equilibrium is supported due to the belief that there are more skilled available workers of a particular type, which is borne out in equilibrium. The third class of models is that of Akerlof (1976, 1985) and Peski and Szentes (2013), where not following the established norm of discrimination against certain groups might result in imposition of social sanctions which cause economic losses, and hence make discrimination a rational response. Our model provides a new mechanism as to how discrimination can persist. In our setting the distribution of ability within the two groups is identical ex-ante and ex-post, there is perfect observability of ability, and there are no social norms to discriminate. Moreover, the nature of the coordination failure highlighted does not allow for a single principal who does not discriminate to reap the unrealized profits, a possibility traditionally assumed by Becker (1957), therefore providing a theoretical rational as to why discrimination can persist. To our knowledge, we are also the first to provide empirical evidence, albeit correlational, concerning the channel of discrimination presented theoretically.

The rest of the paper is organized as follows: Section I presents the static and dynamic theoretical model, Section II presents empirical support of our theory and extends applications to other markets and situations, and Section III concludes.

I. The model

The society consists of individuals i of two types $s_i \in \{A, B\}$. The types A and B form social groups based on visible characteristics which do not influence performance (e.g., race, gender). Individuals of type A and B belong to the finite, large sets $\mathcal{A}$ and $\mathcal{B}$, respectively.[5] Individuals have an ability a_i, where a is distributed uniformly over $[0, 1]$. Ability $a_i \in [0, 1]$ reflects productive capacity and is perfectly observable to all. For sake of simplicity we are dropping the index i in what follows.

Those referred to throughout the paper as "individuals" can opt to engage in one of the two possible kinds of productive activities in the economy ($L \in \{0, 1\}$) - activities that involve establishing *interlinkages* ($L = 1$) with other agents, who are referred to as "principals" $P \in \{p_1, p_2\}$, and activities that do *not* establish interlinkages ($L = 0$). In case the individual i of type s decides to engage in an activity that does not involve establishing interlinkages with other principals in the economy, she earns a net income on her activity equal to the level of her productivity, which is given by:

$$W = G_{NL}(a) = a, \tag{1}$$

where G_{NL} is the production function of activities not involving interlinkages. On the other hand, individuals have the option of engaging in an activity which involves establishing interlinkages with other principals, and earn a gross income equal to:

$$W = G_L(a, C) = \lambda c_{p_1} c_{p_2} a, \tag{2}$$

[5]The assumption of large sets is to ensure that any single individual does not have any market power and collusion cannot take place.

where G_L is the production function for activities involving *interlinkages*. The above production function G_L captures the notion of *interlinkages*. Activities requiring interlinkages imply that the gross income from this activity not only depends upon the individual's own ability a, but also on C. The component $C = (c_{p_1}, c_{p_2})$ captures the interlinkages or productive relations individual i is able to establish with the principals in the economy. Interlinkages refer to the fact that the production in such activities is a joint process and requires input from multiple sources.

We denote by $c_P \in \{0, 1\}$ the decisions by the principal $P \in \{p_1, p_2\}$ of whether to establish productive relations. Thus, we only allow for pure strategies, such that the two principals decide whether to establish the relation ($c_P = 1$) or not ($c_P = 0$). Moreover, it is assumed that in case the principals decide to establish a productive relationship with the individual, they need to make an investment, which is normalized to unity. The incentive for engaging in activities that involve establishing interlinkages arises as complementary investment by the principals results in a boost to productivity, captured by the factor $\lambda(> 1)$ in the production function G_L.

A simple example to fix ideas for activities that are characterized by the need to establish interlinkages, is the market for self-employment.[6] For example, an entrepreneur might require capital in the form of a loan from a bank (lender) to produce goods and also may need to have an agreement with a distributor willing to distribute her goods. The example with the requirement of productive relations with two principals, a distributor and a lender, is only for illustrative purposes and tractability, and could be extended to n-players or include any other contact necessary to setup a

[6]In Section II.D we outline a range of alternative applications of the model beyond self-employment.

successful enterprise (e.g., supplier, landlord to rent office).

The functional form in (2) exhibits an extreme form of complementarity in the actions of the principals P, implying $G_L(a, C) = G_L(a, (c_{p_1}, 0)) = G_L(a, (0, c_{p_2})) = 0$. The intuition is that establishing a relationship with both principals $P \in \{p_1, p_2\}$ is required for the individual to produce, as neither component (loan, distribution route) can be substituted through ability.

A. The static game

Individuals i decide whether they want to enter an activity that involves no interlinkages, or whether to enter into activities involving interlinkages by trying to establish productive relations with the principals. As already noted, individuals who decide to enter activities involving no linkages earn a net income equal to their ability a.

Individuals wanting to enter activities involving interlinkages are randomly matched with a pair of principals, p_1 and p_2, in the market to try to establish productive relations. Principals have an outside opportunity of a risk free investment yielding interest r per unit invested. To establish a productive relationship, and in return for the investment in their activity by the principals, the individual's offer an amount σ_P to each of the principals as repayment for the investment. In case this offer is rejected by any one of the principals, the attempt to enter the activity with interlinkages fails and the individual i faces a fixed cost δ from the effort exerted. She then enters the activity not involving interlinkages and earns a net income of $(a - \delta)$.

The offer, denoted by σ_P, made by the individual to the principal arises

as a solution of a Nash bargaining process over the surplus generated from the activity involving interlinkages compared to the alternative involving no interlinkages. The bargaining takes place between each principal and the individual separately, without communication between p_1 and p_2.[7] In light of the production function given by (2), the outcome is dependent on what decision the other principal (henceforth denoted by $-P$) makes. Hence, P has a belief concerning the likelihood of the other principal accepting the offer as well.[8] If P accepts an offer which the other principal rejects, then he is not able to obtain r from the risk free investment in the given period due to his capital being bound and not yielding any interest.

The production function given by (2) implies that production can take place only if the grand coalition, i.e. of the two principals and the individual, forms. This implies returns to bilateral agreements are zero and hence in the absence of a grand coalition, the individual prefers to enter the activity involving no interlinkages and the principals prefer investing in the risk-free asset. Thus, the disagreement points are exogenously given by their outside option and, contrary to Bennett (1997) and Burguet and Caminal (2011), do not depend on what the individual can achieve in the alternative negotiation. The interaction between the individual and the principal is characterized by a monopoly versus monopsony, commonly referred to as bilateral monopoly. Both have one shot at earning a surplus compared to their outside option. Assuming equal bargaining power and linear utility functions in payoffs, the bargaining solution, resulting in offer

[7]Communication could be allowed between the two principals and would not change our results if we assume costly communication or no reputation effects. Without reputation effects moral hazard problems would arise as there would be no gain from admitting when one was not willing to establish the productive relation.

[8]We disregard higher order beliefs (as in, the lender believes that the distributor believes that etc.), even though they would additionally speed up the contagion-effect.

σ_P, is characterized by the disagreement point given by the outside option, i.e. $d = (d_i, d_P) = (a - \delta, 1 + r)$, and the maximization of:

$$(3)\quad (\sigma_P - 1 - r)(\lambda a - \sigma_P - \sigma_{-P} - a + \delta) \Rightarrow \sigma_P = \frac{(\lambda - 1)a - \sigma_{-P} + r + \delta + 1}{2}.$$

Now assuming i makes the same offer to the two principals, such that $\sigma_P = \sigma_{-P}$, the Nash bargaining solution is defined as:

$$(4)\qquad \sigma^N(a) = \frac{(\lambda - 1)a + r + \delta + 1}{3}.$$

This Nash bargaining solution is a function of a, which we denote as $\sigma^N(a)$. In order for P to accept this offer, we require the amount he expects to be repaid to be at least what he can earn through the risk free investment, i.e. $(1 + r)$.

Let the state variables $x = (a, s)$ be given by ability a and type s of an individual i. The utility of the individual is $V_i(x)$, which will either be given by the utility of not establishing interlinkages $V_i^{NL}(x)$, or by the expected utility of attempting to establish interlinkages $E_C[V_i^L(C, x)]$. Formally, the static game for the individual can be written as

$$\begin{aligned}
&V_i(x) = \max_{L \in \{0,1\}, \sigma \in \mathbf{R}^+} \{V_i^{NL}(x), E_C[V_i^L(C, x)]\} \\
&\textit{subject to} \\
&(5)\qquad \begin{cases} E_C[V_i^L(C, x)] = C(\lambda a - 2\sigma) + (1 - C)(a - \delta) & \text{if } L = 1 \\ V_i^{NL}(x) = a & \text{if } L = 0, \end{cases} \\
&C = c_{p_1} c_{p_2},\ \sigma \geq \sigma^N(a).
\end{aligned}$$

As noted before, the offer σ made by the individual to the principals arises from a process of Nash bargaining, and is the minimum offer that will be accepted by a principal given our assumption of equal bargaining power. However, in the optimization problem of the individual, given by (5), we allow the individual to offer an amount greater than the Nash bargaining solution in case it maximizes her utility. In other words, if an offer of $\sigma^N(a)$ is acceptable to the principal, so is any offer $\sigma \geq \sigma^N(a)$, as it gives the principal a share of the surplus more than commensurate to his bargaining power.

For the principal, the static game concerning his decision $c_P \in \{0, 1\}$ is represented by

$$
\begin{aligned}
& V_P(x, \sigma) = \max_{c_P \in \{0,1\}} \{V_P^{c0}, E_{c_{-P}}[V_P^{c1}(c_{-P}, x, \sigma)]\} \\
& \textit{subject to} \\
& \begin{cases} E_{c_{-P}}[V_P^{c1}(c_{-P}, x, \sigma)] = c_{-P}\sigma + (1 - c_{-P}) & \text{if } c_P = 1 \\ V_P^{c0} = 1 + r & \text{if } c_P = 0. \end{cases}
\end{aligned} \tag{6}
$$

Given $\sigma^N(a)$ from (4) it is now possible to calculate the lowest ability individual, denoted by a', who could possibly offer the principal a share dominating the risk free investment, and is given by:

$$
\sigma^N(a) = \frac{(\lambda - 1)a + r + \delta + 1}{3} \geq 1 + r \Rightarrow a' = \frac{2(1 + r) - \delta}{\lambda - 1}. \tag{7}
$$

Now looking at the participation constraint of the individual, we require $\lambda a - 2\sigma^N(a) \geq a$

$$
\Rightarrow a^* = \frac{2(1 + r + \delta)}{\lambda - 1}. \tag{8}
$$

Since $a^* > a'$ only individuals with $a \geq a^*$ will intend to enter activities involving interlinkages and, as long as there is no taste for or belief about discrimination, will be accepted.[9]

B. Discrimination in the static framework

The static game outlines the decision making process and identifies the lowest ability type a^* from both groups $s \in (A, B)$ who enter into activities involving interlinkages without discrimination.

Now assume that individuals and principals believe that there exist principals with a taste for discrimination.[10] Taste for discrimination can be understood as a cost/disutility which principals with taste for discrimination face when they decide to establish a productive relation with a B-type individual in society, and is captured by the parameter $b(> 0)$.

The belief regarding the presence of taste discriminators implies that the probability of discrimination occurring has to be taken into account while deciding on the optimal course of action. Due to the random matching, the assigned probability of meeting a principal with a taste for discrimination is equal to what is believed to be the share of total taste discriminators among principals. As the share is not observable, decisions are conditioned on beliefs shared by the principals and individuals about the fraction of discriminators amongst principals, which is assumed to be equal to φ. The extensive game form is exhibited in Figure 2.

From the static model without discrimination, we know that the minimum offer that will be accepted by a principal without a taste for dis-

[9]Similar to Lucas Jr (1978), in our model only the most able want to enter the activity involving interlinkages such as entrepreneurship.

[10]How this belief can arise and sustain itself even after all principals with taste for discrimination have died out is discussed in the next subsections.

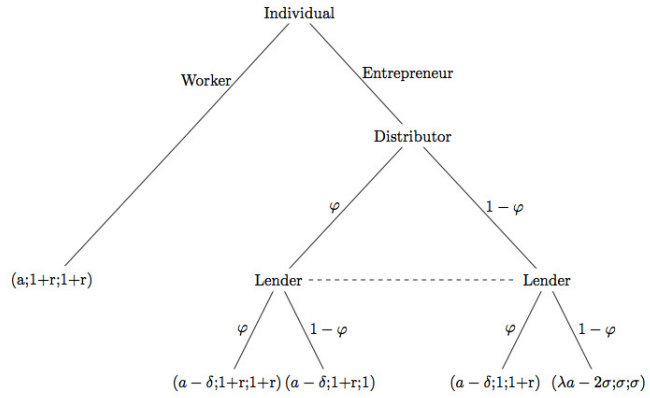


Figure 2. : Extensive game form of decision of B-type individual

crimination is $\sigma^N(a) \geq (1+r)$, i.e. a return greater than the risk free investment. Similarly, observe that any offer $\sigma^N(a) \geq (1+r+b)$ will be always accepted, even by principals with a taste for discrimination, as it compensates for the taste for discrimination. This implies that only offers in the range $(1+r) \leq \sigma^N(a) < (1+r+b)$ could be subject to any discrimination. Individuals and principals are expected payoff maximizers. In light of the above, discrimination is defined in the following manner:

DEFINITION 1: *An offer of $\sigma^N(a) > 1+r$, which is rejected by any principal, is defined as a case of discrimination.*

Individuals of the B-type now take the probability of meeting a discriminator in the market into account while deciding on their optimal course of action.

DEFINITION 2: *Let us denote by a_b^* the lowest ability type such that $\sigma^N(a_b^*) =$*

$1 + r + b$ *and by* a_b *the ability of an individual who is indifferent between offering* $(1 + r + b)$ *to each principal and entering an activity not involving interlinkages.*

All individuals with ability $a \geq a_b^*$ offer $\sigma^N(a) \geq 1 + r + b$ and will never be discriminated against. Individuals in the ability range $a^* \leq a < a_b^*$ potentially face discrimination. These individuals compare the expected payoff from each of the three actions available to them, given by:

$$E[V_i(x)] = \begin{cases} (1-\varphi)^2(\lambda a - 2\sigma^N(a)) + (1-(1-\varphi)^2)(a-\delta) & \text{if } (L=1, \sigma = \sigma^N(a)) \\ \lambda a - 2(1+r+b) & \text{if } (L=1, \sigma = 1+r+b) \\ a & \text{if } L = 0. \end{cases}$$

The expected payoff from applying for the activity involving interlinkages ($L = 1$) and offering the Nash bargaining solution depends on the belief φ about the share of taste discriminators. The payoffs from offering each principal $(1 + r + b)$ and escaping potential discrimination, as well as from entering the activity involving no interlinkages, are certain. Individuals choose whichever profile (L, σ) maximizes their expected payoff.[11]

Principals with no taste for discrimination, when facing an individual of the B-type in the ability range $a^* \leq a < a_b^*$ that offers $\sigma^N(a)$, compare their expected payoff from accepting the offer to the certain payoff from the risk free investment:

$$E[V_P(x)] = \begin{cases} (1-\varphi)\sigma^N(a) + \varphi & \textit{if } c_P = 1 \\ 1 + r & \textit{if } c_P = 0, \end{cases}$$

[11]Recall that an individual is always free to offer a $\sigma \geq \sigma^N(a)$ in case that maximizes his/her expected utility.

and choose the option that gives them the greater expected payoff. Here we can observe how discrimination has spilled over to those principals without a taste for discrimination, as they now take into account the probability of the individual being matched with a principal who is a taste discriminator.[12] In case the principal P accepts and the counterpart $-P$ rejects the individual's offer, the principal P will lose his risk free return r. Thus, the principal P, due to the belief that the individual might face discrimination from principal $-P$, also ends up discriminating against the individual of type-B. Principals with a taste for discrimination equal to b reject the Nash bargaining solution when facing an individual in the ability range $a^* \leq a < a_b^*$ and accept offers from $a \geq a_b^*$.

Observe that taste for discrimination against B-type individuals does not affect the A-types in the market. The A-types play the identical game as in the static framework without discrimination. All individuals of the A-type with $a \geq a^*$ still offer $\sigma^N(a)$ and are accepted.

For the B-types potentially exposed to discrimination, observe that

$$\lambda a - 2\sigma^N(a) > (1-\varphi)^2(\lambda a - 2\sigma^N(a)) + (1-(1-\varphi)^2)(a-\delta), \tag{9}$$

such that the expected payoff without discrimination, ceteris paribus, will always be greater for the individual. Recall from (8), the individual with the lowest ability, who applies to enter activities with interlinkages in the absence of beliefs about discrimination, is given by a^*, which implies:

$$\lambda a^* - 2\sigma^N(a^*) = a^* > (1-\varphi)^2(\lambda a^* - 2\sigma^N(a^*)) + (1-(1-\varphi)^2)(a^*-\delta). \tag{10}$$

[12]In the Online Appendix, when discussing policy interventions, we show that discrimination under certain interventions can even spillover to the A-type as well

Expression (10), combined with the fact that the payoff from offering the Nash bargaining solution is strictly increasing in a, implies some value of $a > a^*$, denoted by a_{db}, will solve the equation such that $(1-\varphi)^2(\lambda a_{db} - 2\sigma^N(a_{db})) + (1-(1-\varphi)^2)(a_{db} - \delta) = a_{db}$.

DEFINITION 3: *Ability a_{db} is the minimum ability of a B-type wanting to apply for interlinkages offering the Nash bargaining solution in the presence of beliefs about discrimination.*

In the equilibrium of the static game with beliefs about discrimination, all A-types with $a \geq a^*$ apply to enter activities involving interlinkages, offer the Nash bargaining solution, and are successful. For the B-types, if $a_{db} < a_b$, only individuals with $a \geq a_{db} > a^*$ apply to enter activities involving interlinkages offering the Nash bargaining solution, and are successful.[13] On the other hand, if $a_{db} > a_b$, individuals with $a \geq a_b > a^*$ apply to enter activities involving interlinkages and are successful. B-type individuals with $a > a_{db}$ offer the Nash bargaining solution, like their A-type counterparts. However, B-type individuals in the ability range $a_b \leq a < a_{db}$ offer $(1+r+b) > \sigma^N(a)$, an offer strictly greater than the ones offered by the A-types in the same ability range.

Beliefs about discrimination, hence, have two potential effects: (i) if $a_{db} < a_b$, individuals of type B in the range $a^* \leq a < a_{db}$, enter activities involving no interlinkages, whereas A-types of the same ability enter activities involving interlinkages and enjoy the associated surplus, and (ii) if $a_{db} > a_b$, individuals of B-type in the range $a^* \leq a < a_b$ enter activities involving no interlinkages, whereas A-types of the same ability enter activities

[13]Recall a_b is defined as the ability type from group B who is indifferent between offering $(1+r+b)$ to each principal and entering an activity not involving interlinkages.

involving interlinkages and enjoy the associated surplus. On the other hand, B-types in the range $a_b \leq a < a_{db}$ enter activities involving interlinkages, but pay a higher price to the principals than the A-types in the equivalent ability range.

C. *The dynamic game and the belief updating process*

In the previous subsection, we assumed that there was an exogenously given belief regarding the probability of meeting a taste discriminator in society. In light of this assumption, we characterized the optimal decisions in a static framework. The logical questions that arise are whether these beliefs are consistent with a model of rational decision makers in a dynamic setting in which beliefs are updated, and how these beliefs regarding taste for discrimination can persist when there are no taste discriminators left in society.

In order to address these questions, we now extend the framework and allow for the game to be repeated every period. Assume that the taste for discrimination arises due to a shock to the taste of a subset of principals in society at time t_0. It is assumed that π_0 proportion of principals develop a taste for discrimination equal to $b(>0)$ against establishing a productive relation with B-type individuals.[14] The origins of the shock which result in creating a taste for discrimination among a subset of principals is not the focus of the paper. An example could be the incidents of September 11^{th} 2001, which resulted in discriminatory actions against Muslims in the US

[14]Allowing principals to have different tastes for discrimination leaves our results essentially unchanged. The case for two different levels of taste for discrimination, b_h (high) and b_l (low) is provided in the Online Appendix.

(e.g., Kaushal, Kaestner and Reimers 2007).[15]

In the dynamic game we assume time to be discrete and indexed by t. A principal P exits the market with exogenous probability ω every period. The probability ω is not known to anybody in society. A principal without a taste for discrimination replaces the exiting principal, such that at some point no principals with a taste for discrimination will be left. Therefore, if we define the share of principals with a taste for discrimination in period $t = 0$ to be π_0, the share of principals π_t with taste for discrimination in period t is $\pi_0(1-\omega)^t$.

Since neither π_t nor ω are common knowledge, decisions are conditioned on beliefs about the share of discriminators amongst principals, which are updated through observations of discrimination in the market.[16] We assume that the event which creates a taste for discrimination initially results in creating a common prior among individuals and principals.[17] The common prior is assumed to have a distribution denoted by η_0, capturing the probability of meeting a principal with a taste for discrimination. The common prior η_0 is modeled as having a beta distribution. More specifically, it is assumed that individuals and principals believe that the share of principals with taste b has a beta distribution with parameters α_0 and β_0. Moreover, we denote the density of the distribution η_0 by θ. The beta distribution captures the belief regarding the probability of meeting a principal with

[15]The assumption that shocks do not work the other way, i.e. people immediately forget the past existence of discrimination due to sudden events today, is justified by the literature on trust and beliefs which shows how persistent past beliefs are in shaping todays action. See section I.E for discussion and references.

[16]The definition of a case of discrimination is provided in Definition 1.

[17]The importance of the initial prior is negligible as we envisage a setting with long term discrimination, thereby reducing the weight the initial prior has in long run beliefs. In an earlier version of the paper beliefs were updated following frequentist approach as in case-based decision theory by Gilboa and Schmeidler (1995).

a taste for discrimination through its expected value, or the mean of the distribution.

ASSUMPTION 1: *The probability parameter capturing the share of lenders with a taste for discrimination equal to b in period t_0 is given by $\theta(\eta_0) \sim beta(\alpha_0, \beta_0)$.*

The above distribution implies that the density function associated with facing a discriminator with taste b is given by:

$$(11)\quad \theta(\eta_0) = \frac{(\eta_0)^{\alpha_0 - 1}(1-\eta_0)^{\beta_0 - 1}}{beta(\alpha_0, \beta_0)} = \frac{(\alpha_0 + \beta_0 - 1)!}{(\alpha_0 - 1)!(\beta_0 - 1)!}(\eta_0)^{\alpha_0 - 1}(1-\eta_0)^{\beta_0 - 1}$$

The beta distribution gives us a density on $[0, 1]$, which captures the beliefs held by individuals and principals regarding η_0. As individuals and principals need to decide on optimal actions based on their beliefs, and all individuals and principals are assumed to be risk neutral, individuals and principals use the expected value of the distribution which is given by $E(\eta_0) = \frac{\alpha_0}{\alpha_0 + \beta_0}$.[18]

The belief updating process of principals and individuals is characterized by a standard Bayesian approach. Assume that in period 1, n_1 individuals of the B-type applied and k_1 cases of discrimination were observed in the market. Out of the total of n_1 cases, assume that $n_{1b} (\leq n_1)$ cases involve offers $\sigma \geq \sigma^N(a)$, such that $(1 + r) \leq \sigma < (1 + r + b)$. This implies the total number of people who could potentially be discriminated against is n_{1b}.

ASSUMPTION 2: *It is assumed that all market transactions in terms of the offers made and rejected are common knowledge.*[19]

[18]This is the point at which the density of the distribution takes its highest value.

[19]We relax this assumption in the Online Appendix to allow individuals and principals

Given that out of the potential n_{1b} cases, k_1 cases exhibit discrimination, we can define the posterior density function for the individuals and principals in society. The posterior function for $\theta(\eta_0)$ is given by $\theta(\eta_1|k_1) \sim beta(\alpha_0 + k_1, \beta_0 + n_{1b} - k_1)$.

The above outlines the Bayesian belief updating procedure used by individuals and principals regarding the probability of meeting a principal with taste for discrimination. We can denote the posterior distribution for any period $T > t_0$, given the total number of B-type individuals who make offers $\sigma^N(a)$, such that $(1 + r) \leq \sigma^N(a) < (1 + r + b)$, in order to become entrepreneurs, and the cases of discrimination observed in the market. The posterior probability density is given by $\theta(\eta_T|\sum_{t=1}^{T} k_t) \sim beta(\alpha_0 + \sum_{t=1}^{T} k_t, \beta_0 + \sum_{t=1}^{T} n_{tb} - \sum_{t=1}^{T} k_t)$. The associated expected value or the point probability estimate used by individuals and principals to make their optimal decision is given by:

$$E(\eta_T) = \frac{\alpha_0 + \sum_{t=1}^{T} k_t}{\alpha_0 + \beta_0 + \sum_{t=1}^{T} n_{tb}}. \tag{12}$$

DEFINITION 4: *Let $\varphi_t = E(\eta_{t-1})$, such that φ_t is the probability that individuals and principals assign to the existence of a principal with taste for discrimination b in period t.*

The decision-making rules of individuals and principals imply that the probability of entering an activity involving interlinkages for a B-type individual in any period T will depend upon her ability a, the actual share of taste discriminators $\pi_0(1 - \omega)^T$, and the beliefs φ_T regarding the share of taste

to observe only a subset of all the market transactions and show that the results remain essentially unchanged.

discriminators in society. We can thus express the probability of entering an activity involving interlinkages for a B-type individual as a function of the above three factors, i.e. $f(a, \pi_0(1-\omega)^T, \varphi_T)$. It is easy to see that the probability of entering an activity involving interlinkages is increasing in ability and declining in the actual share and the belief regarding the proportion of taste discriminators in society, i.e. $f_1 > 0$, $f_2 < 0$, and $f_3 < 0$, where the subscripts refer to the first, second, and third argument of the function. In Section II we explicitly test for the predictions of our model using the above function f.

D. Characterization of the dynamic steady state equilibrium under no remaining taste for discrimination

The channel of discrimination that we put forth works on the premise that even once all principals with taste for discrimination have died out, to discriminate against members of group B may remain as the optimal action. In what follows we address whether discrimination can exist, and if it can, under what conditions does it exist, for how long does it persists, and in what form does it manifest itself. Let us denote by T^* the first period in which no principals with taste for discrimination remain in the economy. The probability density function, given the beliefs and the Bayesian updating rule used, for meeting a principal with a taste for discrimination is given by:

$$beta(\alpha_0 + \sum_{t=1}^{T^*-1} k_t, \beta_0 + \sum_{t=1}^{T^*-1} n_{tb} - \sum_{t=1}^{T^*-1} k_t). \tag{13}$$

The probability point estimate for meeting a discriminator is given by

$$\varphi_{T^*} = \frac{\alpha_0 + \sum_{t=1}^{T^*-1} k_t}{\alpha_0 + \beta_0 + \sum_{t=1}^{T^*-1} n_{tb}}. \tag{14}$$

It is clear that all $B-$type individuals with $a \geq a_b^*$ will offer their Nash bargaining solution and be accepted.[20] The form of discrimination and the length for which it will persist after all principals with taste for discrimination have died out will depend on φ_{T^*}, i.e. the belief on the probability of the individual meeting a taste discriminator.

The point probability estimates in period T^* are a function of the initial beliefs (α_0, β_0), the actual share of taste discriminators π_0, and the rate ω at which principals with a taste for discrimination exit the market in every period. If we assume that the initial beliefs are a function of the actual share of taste discriminators, i.e. $\alpha_0(\pi_0)$ and $\beta_0(\pi_0)$, then we can write $\varphi_{T^*} = z(\pi_0, \omega)$. In the proposition that follows, we highlight the various forms in which discrimination manifests itself and persists in the multiple steady state equilibria depending on φ_{T^*} after no principals with a taste for discrimination are left.

PROPOSITION 1: *1) Let φ_{T^*} be such that no individual of the type B in the range $a^* \leq a < a_b^*$ prefers the Nash bargaining solution to wage employment at time T^*. In such a scenario discrimination based on beliefs persists forever. Discrimination manifest itself in two forms: (i) B-types being underrepresented, relative to A-types, in activities involving interlinkages at the lower tail of the ability distribution and (ii) B-types, in the middle ability ranges, pay a strictly higher fee to*

[20] Recall $\sigma^N(a_b^*) = 1 + r + b$ and $\sigma^N(a)$ is increasing in a.

establish interlinkages than the A-types with similar ability.

2) *Let φ_{T^*} be such that some individual of B-type with ability in the range $a^* \leq a \leq a_b^*$ strictly prefers the Nash bargaining solution to obtaining wage employment. This implies that in the long run discrimination will not persist. However, B-types are penalized in the form of lower participation rates relative to A-types for a finite duration before beliefs about discrimination disappear from society.*

Proof in the Appendix.

The equilibrium outcome in which discrimination persists forever crucially depends on who is the lowest B-type deciding to apply for the activity involving interlinkages when the last principal with taste for discrimination dies out. This occurs when all individuals of the B-type, whose Nash bargaining solution is not sufficiently high to compensate the taste for discrimination, i.e. $a < a_b^*$, decide to enter activities involving no interlinkages rather than seek to establish productive relations. Even if an offer which could be subject to potential discrimination were to be made, it would be rejected due to beliefs about discrimination being prohibitively high. Therefore, this equilibrium is even stable under the trembling hand. This in turn implies that beliefs remain frozen at the current level and hence all individuals with ability levels $a^* \leq a < a_b$ will always prefer entering activities involving no interlinkages.

However, if when the last principal with taste for discrimination dies out, the lowest B-type who decides to enter an activity involving interlinkages is one whose Nash bargaining solution is not sufficient to compensate the taste for discrimination, i.e. $a < a_b^*$, implies discrimination will not persist in the long run. Now that all principals with taste for discrimination have died out

and beliefs are not prohibitively high, all offers made by individuals seeking to establish productive relations will be accepted. As this includes individuals whose offers could have been subject to potential discrimination, but are not (as no taste discriminators are left), the next period beliefs about discrimination will be lower after beliefs have been updated. As every period all offers are accepted, in the long run the belief about discrimination will converge to zero.

E. Persistence of beliefs as collective memories

The model presented above assumes that once the equilibrium set of beliefs have been established they can persist over time under certain conditions. The crucial question then arises as to how and why beliefs regarding the presence of discrimination might tend to persist? We interpret transmission of beliefs in our model as happening through intergenerational transmission of collective memory regarding discrimination.

The contemporary usage of the term collective memory can be traced back to Emile Durkheim (1859-1917), and his student Maurice Halbwachs (1877-1945), who published the seminal study titled *The social framework of memory* in 1925. The concept of memory has been constructed in the literature as to how the mind works in a society and how their operations are structured by social arrangements. Halbwachs argues: "It is in society that individuals normally acquire their memories. It is also in society that they may recall, recognize and localize their memories" (Halbwachs 1992, 38). Formulation of memories regarding the past are hence affected by transmission of cultural beliefs and norms in society.

Beliefs regarding discrimination can be seen to fulfill the two important cri-

teria to be categorized as collective memories. First, events which influence collective memory are widely documented and recorded in these societies (Griffin and Bollen 2009). For the case of discrimination against blacks in the US or Dalits in India, these events have been widely recorded and recollected. Second, a consensual view of the recollected past. The presence of affirmative action policies in the US and India serve as clear signals of consensus among policymakers and the public at large concerning the need to address previous wrongs.

Beliefs regarding discrimination being transmitted as collective memory through generations can also be rationalized by economic models of cultural transmission such as in Bisin and Verdier (2001) and Dessí (2008). They show that transmission of existing beliefs by parents to their offspring would be consistent with maximizing the utility of children or preserving their cultural traits. Finally, the importance of history, culture, and past events such as discrimination in shaping today's beliefs, behavior, and outcomes, has also been demonstrated in the empirical literature (Nunn and Wantchekon 2011; Voigtländer and Voth 2012; Alesina, Giuliano and Nunn 2013) and brought forth theoretically in Argenziano and Gilboa (2012).

The above discussion highlights the fact that beliefs regarding discrimination could be understood as collective memories that are passed on from one generation to another, which can be remarkably similar for long stretches of time.

F. Welfare effects

Now, let us consider the loss of welfare to society when discrimination persists due to beliefs, despite no taste for discrimination remaining in the

economy. The efficiency loss not only affects the B-type individuals, who become involved in activities involving no interlinkages instead of establishing productive relations (which would be the social optimum), but also the principals, who lose out on opportunities of receiving offers yielding more than the risk free investment r. Assuming that the necessary and sufficient condition of proposition 1.1 hold and discrimination persists, we can quantify the deadweight loss of each period to the B-type as $\sum_{a\geq a^*}^{a_b}[(\lambda a - 2\sigma^N(a)) - a]$, while the deadweight loss to principals is $\sum_{a\geq a^*}^{a_b} 2[\sigma^N(a) - (1+r)]$. Additionally, a wealth transfer takes place, as B-types offering $(1 + r + b)$ to escape discrimination are paying a higher price than the equivalent A-type, from which principals are profiting as they are receiving more than the Nash bargaining solution.[21] The transfer from the B-types to the principals is $\sum_{a\geq a_b}^{a_b^*} 2[(1 + r + b) - \sigma^N(a)]$.

PROPOSITION 2: *In an equilibrium as in proposition 1.1, on average both principals earn lower profits by discriminating.*

Proof in the Appendix.

In the top panel of Figure A1 in the Appendix the difference between net earnings of the A and the B-type in function of their ability is illustrated. The dark shaded area is the deadweight loss caused by individuals of the B-type not entering activities involving interlinkages due to beliefs about discrimination, whereas the light shaded area illustrates the transfer caused by the higher price individuals of the B-type are paying in order to escape discrimination when choosing activities involving interlinkages.

[21]This theoretical prediction is consistent with the finding of the US Department of Justice that Countrywide charged more than 200,000 black and Hispanic borrowers higher fees and interest rates than comparable whites with similar credit histories between 2004 and 2008 leading to the Bank of America paying a settlement of 335M$.

In the GSS dataset the pattern of average income in constant dollars of the self-employed by highest educational degree attained, exhibits a striking similarity with our theoretical prediction. As can be seen in bottom right panel of Figure A1 in the Appendix, on average blacks earn less than whites in self-employment for all but those that obtained a graduate degree, which is the highest degree coded in the dataset.

II. Data and empirics

As foreshadowed in our discussion in the theoretical section, we empirically investigate the market for self-employment in the US, an occupation characterized by the need to establish interlinkages across markets. We analyze whether the patterns of self-employment of blacks, as compared to the white population, are correlated to differences in taste for and belief about discrimination across regions and time in the US.

The empirical literature dealing with discrimination and self-employment in the US documents the differences in participation and returns between ethnic groups (Moore 1983; Borjas 1986; Bailey and Waldinger 1991; Fairlie 1999; Fairlie and Meyer (1996, 2000); Blanchflower, Levine and Zimmerman 2003; Fairlie and Robb 2008; Blanchflower 2009). In line with our theoretical predictions the above mentioned studies find that black males have lower self-employment rates, blacks are more likely to have loan applications rejected, and pay higher interest rates on loans than comparable white males do. For the case of wage employment, premarket skills measured by the Armed Forces Qualification Test (AFQT) score have been shown to account for most of the black-white wage gap (Neal and Johnson 1996). However, for the case of self-employment, Fairlie (2002) shows that controlling for AFQT

test scores does not significantly reduce the black-white gap concerning self-employment rates, suggesting that discrimination might have a role to play in explaining the observed differences in self-employment rates.

A. Data

We use the General Social Survey (GSS) from 1972-2012 along 29 questionnaires to provide empirical support for the predictions of our theoretical framework. The data allows us to construct proxies for the belief about and taste for discrimination parameters in our model. We construct two proxies of taste for discrimination by computing the share of whites by year and region that express taste for discrimination. We define taste discriminators to be:

1) Whites answering "yes" to "Do you think there should be laws against marriages of Blacks and Whites?"

2) Whites who are "very" or "somewhat opposed" when asked "What about having a close relative marry a Black person?"

In order to construct a proxy for beliefs regarding discrimination, we take the share of the sample, for each year and region, answering the following question with "yes":

- "On the average Blacks/African-Americans have worse jobs, income, and housing than White people. Do you think these differences are mainly due to discrimination?"

Unfortunately, neither of these questions is asked throughout all survey years, which, depending on the specification, restricts our sample size to between 14,719 and 23,895 observations. In Figure 1 the two measures of taste

for discrimination, beliefs about discrimination, and the self-employment rates of blacks and whites are plotted from 1972-2012 for those years where the corresponding questions were included in the surveys. The discrimination measures are from the GSS dataset, whereas for expositional purposes self-employment rates by race are obtained from the Current Population Survey (CPS) March supplement provided by the Integrated Public Use Microdata Series (King et al. 2010).

Tastes for discrimination seem to decline linearly. Beliefs about discrimination, on the other hand, have remained remarkably stable, just as the gap in self-employment rates between blacks and whites. Not decomposing by region, beliefs about discrimination among whites peak in 1985 at 45% and reach its lowest point in 2004 at 34%. Our first measure for taste for discrimination among whites declines from 39% in 1972 to 10% in 2002. The second measure declines from 66% in 1990 to 21% in 2012.[22]

The usage of survey responses is susceptible to the problem that responses to delicate questions, such as those concerning discrimination, can be subject to a social desirability bias. A respondent might claim not to have discriminatory taste due to social desirability, which might not reflect real preferences. In order to validate that we are capturing a real trend in discriminatory taste, in Figure 3 we plot our second measure of taste for discrimination at the aggregate level against a range of racially-motivated hate crimes committed in the US against blacks (namely the number of total victims, murder and manslaughter, forced rape, aggravated assault, simple assault, and intimidation).[23] The hate-crime statistics are obtained from the

[22]In the Online Appendix we decompose the time trends of belief about discrimination by region.

[23]We do not plot the first measure for taste for discrimination because it only overlaps with the available hate crime statistics for only four years. However, the correlations are

Federal Bureau of Investigation (FBI) Uniform Crime Reports for the years 1996-2012.[24] The proxy for taste for discrimination aggregated at the national level (grey dashed line) follows a downward trend closely resembling the downward trend for all racially-motivated hate crimes against blacks with correlations ranging between 0.76 for forced rape to 0.91 for murder and manslaughter. Racially-motivated hate crimes could be seen as extreme expressions of discriminatory taste, wherefore the exhibited patterns strengthen the validity of our taste for discrimination measure. [25]

B. The method

Following the theoretical model we derive the equation for the estimation of the probability of individuals being self-employed as a function of ability a, the proportion π_{tqs} of principals with a taste for discrimination at time t in region q against group s, the proportion φ_{tqs} with beliefs about discrimination at time t in region q against group s, and a vector of individual characteristics X_i with associated parameter vector γ. As a proxy for ability we use years of schooling. Therefore, the probability of individual i at time t in region q and of group s being self-employed $se_{itqs} \in \{0, 1\}$ we define as

$$Prob(se_{itqs} = 1|a, \varphi_{tqs}, \pi_{tqs}, X_i) = f(a, \varphi_{tqs}, \pi_{tqs}, X_i). \tag{15}$$

similarly high and available upon request.

[24]For FBI hate-crime statistics against blacks see Table 7 of the reports to be found at `http://www.fbi.gov/about-us/cjis/ucr/hate-crime`.

[25]However, this downtrend could simply reflect an overall downward trend in criminal activity. In order to reject this hypothesis, we normalize the occurrence of racially-motivated hate crimes by the total frequency of arrests of white individuals for murder and manslaughter as well as aggravated assault. Unfortunately, we do not have data on arrests for all of the before mentioned crimes. As can be seen in the Online Appendix, the patterns corroborate the assumption that our measure is a reasonable proxy for taste for discrimination by exhibiting high correlations between taste for discrimination and the normalized hate-crimes of 0.89 and 0.85.

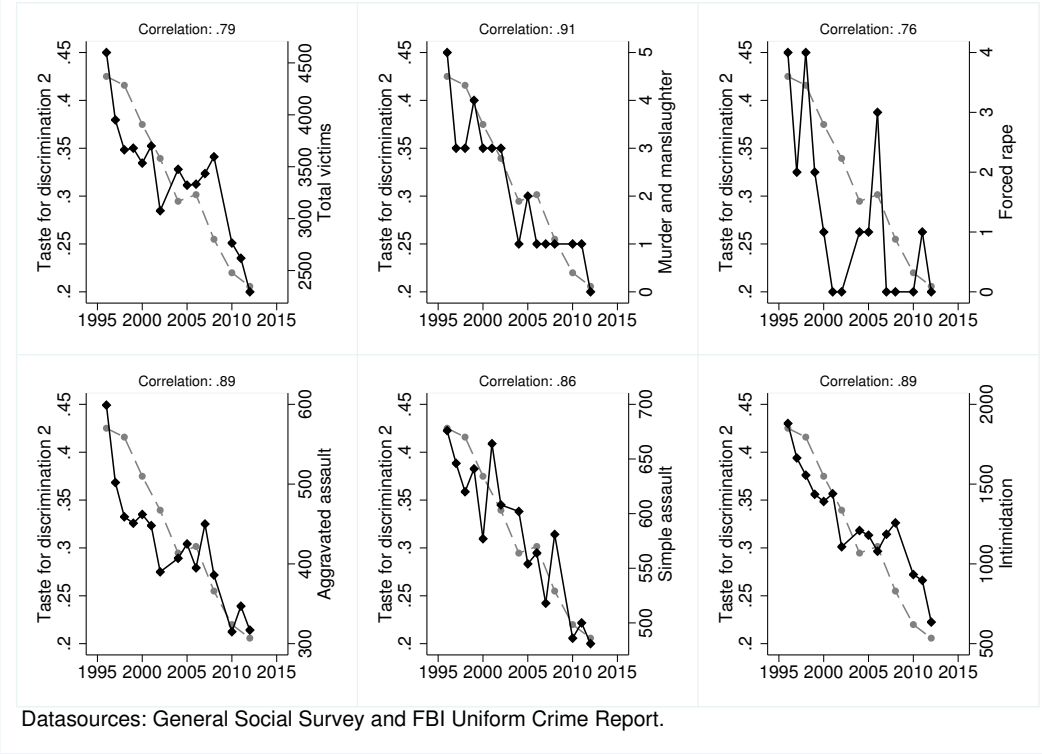


Figure 3. : Racially-motivated hate crimes (solid black line) versus taste for discrimination (grey dashed line)

Using a logit regression, we can define the estimated probability as

$$Prob(se_{itqs}|a, \varphi_{tqs}, \pi_{tqs}, X_i) = \frac{e^{g(a,\varphi_{tqs},\pi_{tqs},X_i)}}{e^{g(a,\varphi_{tqs},\pi_{tqs},X_i)} + 1}, \tag{16}$$

where $g(a, \varphi_{tqs}, \pi_{tqs}, X_i) = \beta_0 + \beta_1 a + \beta_2 \varphi_{tqs} + \beta_3 \pi_{tqs} + \gamma X_i + \varepsilon_{itqs}$ and ε_{itqs}, the error term, is a binomially distributed random variable.

The proportion of principals with a taste for discrimination π_{tqs} and the proportion with beliefs about discrimination φ_{tqs} take the value zero for white individuals, i.e. for $s = A$. We restrict our sample to white and black respondents who are not students or retired, while assuming no differences

in preferences to become self-employed.[26] In the logit regression, estimating self-employment we control for gender, age, age squared, and whether the father was self-employed. All specifications include time and region fixed effects.

C. Results

The reports of the baseline regression are reported in Table 1. Years of schooling, our proxy for ability, and all controls have the expected sign and are significant at the 1% level in all specifications. The probability of being self-employed is increasing in years of schooling, hump shaped in age, females are less likely to be self-employed, and having a father that was self-employed increases the probability of self-employment.

In columns (1) and (2), we show that either proxy for taste for discrimination against blacks is a significant negative correlate of self-employment only as long as the proxy for belief about discrimination does not enter the model. Once belief about discrimination enters the model, either proxy for taste for discrimination becomes insignificant as can be seen in columns (3) and (4). The variable representing belief about discrimination is significant at the 1% level when paired with taste for discrimination. In columns (5) and (6), we add a race dummy for blacks to validate that unobservables correlated with being black are not what are actually driving our results. The race dummy turns out to be insignificant, whereas belief about discrimination remains a significant negative correlate.

In column (7) we pair belief about discrimination with the black race

[26]In the International Social Survey on Work Orientation III, we find that 71% of blacks versus 58% of whites in the labour force in the US would choose self-employment if they could choose between different kinds of jobs, suggesting that our estimates might even be underestimating effects of discrimination.

Table 1—: Baseline logistic regression

Dependent variable: Self-employment							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Belief about discrimination			-2.781***	-3.947***	-3.308**	-2.947*	-1.638*
			(0.682)	(1.025)	(1.296)	(1.594)	(0.966)
Taste for discrimination 1	-4.240***		0.084		-0.156		
	(0.851)		(1.221)		(1.331)		
Taste for discrimination 2		-2.161***		0.780		0.837	
		(0.436)		(0.806)		(0.797)	
Black					0.164	-0.229	-0.330
					(0.340)	(0.283)	(0.230)
Years of schooling	0.049***	0.037***	0.048***	0.036***	0.048***	0.036***	0.036***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.007)
Female	-0.696***	-0.694***	-0.694***	-0.692***	-0.695***	-0.692***	-0.727***
	(0.051)	(0.052)	(0.051)	(0.052)	(0.051)	(0.052)	(0.041)
Age	0.097***	0.095***	0.098***	0.095***	0.099***	0.095***	0.097***
	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)	(0.009)
Age squared x 1,000	-0.754***	-0.693***	-0.766***	-0.699***	-0.766***	-0.698***	-0.744***
	(0.106)	(0.109)	(0.106)	(0.109)	(0.106)	(0.109)	(0.087)
Father was self-employed	0.630***	0.618***	0.626***	0.613***	0.627***	0.612***	0.627***
	(0.052)	(0.055)	(0.052)	(0.055)	(0.052)	(0.055)	(0.043)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.063	0.063	0.065	0.065	0.065	0.065	0.065
Observations	16104	14719	16104	14719	16104	14719	23895

Note: All regressions include a constant. Standard errors are in parentheses.
*, ** and *** significant at 10, 5 and 1 % significance level, respectively.
Source: General Social Survey.

dummy. Again observe that belief about discrimination is a negative and significant correlate of self-employment, while the race dummy is negative but insignificant. If we were to interpret the correlation as causal, the expected effect on an average black male of eliminating belief about discrimination from the sample average of 23% to 0% would raise the self-employment probability from 7.7% to 10.9%, which is an increase of 42%. The magnitude of the effect of belief about discrimination becomes clear when we calculate the probability of the average black male being self-employed whilst removing the effect of unobservables correlated with race, but holding constant belief about discrimination at the average level of 23%. Here the increase in the probability of self-employment is lower in magnitude, increasing from 7.7%

to 10.4%. Eliminating both, belief about discrimination and unobservables correlated with race, predicts a self-employment probability of 14.5% for the average black male, which is still lower than then the 15.9% probability of the average white male. This gap can, amongst others, be attributed to lower levels of education and demographic factors. While a fully causal interpretation is farfetched, the results indicate that well established beliefs about discrimination might be sufficient to lower the probability of self-employment for blacks.[27]

D. Additional stylized evidence and applications of the theoretical framework

In this subsection, we present further evidence in the form of recent findings in the empirical and behavioral literature that our theoretical framework can reconcile. We then go on to highlight how the presented framework can be also useful in analyzing issues such as the phenomenon of racial tipping points in American neighborhoods.

Alesina, Lotti and Mistrulli (2013) find that banks in Italy charge self-employed women more than self-employed men for credit. They find that characteristics such as riskiness, type of business, or differential bank choice cannot explain their result. They also find that the effect is not restricted to any particular geographical region and taste based indicators of discrimination cannot explain the observed pattern. As women businesses need to establish interlinkages, beliefs of banks that potential productive *male*

[27]Neither including a measure for statistical discrimination nor excluding farmers from self-employment changes the results qualitatively. The results can be found in the Online Appendix.

links might discriminate against women, might result in banks discriminating against women too. Consistent with our theoretical model the authors find that that banks discriminate more against women in sectors, where men dominate, and can be interpreted as being more likely to be matched with a discriminatory male link.

The mechanism put forth is also a plausible explanation for features highlighted in data for the market for self-employment in India and Sweden. For instance, why the Schedule Castes (SCs) and Schedule tribes (STs), the socially most disadvantaged groups in India are relatively more underrepresented in urban rather than rural areas in terms of non-farm enterprise ownership, even though discrimination is higher in rural areas (Iyer, Khanna and Varshney 2013). Why in Sweden, one of the countries where women's labor force participation rate is very high and only 0.4% of the male population strongly agree that men make better business executives than women, has among the lowest level of self-employment for women in the EU.[28] The fact that beliefs about discrimination are higher in urban rather than rural areas in India, and remain high in Sweden concerning women, could be an important explanatory factor.[29]

Daskalova (2013) documents in a lab experiment that people who do not discriminate when making decisions individually, discriminate while making joint decisions due to beliefs about what their co-decision maker

[28]The wave of 2005-2007 of the World Values Survey exhibits that 43% of the Swedish population mention "Discrimination against women and girls" as one of the two most pressing problems facing the country. Moreover, Swedish males have the lowest bias against women across all 39 countries in the sample, while beliefs about discrimination are the second highest.

[29]Observe that coordination failures in urban markets are more likely as they are anonymous, so even if taste for discrimination is higher in the rural than urban settings, it could well be the case that the coordination failures in urban areas outweighs the taste for discrimination effect in rural areas, leading to the outcome observed in the data.

will do. Albrecht et al. (2013) find that in the lab individuals are conservative in updating their beliefs, which points to another channel through which beliefs regarding discrimination might become sticky over time and be an important determinant of outcomes for the discriminated group.

Our model is also applicable to a range of markets with strategic complementarities. The dominance of particular ethnic groups in certain professions (Greif 1989, 1993; Banerjee and Munshi 2004) might be explained through our mechanism as ethnic enclaves might help secure complementary support from other individuals and overcome coordination failures.[30]

Card, Mas and Rothstein (2008) assume that when black people move into a neighborhood, white neighbors with a distaste for blacks will change neighborhoods. Anticipating a decrease in housing prices, people without a distaste for black neighbors will also sell their property and move. We show that the presence of neighbors with a distaste for black neighbors is not required to trigger the segregating dynamics, the belief is sufficient, hence providing an alternative explanation for the phenomenon of racial tipping points in the United States.

III. Conclusion

In this paper we show that even once taste for discrimination and statistical discrimination were to cede to exist in society, discrimination can persist due to remaining beliefs making discrimination the best-response, a much weaker condition than traditionally assumed in the literature.

[30]This benefit, however, has to be weighed against the restriction on occupational choice that might arise due to ethnic enclaves being effective gate keepers to certain professions.

The theoretical mechanism put forth is relevant for markets characterized by the need to establish productive relations or interlinkages with other agents in society in order for the production process to be carried out. It is shown that in such markets the presence of beliefs regarding the existence of taste discriminators, even when no agents with taste for discrimination exist in society, can result in agents exhibiting discriminatory behavior in equilibrium. Discrimination arises as a rational response to the belief that other agents might discriminate, which would impose losses due to the complementarity in the production process. The model shows lower participation and payoff to the discriminated group in markets characterized by the presence of interlinkages.

Suggestive empirical evidence in support of the theoretical framework is provided by analyzing the market for self-employment, a market characterized by the need to establish productive relations to be able to operate and be successful. The outcomes predicted by the model, in terms of participation rates and incomes for the self-employed for the discriminated group being lower, are confirmed using data from the General Social Survey 1972-2012 of the US and creating proxies for taste and beliefs regarding discrimination. We validate that the downward timetrend of our proxies of taste for discrimination do not necessarily reflect a social desirability bias, as the proxies are strongly correlated with racially-motivated hate crimes against blacks. A simple logit model shows that beliefs about discrimination are a significant negative correlate of self-employment for blacks, even after controlling for individual level characteristics, as well as region and year fixed effects.

The nature of discriminatory coordination failures does not allow market forces to overcome discrimination and may require alternative policy

tools. The various mechanisms through which discrimination manifests its dynamic linkages in terms of cross market and intergenerational effects, and the tendency to persist through cumulative and belief based channels, need to be understood and explored in order to develop policies aimed at eradicating discrimination and achieving equal treatment and opportunities.

Appendix

A1. Proofs

PROOF:

Proposition 1

1) First observe that only B-type individuals in the ability range $a^* \leq a < a_b^*$ offer $1 + r \leq \sigma^N(a) < 1 + r + b$, and hence can potentially face discrimination. The fact that there exists no B-type in the range $a^* \leq a < a_b^*$ that prefers to offer the Nash bargaining solution while attempting to enter activities characterized by interlinkages implies:

$$\text{(A1)} \quad (1 - \varphi_{T^*})^2(\lambda a - 2\sigma^N(a)) + (1 - (1 - \varphi_{T^*}^h)^2)(a - \delta) - a < 0,$$

where the left hand side (LHS) is the expected net payoff from offering the Nash bargaining solution minus the payoff from entering the activity involving no interlinkages. The LHS is strictly increasing in a, which implies that if it is not satisfied for a_b^* then it is not satisfied for all $a \leq a_b^*$. This implies that all B-types with a such that $\lambda a - 2(1 + r + b) \geq a$ (or all $a_b \leq a \leq a_b^*$) offer a share equal to $(1 + r + b)$, while being accepted by the principals and success-

fully enter activities characterized by interlinkages. All B-types with $a^* \leq a < a_b$ will be unable to offer a share to compensate the taste of discriminators, and hence enter activities involving *no* interlinkages. Moreover, note as now all individuals of the B-type with $a \geq a_b^*$ offer $\sigma^N(a) \geq (1+r+b)$, and all B-types with $a_b \leq a \leq a_b^*$ offer $(1+r+b)$, this implies that from period T^* onwards there will be no offers by a B-type made within the range of $(1+r)$ to $(1+r+b)$. Therefore, beliefs will remain frozen at the current level implying the above equilibrium will persist for ever.

2) Let us denote by a_{lb} as the lowest B-type in the range $a^* \leq a < a_b^*$ at T^* who prefers offering the Nash bargaining solution and attempts to enter the activity involving interlinkages. As the Nash bargaining solution is strictly increasing in a, it implies that all individuals with $a \geq a_{lb}$ offer the Nash bargaining solution in period T^*. This means that all B-types in the ability range $a_{lb} \leq a \leq a_b^*$ offer the Nash bargaining solution and are accepted as no more principals with taste for discrimination exist. Assume the total number of cases subject to potential discrimination are n_{pot} and as no taste discriminators exist, the actual cases of discrimination are zero. This implies that the point estimates in the next period T^*+1 for meeting a discriminator is given by $\varphi_{T^*+1} = \frac{\alpha_0^l + \sum_{t=1}^{T^*} k_t}{\alpha_0^l + \beta_0^l + \sum_{t=1}^{T^*} n_{tb} + n_{pot}}$. Therefore, $\varphi_{T^*+1} < \varphi_{T^*}$, implying the lowest type who applies in T^*+2 is such that $a < a_{lb}$, or generalizing $\varphi_{T^*+t} < \varphi_{T^*}$ for all $t > 0$, or $\frac{d\varphi_t}{dt} < 0$ for all $t > T^*$. Hence, at some point $\varphi_T \to 0$, implying all B-types with $a \geq a^*$ apply and enter activities characterized by interlinkages, wherefore discrimination does not persist in society. The number of periods for which discrimination

will persist is as a function of $\varphi_{T^*} = f(\pi_0, \omega)$ and the B-type with the lowest a preferring the Nash solution to wage employment in T^*.

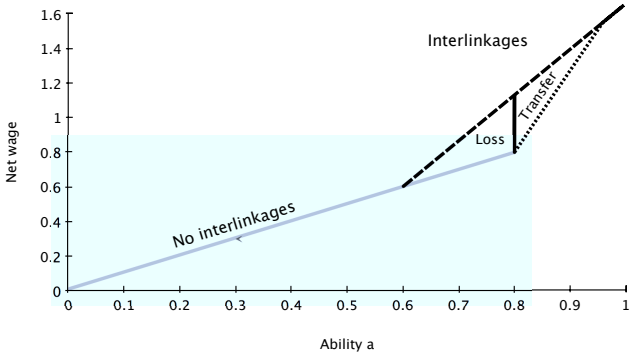
PROOF:

Proposition 2

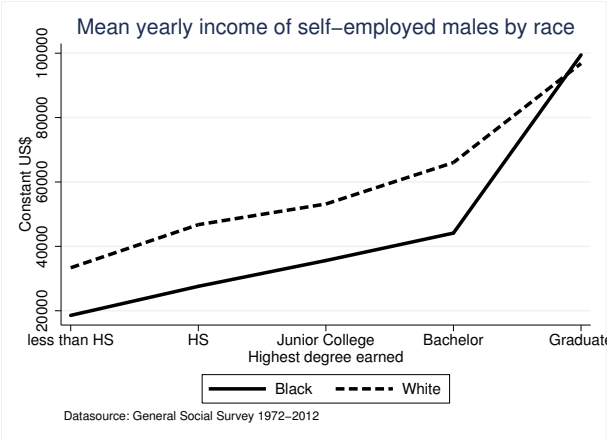
Assume the set $\mathcal{B}$ is large enough that individuals' abilities can be approximated by a continuous distribution on the interval $[0, 1]$. Proposition 1.1 shows that all individuals of the B-type with $a^* \leq a < a_b$ enter activities not involving interlinkages, whereas A-types of the same ability enter activities involving interlinkages. Moreover, B-type individuals with $a_b \leq a < a_b^*$ offer $1 + r + b > \sigma^N(a)$.

Now we can write the transfer, which principals receive from individuals wanting to escape the discrimination as an integral over the ability range who pay the higher fee: $\int_{a_b}^{a_b^*} 2(1 + r + b - \sigma^N(a))da = \int_{a_b}^{a_b^*} 2(1 + r + b - \frac{(\lambda-1)a+r+\delta+1}{3})da$. Therefore, the gain to principals would be $\frac{1}{3}((a_b^* - a_b)(4 + 4r + 6b - 2\delta) - (\lambda - 1)((a_b^*)^2 - a_b^2)))$. Now the loss to principals, due to able individuals entering activities not involving interlinkages, can be written as: $\int_{a^*}^{a_b} 2[\sigma^N(a) - (1 + r)]da = \int_{a^*}^{a_b} 2[\frac{(\lambda-1)a+r+\delta+1}{3} - (1 + r)]da \Rightarrow Loss = \frac{1}{3}((a_b^2 - (a^*)^2)(\lambda - 1) + (a_b - a^*)(2\delta - 4r - 4))$. In order for the loss to be at least as big as the gain we require: $((a_b^* - a_b)(4 + 4r + 6b - 2\delta) - (\lambda - 1)((a_b^*)^2 - a_b^2) \leq ((a_b^2 - (a^*)^2)(\lambda - 1) + (a_b - a^*)(2\delta - 4r - 4))$. Rearranging we get $(4 + 4r - 2\delta)(a_h^* - a^*) + 6b(a_b^* - a_b) \leq ((a_b^*)^2 - (a^*)^2)(\lambda - 1)$. Now substituting $a^* = \frac{2(1+r+\delta)}{\lambda-1}$, $a_b^* = \frac{2(1+r)+3b-\delta}{\lambda-1}$, and $a_b = \frac{2(1+r+b)}{\lambda-1}$ we find that for this to hold the condition is $b \geq \delta$. But since no discrimination exists when $\delta > b$ (because in that case the offer $\sigma^N(a) > 1 + r + b$ for all $a \geq a^*$) gains can never be greater than losses if we are in an equilibrium as in proposition 1.1.

A2. Figures



(a) In the example individuals of A-type become self-employed (black dashed line) when ability $a \geq 0.6$. B-type individuals only with $a \geq 0.8$ become self-employed (black dotted line) and those with $0.8 \geq a \geq 0.9$ pay higher rates resulting in lower net earnings.



(b) Mean yearly income of black and white self-employed by educational degree

Figure A1. : Income from interlinkages by type in model and data

*

REFERENCES

Aigner, Dennis J, and Glen G Cain. 1977. "Statistical theories of discrimination in labor markets." *Industrial and Labor relations review*, 30(2): 175–187.

Akerlof, George. 1976. "The economics of caste and of the rat race and other woeful tales." *Quarterly Journal of Economics*, 90(4): 599–617.

Akerlof, George. 1985. "Discriminatory, status-based wages among tradition-oriented, stochastically trading coconut producers." *Journal of Political Economy*, 93(2): 265–276.

Albrecht, Konstanze, Emma Von Essen, Juliane Parys, and Nora Szech. 2013. "Updating, self-confidence, and discrimination." *European Economic Review*, 60: 144–169.

Alesina, Alberto, Francesca Lotti, and Paolo Emilio Mistrulli. 2013. "Do women pay more for credit? Evidence from Italy." *Journal of the European Economic Association*, 11(s1): 45–66.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the Origins of Gender Roles: Women and the Plough." *Quarterly Journal of Economics*, 128(2): 469–530.

Argenziano, Rossella, and Itzhak Gilboa. 2012. "History as a coordination device." *Theory and Decision*, 73(4): 501–512.

Arrow, Kenneth J. 1973. "Discrimination in Labor Markets." , ed. O. Ashenfelter and A. Rees, Chapter The Theory of Discrimination, 3–33. Princeton: Princeton University Press.

Bailey, Thomas, and Roger Waldinger. 1991. "Primary, secondary, and enclave labor markets: A training systems approach." *American Socio-*

logical Review, 56(4): 432–445.

Banerjee, Abhijit, and Kaivan Munshi. 2004. "How efficiently is capital allocated? Evidence from the knitted garment industry in Tirupur." *The Review of Economic Studies*, 71(1): 19–42.

Basu, Kaushik. 2010. *Beyond the invisible hand: Groundwork for a new economics.* Princeton University Press.

Becker, Gary S. 1957. *The economics of discrimination.* University of Chicago Press.

Bennett, Elaine. 1997. "Multilateral bargaining problems." *Games and Economic Behavior*, 19(2): 151–179.

Bisin, Alberto, and Thierry Verdier. 2001. "The economics of cultural transmission and the dynamics of preferences." *Journal of Economic Theory*, 97(2): 298–319.

Blanchflower, David G. 2009. "Minority self-employment in the United States and the impact of affirmative action programs." *Annals of Finance*, 5(3-4): 361–396.

Blanchflower, David G, Phillip B Levine, and David J Zimmerman. 2003. "Discrimination in the small-business credit market." *Review of Economics and Statistics*, 85(4): 930–943.

Borjas, George J. 1986. "The self-employment experience of immigrants." Working Paper 1942, National Bureau of Economic Research, Inc.

Burguet, Roberto, and Ramon Caminal. 2011. "Simultaneous Nash bargaining with consistent beliefs." Barcelona GSE Working Paper Series Working Paper no 521.

Card, David, Alexandre Mas, and Jesse Rothstein. 2008. "Tipping and the Dynamics of Segregation." *Quarterly Journal of Economics*,

123(1): 177–218.

Coate, Stephen, and Glenn C Loury. 1993. "Will affirmative-action policies eliminate negative stereotypes?" *American Economic Review*, 83(5): 1220–1240.

Daskalova, Vessela. 2013. "Discrimination, Social Identity, and Coordination: An Experiment." mimeo.

Dessí, Roberta. 2008. "Collective memory, cultural transmission, and investments." *American Economic Review*, 98(1): 534–560.

Fairlie, Robert W. 1999. "The absence of the African-American owned business: An analysis of the dynamics of self-employment." *Journal of Labor Economics*, 17(1): 80–108.

Fairlie, Robert W. 2002. "Drug Dealing and Legitimate Self-Employment." *Journal of Labor Economics*, 20(3): 538–537.

Fairlie, Robert W, and Alicia M Robb. 2008. "Race and entrepreneurial success: Black-, Asian-, and White-owned businesses in the United States." *MIT Press Books*, 1.

Fairlie, Robert W, and Bruce D Meyer. 1996. "Ethnic and racial self-employment differences and possible explanations." *Journal of Human Resources*, 31(4): 757–793.

Fairlie, Robert W, and Bruce D Meyer. 2000. "Trends in self-employment among white and black men during the twentieth century." *Journal of Human Resources*, 35(4): 643–669.

Fryer, Roland, and Matthew O Jackson. 2008. "A categorical model of cognition and biased decision making." *The BE Journal of Theoretical Economics*, 8(1): 1935–1704.

Gilboa, Itzhak, and David Schmeidler. 1995. "Case-based decision the-

ory." *Quarterly Journal of Economics*, 110(3): 605–639.

Greif, Avner. 1989. "Reputation and coalitions in medieval trade: evidence on the Maghribi traders." *The Journal of Economic History*, 49(4): 857–882.

Greif, Avner. 1993. "Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition." *American Economic Review*, 83(3): 525–548.

Griffin, Larry J, and Kenneth A Bollen. 2009. "What do these memories do? Civil rights remembrance and racial attitudes." *American Sociological Review*, 74(4): 594–614.

Halbwachs, Maurice. 1992. *On collective memory.* University of Chicago Press.

Iyer, Lakshmi, Tarun Khanna, and Ashutosh Varshney. 2013. "Caste and Entrepreneurship in India." *Economic and Political Weekly*, 48(6): 52–60.

Kaushal, Neeraj, Robert Kaestner, and Cordelia Reimers. 2007. "Labor market effects of September 11th on Arab and Muslim residents of the United States." *Journal of Human Resources*, 42(2): 275–308.

King, Miriam, Steven Ruggles, Trent Alexander, Sarah Flood, Katie Genadek, Matthew B Schroeder, Brandon Trampe, and Rebecca Vick. 2010. *Integrated public use microdata series, current population survey.* Minneapolis: University of Minnesota.

Lucas Jr, Robert E. 1978. "On the size distribution of business firms." *The Bell Journal of Economics*, 9(2): 508–523.

Lundberg, Shelly, and Richard Startz. 1983. "Private discrimination and social intervention in competitive labor market." *American Economic*

Review, 73(3): 340–347.

Lundberg, Shelly, and Richard Startz. 1998. "On the persistence of racial inequality." *Journal of Labor Economics*, 16(2): 292–323.

Mailath, George J, Larry Samuelson, and Avner Shaked. 2000. "Endogenous inequality in integrated labor markets with two-sided search." *American Economic Review*, 90(1): 46–72.

Moore, Robert L. 1983. "Employer discrimination: Evidence from self-employed workers." *The Review of Economics and Statistics*, 65(3): 496–501.

Neal, Derek A, and William R Johnson. 1996. "The role of pre-market factors in black-white wage differences." *Journal of Political Economy*, 104(5): 869–95.

Nunn, Nathan, and Leonard Wantchekon. 2011. "The slave trade and the origins of mistrust in Africa." *American Economic Review*, 101(7): 3221–3252.

Peski, Marcin, and Balázs Szentes. 2013. "Spontaneous discrimination." *American Economic Review*, 103(6): 2412–2436.

Phelps, Edmund S. 1972. "The statistical theory of racism and sexism." *American Economic Review*, 62(4): 659–661.

Rosén, Åsa. 1997. "An equilibrium search-matching model of discrimination." *European Economic Review*, 41(8): 1589–1613.

Smith, Tom W, Peter Marsden, Michael Hout, and Jibum Kim. 2012. *General Social Surveys*. National Opinion Research Center.

Voigtländer, Nico, and Hans-Joachim Voth. 2012. "Persecution perpetuated: the medieval origins of anti-Semitic violence in Nazi Germany." *Quarterly Journal of Economics*, 127(3): 1339–92.