A Political Economy of Social Discrimination

Dewan, Torun and Wolton, Stephane

London School of Economics and Political Science, London School of Economics and Political Science

8 June 2019

A Political Economy of Social Discrimination*

Torun Dewan† Stephane Wolton‡

Link to most recent version

June 8, 2019

Abstract

From burqa ban to minaret ban, from right to detain suspected illegal immigrants to restricting the help to migrants, the number of social laws specifically targeting a tiny proportion of citizens has raised in recent years across Western democracies. These symbolic policies, we show, are far from being innocuous: they can have far reaching consequences for large parts of the population. By raising the salience of certain social traits (e.g., Muslim identity) these laws can create a labour market loaded in favor of the majority (e.g., the non-Muslims), yielding higher unemployment rates and spells for minority citizens. These deleterious effects arise even absent any form of bias against, or uncertainty about, minority workers. Instead they are fully driven by social expectations about behavior and are best understood as a form of *social discrimination*. Importantly, we establish conditions under which a plurality of the citizenry demands the implementation of symbolic policies anticipating their labor market consequences. We further highlight that the implementation of symbolic policies is always associated with less redistribution and can be coupled with lower tax rates. We discuss several policy recommendations to limit the possibility of social discrimination arising.

*For helpful comments and suggestions, we thank Avi Archarya, Vessela Daskalova, Saurabh Pant, and seminar and conference participants at Caltech, Konstanz University, Northwestern University – Meds, Princeton University, Stirling University, 2018 MPSA Annual Conference, Petralia Workshop 2018, IBEO Workshop 2018, the 2018 Comparative Politics and Formal Theory Conference at Yale University, 2019 King's College Fair Play Conference, 2019 TSE/IAST Conference in Political Science and Political Economy. All remaining errors are the authors' responsibility

†London School of Economics and Political Science. Email T.Dewan@lse.ac.uk

‡London School of Economics and Political Sciecne. Email: S.Wolton@lse.ac.uk

"Do you know how many lawyers, doctors, and engineers have come out of these blocks? I see so many people studying, trying to become part of this country, but suddenly we are not good enough because we don't eat pork."

Mina, 30, about Agervang a Muslim dominated housing estate from news article "Denmark swings right on immigration — and Muslims feel besieged" (Guardian, 10 June 2018)

"Ever since this government has come in, I feel like people look at me and see a Muslim for the first time."

An Indian butcher referring to the beef ban from news article"Why a crackdown on Indian cattle trade is seen as anti-muslim" (Independent, 31st July 2018).

"If [the ban on full-face coverings] was intended as a contribution in the fight against conservative Islam, then I can only say: it's gone belly up."

Hermann Greylinger of the Austrian police union from news article "Austrian full-face veil ban condemned as a failure by police" (Guardian, 27th March 2018)

1 Introduction

That majorities can use the political process to discriminate against minorities is well understood, at least since Mill (1859) who coined the phrase "the tyranny of the majority." Discrimination may take the form of targeting goods and benefits in ways that marginalise minority groups, or through direct expropriation of their assets. In this paper we investigate a more subtle form of discrimination that can arise when political leaders choose to make minority identity a salient aspect of political choice. As examples, consider the introduction of policies (equivalently, the use of rhetoric) that highlights minority practice. This might take the form of forcing minority groups to identify as such via the adoption of specific clothing (equivalently prohibiting them from doing so), placing restrictions on their practices, or forcing them to adapt to behaviour favoured by the majority.

There is a long history of such policy. The Fourth Lateran Council of 1215 proclaimed that Jews and Muslims must wear clothing that distinguished them from Christians. According to Appiah (2018), the Cagots of the French and Spanish Pyrenees had no physical features, names or language distinguishing them from their neighbours, but were forced to identify themselves "with badges pinned to their clothing, often duck or goose feet, or fabric fascimiles." More recently, in 2016, 30 municipalities in France introduced a ban on the use of the Burkini, a full body length swimsuit worn by muslim women at the seaside. In Austria, Netherlands, Denmark and elsewhere bans have been put in place on Kosher and Halal produce. In 2014 a beef ban was introduced in several Indian states that, while not affecting the Hindu population who do not consume it, prohibited Muslims and Christians from consuming a relatively cheap form of protein.

An immediate consequence of symbolic policy, such as the proclamation of the Latern Council, or those with respect to the Cagots, is to make salient who is a member of the majority and who is not. Similarly, as Terrence G. Peterson (quoted in the *New York Times*, 18 August 2016) puts it, the burkini ban is "a way to police what is French and what is not French". More generally, these symbolic policies, even when enforcing majority patterns of behaviour (as in the beef ban in Indian states) make identity a focus of political discourse and economic interaction, heightening awareness of identity and the perception of others' identity. For example, French politics was about little other than the burkini in the summer of 2016, a period that (according to the Collective against Islamophobia in France which filed a complaint with the Counseil d'Etat, France's highest administrative court) " witnessed a hysterical political islamophobia that pits citizens against one another." The introduction of symbolic policies then can have a direct impact (as noted in our opening quotes) and it is this: a worker is no longer identified by their trade be it butcher, doctor or engineer; instead they are identified as a member of a specific group and judged as a member of that group irrespective of their human capital.

Building on this insight, we study the labour market consequences of these policies and in doing so show their effect to be far from innocuous. In fact, they can have a significant negative impact on minority welfare. Symbolic policies raising the salience of identity generates a labor market loaded in favor of majority workers and to the detriment of minority workers and employers. While majority workers find employment more easily, their minority counterparts experience higher unemployment and longer unemployment spells and minority owned firms see their profits decrease substantially.

This finding contrasts with official justifications for the introduction of symbolic policies that stem from first principles. For example, France's minister for women's rights, Laurence Rossignol, argued that "[The burkini] is the symbol of a political project that is hostile to diversity and women's emancipation." Whereas the French Prime Minister described the burkini as "a symbol of enslavement of women."[1] Similarly, the recent headscarf ban in Austrian school has been justified on the grounds that "the

[1]The burkini ban: what it really means when we criminalise clothes, *Guardian, 24 August 2016*

measure was necessary to free girls from subjugation."[2] According to this view, enforcement of the burkini or niqab bans was necessary in order to emancipate muslim women. They were to "be forced to be free" according to Rousseau's famous dictum.[3] Such normative claims are contentious. From a liberal perspective it is hard to understand why representatives of the state should take a view on what women wear to the beach, or what Muslims have for lunch. Moreover, the target audience of these contentious symbolic policies are often very small: it is estimated the the headscarf ban affected just 150 women in Austria and between 150 and 200 women in Denmark.[4]

Our assumption, instead, is that the the impact of these policies makes minority markers salient. Building on that assumption we show that their introduction can serve the electoral interests of political leaders.–and we explore the conditions under which this is so. At first glance, this might not appear as a novel finding. The large literature on conflict documents many such cases. Indeed, in discussing the causes of ethnic conflict, Fearon and Laitin (2000) remark that "political elites use violence to construct antagonistic ethnic identities, which in turn favour more violence" with the masses following. Our contribution is in showing that similar strategies are deployed in democracies where redistribution takes place via political competition rather than through conflict. Furthermore, a critical difference emerges in the role played by leaders in these contexts. In contrast to the top-down strategies typically deployed by warlords, who use identity to mobilise and galvanise fighters for their cause, we show that democratically elected leaders respond to the demands of majority workers when making identity salient to political choice. The demand is, in turn, due to the advantages that such workers face when the labour force is segregated.

Our insights into the cause and consequences of symbolic policies stem from a political economy model that connects political choice via (two) party competition and a labor market. Specifically we study a dynamic model where politicians compete by choosing whether to introduce a "symbolic" policy or not alongside a proportional tax rate that is uniformly distributed. Once that policy environment is set, workers who vary by their type (given by their membership of a majority or minority group and their productivity) and firms are (randomly) matched. Productivity of a worker-firm relationship

[2]Austria approves headscarf ban in primary schools, *Guardian, 16 May 2018*

[3]In his *Du Contrat Social: Ou Principes du Droit Politique* published in 1762, Jean Jaques Rousseau states that"Whoever refuses to obey the general will will be forced to do so by the entire body, this means merely that he will be forced to be free."

[4]For Austria, see "Austrian full-face veil ban condemned as a failure by police, " *The Guardian*, March 27th 2018. The same article notes that the legislation was criticised by police after it mainly resulted in the issuing of warnings against people wearing smog masks, skiing gear and animal costumes. For Denmark, see "Protests in Denmark as 'burqa ban' comes into effect," *The Guardian*, August 1st 2018.

is match-specific and is unaffected by identity markers. Both majority and minority workers are ex-ante the same, leaving no room for any form of statistical discrimination. Employment relations are voluntary: both firm and worker must agree to a match before productivity takes place. We make two crucial assumptions regarding identity. First, a worker or employer's physical (minority or majority) traits are salient, and market relations can be conditioned on identity, only in an environment where the elected leader implements the symbolic policy. Second, we build on earlier work by Peski and Szentes (2013) in assuming that identity traits are not just physically fixed, but also socially malleable; they can be transferred across groups so that a majority(minority) employer who hires a minority(majority) employee might take on their social identity and vice-versa.

Our first results connect symbolic policies and labor market discrimination. Whereas historical episodes of overt work discrimination against minorities are common, such discrimination can also emerge endogenously via choices made by employers and workers in a climate influenced by the policy choices and rhetoric of politicians that makes identity focal. Labor market discrimination arises when majority (minority) employers will only hire majority (minority) workers and majority (minority) workers only accept employment from majority (minority) employers. We explore when such discriminatory strategies are best responses. This depends upon the relative proportion of majority workers and employers. There exists a "zone of discrimination" where the relative proportion of majority workers to employers is neither too large or too small. Within this zone, a segregated labor market where majority (minority) workers are hired by majority (minority) owned firms can emerge. The existence of this zone provides bad news for minorities. Within it, the level of minority unemployment and the duration of their unemployment spells are higher than those of their counterparts in the majority and than would be the case in a more benign environment. Furthermore, we show that workplace discrimination delivers a productivity gap between firms owned by majority employers and those owned by minority ones.

Notably, these effects – a segregated labour market with relatively poor employment prospects for minority workers and a productivity gap between majority owned and minority owned firms – arises solely from the adoption of identity-based employment strategies. So these phenomena may be entirely unconnected to differences in human capital in subgroups of the population. Moreover, they need not stem from taste-based discrimination that would arise should majority members intrinsically dislike interacting with their minority counterparts or from statistical discrimination that would arise should a decision-maker uses observable characteristics (such as physical traits that allow her to categorize agents) as a proxy for otherwise unobserved human capital that is relevant to economic outcomes. Instead, the discrimination we uncover is fully driven by social expectations: the anticipated consequences of hiring minority workers, working for a minority owned firms for majority employers and workers, respectively. As such, the phenomena we describe are best understood as a form of social discrimination.

We further show labour market discrimination can benefit majority workers thereby introducing demand for such policy. This demand (we again stress) is not driven by intrinsic taste for discrimination (majority workers do not care what minorities eat or how they dress). It arises due to the anticipated labour market effects. Our focus on the demand for social discrimination reveals that it need not arise due to coordination failure or commitment problems—it might be desired by workers.

We use our political economy model to determine when symbolic policies are implemented and their impact on redistribution. A necessary condition for symbolic policies to pass democratically is that the proportion of majority workers is neither too small or too large. When it is too small, the majority workforce lacks the political clout to force politicians to deliver. When it is too large then labor market discrimination offers insufficient benefits. Indeed, in a segregated labor market, majority workers are unable to take up employment in minority firms. The gain from discrimination (better employment prospects with majority employers) dominates the loss (no work relationship with minority employers) only if the proportion of minority owned firms is not too large. In particular, we show that the demand for symbolic policies and its associated social discrimination is positive only if the minority is poorly integrated: there are relatively more minority workers than minority firms. A surprising implication is that (when there is no explicit prejudice against minorities) policies designed to increase the size of the minority (perhaps through immigration) can make social discrimination less likely.

Symbolic policies do not only affect labor market outcomes, they also impact redistribution through two distinct channels. First, labor market discrimination changes the fiscal landscape. The revenue from taxes on production are lower when feasible and productive working relations involving social interaction between majority and minority are prohibited. Second,because majority workers face higher employment probability with discrimination than without, the cost of taxation increases, reducing their demand for redistribution. Combining these two effects, redistribution is always lower with discrimination and, sometimes, the tax rate can also decrease. As a result, minority workers suffer doubly from symbolic policies, their employment prospects diminish because of social discrimination and the redistributive transfers are reduced when they need it most.

Our equilibrium analysis delivers several central messages that highlight how political and economic institutions interact in ways that harm minorities. We conclude our analysis by focussing on policy measures that might alleviate these negative welfare effects.

2 Literature review

Our paper relates to the economics literature on identity and its effect on discrimination and redistribution. Here we highlight some of these papers and our contribution to this literature.

We relate to a classic literature on discrimination in the labour market that goes back to Becker (2010) who focussed on taste based discrimination.[5]. That identity can be a source of discrimination in equilibrium even though employers have no taste for it is shown by Arrow et al. (1973) and Phelps (1972) who developed the notion of statistical discrimination. In these models employment discrimination occurs when employers believe that physical attributes are correlated with human capital investments. Our paper suggests another source of labor discrimination. Even though all agents are known to be identical (no room for statistical discrimination), even though no individual harbours a distaste for minorities (no room for taste-based discrimination), discrimination can arise due to discriminatory norms in the labour market.

Several contributions have studied how payoff-irrelevant identity traits can serve as a focal point in labor market interactions. Starting with Akerlof's (1976) analysis of castes, several papers have shown how the fear of punishment by one's own group can create segregation in the labour market. In Bramoullé and Goyal (2016), fear of being ostracized by in-group members sustains favouritism whereby a firm always hires workers with the same identity trait as its owner even if more productive matches are available. Choy (2017) shows how identity can endogenously generate a social hierarchy, whereby some groups have superior outcomes and cooperate more than other groups. In his work, segregation is sustained because members of upper groups are judged less trustworthy than their peers when interacting with members from groups of lower standing. Discrimination can also be sustained as a form of self-fulfilling prophecy. In Harbaugh and To (2014), discrimination arises because minority consumers expect to be cheated by majority-owned firms and, therefore, refuses to invest in exploring possible beneficial relationship. This insight is close in spirit to Eeckhout (2006) who shows that expectations of low level

[5]Recent contributions have shown how even if the vast majority of majority members are unbiased, the simple fear of meeting a biased members can lead to a complete breakdown of relationships between minority and majority (e.g., Basu, 2005; Ramachandran and Rauh, 2013)

of collaboration in mixed marriages can sustain higher level of cooperation in same-race relationships. Unlike Eeckhout (2006), however, the equilibrium characterized in Harbaugh and To (2014) is always pareto-dominated. Cavounidis and Lang (2015) describes a labor market equilibrium in which firms monitor more closely black workers, which decreases the average productivity of unemployed minority members (as less productive workers are screened and fired), and sustain the monitoring of employers in the first place. Finally, Kamphorst and Swank (2016) shows how discrimination can sustain higher average effort by employees.

We choose another avenue to sustain discrimination as an equilibrium outcome. We build upon work by Peski and Szentes (2013). They developed the notion that the perceived identity of a player changes as a result of his/her social interactions. In that paper, agents who are matched must decide whether or not to enter into a profitable relationship with each other. Each agent has a fixed physical identity and a malleable social one. Social color conveys information about who the agent has partnered with in her employment history. As in our paper, Peski and Szentes show that discrimination can arise spontaneously in equilibrium: Agents with the majority trait fear the consequences of social contamination that would leave them enjoying less opportunities should they interact with minority members. There are several technical differences between our approach and theirs. In Peski and Szentes (2013), an individual is randomly allocated to the role of employer or employee upon matching. Instead, we consider a more canonical labor market approach where the population is divided between firm owners and workers. This allows us to determine the full set of economic consequences of social discrimination which cannot be foreseen from Peski and Szentes's paper (Peski and Szentes focus on the properties of equilibrium with discrimination rather than their consequences). In addition, we embed our modified Peski and Szentes's framework in an institutional environment that encompasses a labour and a political market and we explore the demand and supply of social discrimination (in Peski and Szentes, 2013, there is no demand for discrimination as all agents lose when identity is salient). A further contribution is in exploring the labour market and fiscal consequences of discrimination. Finally, our model also suggests remedies to ameliorate the negative consequences on minority welfare.

In our set-up, individuals' labour market and political decisions are a function of identity and its salience. As such, our work connects with the large literature pioneered by Akerlof and Kranton (2000) (see also Akerlof and Kranton, 2010). Austen-Smith and Fryer Jr (2005) explore the "acting white" phenomena whereby human capital investments are stigmatised by peers. In their model agents face a tradeoff in that signals that induce high wages (educational attainment) are also those that induce peer rejection. In Eguia (2017), investment in the majority group identity attributes from minority members is used as a screening device. Only minority members who show high levels of investment are assimilated and benefit from labor market opportunities available to majority members. These also happen to be the most productive minority members who have the most to gain from assimilation. In a related contribution, Schnakenberg (2013) highlights how individuals can use symbolic political behavior (e.g., participation in protests) to signal the strength of their attachment to their identity. This signal improves their interactions with members of their own group, but comes at the cost of deteriorating relationships with out-group members. In Carvalho (2012), investment in identity attachment, in turn, serves as a self-commitment device to avoid yielding to temptation. In contrast, in our paper, identity is not so much an individual choice as a social and political construct. Identity becomes politically salient when politicians choose to make it so, anticipating that workers and employers will condition their behaviour on it.

Our paper studies the relationship between identity and redistribution and so relates to several papers in political economy with a similar focus. Levy (2004) develops a model with endogenous party formation showing that, when an identity dimension exists, party formation can lead to targeted redistribution to groups with a majority identity; while Fernandez and Levy (2008) shows empirical support for the model. Relatedly, Krasa and Polborn (2014) and Matakos and Xefteris (2017) show how a candidate's attributes (e.g., her/his race as in Matakos and Xefteris, 2019) and how they are viewed by the electorate can affect the politician's position on redistributive issues (see also Landa and Duell, 2015, for the role of identity traits in a political agency set-up). In these contributions, identity is fixed. Instead, Shayo (2009) analyzes agents identification with class or nation and its relationship to redistribution. In a similar setting, Grossman and Helpman (2018) study the effect of identity on the demand for protectionist trade policy. In turn, Penn (2008) look how institutions can shape identity choice, while Huber (2017) explores how the social structure of a society relates to the salience of class and ethnicity. Our contribution to this literature is twofold. We endogenize the dimensions of political competition by allowing politicians to choose the salience of identity politics and we microfound the effect of identity on taxation and redistribution via social discrimination in the labour market.

3 Set-Up

We consider a set-up with two candidates A and B and a population of mass 2 which incorporate a one-shot electoral competition stage followed by infinitely-repeated labour market stages. The population is evenly divided between workers and employers. However, while all workers have the right to vote, only a mass f of employers are citizens, the remaining correspond to foreign-owned firms or public administration. The population is also characterized by its physical identity: actors exhibit either majority trait M or minority trait m. This identity is based on identifiable features, be it religious (e.g., Catholics v. Muslims), racial (e.g., White Europeans v. North Africans), or even first-names (e.g., the experiments in Adida, Laitin, and Valfort, 2016). Physical identity is not always salient socially. Its importance depends on the policies in place in the country, which are a function of the electoral game that we now describe.

Electoral competition

Competition takes place between two candidates A and B who are office-motivated. They receive a payoff normalized to 1 from being in office (without loss of generality) and 0 otherwise. The candidate obtaining the most votes wins office (ties are decided by a fair coin toss). In order to be elected, candidate $J \in \{A, B\}$ proposes a platform q_J. This contains two policies.

The first policy offered by candidate J is a proportional tax rate on income $\tau_J \in [0, 1]$. Revenues from taxation, which we denote $R(\tau)$, are uniformly redistributed to workers (supposing that all citizens receive transfer does not change our results). In particular, candidates cannot propose targeted transfers to workers according to their (majority or minority) identity. We assume that there is some deadweight loss of taxation (in a reduced form) so that total transfers are $T(\tau) = K\big(R(\tau)\big)$, with $K(\cdot)$ $\mathcal{C}^\infty$ on $\mathbb{R}$, strictly increasing and concave, and satisfying $K'(0) < 1$ (for example, the function $K(x) = 1 - \exp(-\lambda x)$ with $\lambda < 1$ satisfies all assumptions). This deadweight loss assumption means that, as is common in the literature, we avoid full taxation.

The second policy proposed by candidate J is a symbolic policy $d_J \in \{0, 1\}$ targeting the physical minority. As discussed in the introduction, examples of such policies include a ban on wearing the burqa in public places, eating beef and so on. The effect of these policies is to make the physical identity, majority trait M or minority trait m, socially salient.

After observing the platforms q_J, the mass $1+f$ of citizens casts a ballot for one of the two candidates. We assume that, when indifferent, citizens vote against a candidate proposing the symbolic policy (if his opponent is not) and when indifferent candidates do not propose the symbolic policy. This assumption can be understood as due to (negligible) costs of implementing such policy. Or, alternatively, players might suffer a (negligible) moral disutility cost from implementing policies that could be viewed as discriminatory. This assumption also guarantees that the symbolic policy is implemented only if it impacts the labour market and lead to discrimination. However, all our results would hold as long as the symbolic policy *increases* the probability that actors condition their labor market behavior on identity traits.[6]

After the election is held and platforms are implemented, citizens and non-citizen employers interact in the labour market and so we now describe these interactions.

Labour market

The labour market takes the form of an infinitely repeated game with discrete time periods denoted by $t \in \{1, 2, \dots\}$. Employers own a single firm with one position to fill. In what follows we, thus, use interchangeably the terms 'firms' and 'employers'.[7] If the position is filled in period t, the firm produces and the position remains filled the following period with probability $1 - \delta$, $\delta \in (0, 1)$. If the position is unfilled at t, then the firm is matched with a randomly drawn worker from the pool of unemployed. Both the would-be employee and the employer must agree to enter a working relationship and, if they do so, the firm produces in period $t + 1$. If either does not agree then the position remains unfilled at $t + 1$.

A match produces a quantity $\theta \in [0, 1]$ sold at an exogenous price of 1. We interpret θ as the worker's and thus firm's productivity. Both the workers and the employers observe θ before agreeing whether to enter a working relationship. If they do so, then in the next period, the revenue $1 \times \theta$ is split between a fixed exogenous wage $w \in (0, 1)$ for the worker and the remainder to the employer. A non-employed worker earns 0 on the labour market as does an employer whose firm's opening is unfilled.

[6]Our breaking rule for indifference guarantees that there is a direct link between the symbolic policy and discrimination in the labor market. Using the language of empirical analysis, we focus on the average treatment effect rather than the intention to treat effect.

[7]Combining this aspect of our setup with a mass 1 of employers dramatically simplifies our analysis. It guarantees that there are as many jobs as workers. The model could be extended to firms having more than one position, but this would complicate the identity transmission discussed below.

We assume that productivity is match-specific. Denote θ_k a worker's productivity in match k (whether there is a working relationship). We assume that θ_k is drawn i.i.d. for each match according to the Cumulative Distribution Function $S(\cdot)$. We impose that $S(\theta) = \theta$ (i.e., productivity is drawn from a uniform distribution) for ease of analysis.

Recall that transfers are uniformly redistributed to workers so the per-period payoff of an employer is:

$$U^F = \begin{cases} (1-\tau)(\theta - w) & \text{if position filled} \\ 0 & \text{if position open} \end{cases}$$

The per-period of a worker assumes the following form:

$$U^W = \begin{cases} (1-\tau)w + T(\tau) & \text{if employed} \\ T(\tau) & \text{if unemployed} \end{cases}$$

Because we consider an infinitely-repeated labour market, we suppose that all players discount the future with discount factor $\beta \in (0,1)$ so that all continuation values are well-defined.

The per-period timing in the labour market is as follows.

Filled position	**Open position**
1. Production realized	1. No production
2. Working relationship breaks down with probability δ. If so, position becomes open	2. Match occurs with available workers, employer observes the match-specific θ Matched employer and worker decide whether to enter working relationship. If so, position becomes filled
3. Payoffs realized, move to next period	3. Payoffs realized, move to next period

Our labor market model includes two forms of friction. First, matches are persistent. This is equivalent to assuming that labour laws make lay-offs difficult (e.g., require just cause). Second, production is delayed by one period after an employer and a worker agree to enter in a working relationship. Frictions are important to generate our results below. However, only one form of friction is necessary for discrimination to arise. Results would hold (though equilibrium conditions would obviously change) in the current timing if matches last one period (i.e., $\delta \to 1$). Similarly, our conclusion would be substantially unaffected if production starts immediately as long as the employment relationship can last.

When describing the labour market subgame, note that we have made no mention of identity. Indeed, it is important to stress that a worker's productivity, that is known by her employer, does not depend on identity (neither her own or that of her employer). This rules out statistical discrimination in employment practice. Further, a worker or employer's per-period payoff is independent of the identity of whom they work with. Thus, in our framework, there is no room for taste-based discrimination either. Discrimination, should it arise, can only be a phenomenon related to our notion of social identity that we now detail.

Physical and social identity

We assume that physical identity becomes socially salient only if a social discriminatory policy is implemented. More specifically, we follow Peski and Szentes (2013) and assume that each citizen is characterized by a two-dimensional type $(\phi, s) \in \{m, M\} \times \{\emptyset, m, M\}$.

The first coordinate (ϕ) corresponds to a citizen's physical identity that is fixed. We assume that there is a proportion $\alpha^W > 1/2$ of workers with physical majority trait and a proportion $\alpha^F > 1/2$ of employers with physical majority trait. The proportion of employers with physical majority trait is the same among citizens and non-citizens (to avoid carrying around too many parameters) though our results remain substantially unchanged were this not so.

The second coordinate (s) corresponds to a citizen's social identity which can vary with social interactions. Unlike Peski and Szentes, we assume that social identity is not always salient. Absent the implementation of symbolic policy, all citizens look the same and their social identity type is $s = \emptyset$. If, by contrast, the symbolic policy is implemented by the winning candidate then social identity becomes salient $s \in \{m, M\}$. Both the employer and the employee then observe their respective social identity before deciding whether to enter in a working relationship. At the onset, a worker or employer's social identity is simply her physical identity. However, social identity can change as a result of social interactions.

Our process of identity transmission follows Peski and Szentes (2013). Take employer l working with worker h. Upon break down, employer l's social identity remains unchanged with probability $1 - \rho \in (0, 1)$. With probability, ρ, employer l acquires acquires his employee's identity $s_k \in \{m, M\}$. In turn, worker h's social identity remains unchanged with probability $1 - \gamma \in (0, 1)$ and takes her employer's social identity with probability γ.

This approach allows us to define four different categories of citizens anticipating slightly our discrimination result. First, we have the pure majority with both physical and social majority traits $(\phi, s) = (M, M)$. Second, we have the tainted majority with social minority trait: $(\phi, s) = (M, m)$. These are the majority members referred to as 'jew-lover,' 'nigger-lover,' or 'muslim-lovers,' etc. depending on the circumstances. Third, we can define the assimilated minority as citizens with minority physical trait and majority social identity: $(\phi, s) = (m, M)$. Finally, we refer to the excluded minority when both traits are minority traits: $(\phi, s) = (m, m)$.

A few remarks are in order. First, we assume that physical type M constitutes a majority of the population for both workers and employers. However, we do not impose that the proportion is the same among the two groups (i.e., we can have $\alpha^W \neq \alpha^F$). Second, a citizen's identity changes only if his social identity differs from the social identity of those he interacts with. That is, we assume that social identity is fully determined by social interactions (for tractability reasons, unlike Peski and Szentes, 2013, we do not allow for reversal whereby an agent's social identity can return to her physical one). We are thus considering relatively close net communities where all inhabitants have a long memory of social interaction. Third, we assume that work relationships are a prominent form of social interactions. Finally, for ease of exposition, we do not impose restriction on ρ and γ so full identity swaps are possible.

Equilibrium concept

The equilibrium concept is stationary Subgame Perfect Nash Equilibrium. We add an additional requirement in supposing that the labour market adjusts immediately. That is, workers and employees only consider their (ex-ante) expected payoff from the labour market steady state when making their electoral decision (this allows us to drop the time subscript).

4 Labour market analysis

We first consider outcomes in the labour market when social identity is not salient and so plays no role in hiring decisions. A worker has no incentive to decline a job offer since it only entails a wage loss. Employers then have full power when deciding whether to fill their position. Since social identity is not salient, all employers are identical, and differences in workers arise only due to the specific matches in the labour market. We denote V^o as the continuation value of an employer when he has an opening and $V^f(\theta)$ as the firm's continuation value when the position is filled with a worker characterized by productivity θ.

In a period where the firm position is filled the employer obtains a net profit $(1-\tau)(\theta-w)$. With probability $1-\delta$ this profitable match persists and the employer obtains $V^f(\theta)$ tomorrow (discounted by β). If not (with probability δ), the match breaks and so the employer has an opening tomorrow and obtains V^o. Bringing these elements together we have:

$$\begin{aligned} V^f(\theta) =& (1-\tau)(\theta-w) + (1-\delta)\beta V^f(\theta) + \delta\beta V^o \\ \Leftrightarrow \quad V^f(\theta) =& \frac{(1-\tau)(\theta-w) + \delta\beta V^o}{1-\beta(1-\delta)} \end{aligned} \tag{1}$$

When a firm has an opening it does not produce, its profit is zero in this period, and the employer only receives transfers. It is then matched with an unemployed worker with productivity θ and must decide whether to hire him and so obtain a payoff next period of $V^f(\theta)$. If not, the position remains unfilled and the employer's continuation value is V^o. Thus, when she has an opening, the employer's continuation value is

$$V^o = 0 + \beta E_\theta\big(\max\left\{V^f(\theta), V^o\right\}\big) \tag{2}$$

Of course, when the employer has an opening he does not know with which worker he will be matched. Further since the productivity is match-specific, the relevant distribution is the ex-ante distribution of productivity (i.e., $S(\cdot)$).

Obviously, employers never employ a worker with productivity less than w since it always leads to negative revenue. Less obvious is that frictions in the labour market might lead to an an employer foregoing profit. Specifically, she faces a trade-off between enjoying profit $\theta - w > 0$ today and missing out on a more productive worker ($\theta' > \theta$) tomorrow. If a worker's productivity is too close to w, the loss in term of future opportunity dominates that from the immediate profit that would be lost. Hence, an employer hires a worker if and only if his productivity θ is above a hiring threshold θ^{ND}, strictly greater than the wage w.[8]

Having shown that these value functions are well defined, and building on the arguments presented here, straightforward analysis provides the following result:

[8] We note that even if matches are not persistent, employers only hire workers who generate a strictly positive profit (in formal terms, $\lim_{\delta\to 1}\theta^{ND} > w$). This is due to the timing of labour market interactions. If a hire takes place in period t, production only starts in period $t+1$. Hence, even if matches break after the producing period $t+1$, there is an opportunity cost to hire a low productivity worker (just above w).

Lemma 1. *When social identity is not salient, in the unique equilibrium:*

(i) A worker always accepts to enter in a working relationship;

(ii) There exists a hiring threshold $\theta^{ND} \in (w, 1)$ *such that an employer hires a worker if and only if* $\theta \geq \theta^{ND}$.

We now turn to the labour market equilibria when social identity is salient. When this is so, all actors might simply ignore social identity and play the same strategy as before (when social traits are not salient): An equilibrium without work discrimination always exists. Alternatively, workers and employers might condition their labour decisions on social identity. There are many reasons to focus on this particular assessment. First, Peski and Szentes (2013) and Rosén (1997) both show in distinct settings that a non-discriminatory equilibrium is not stable when identity can serve as a focal point (though it is not clear whether this property holds in our set-up). Second, if social identity has no implication for labour market outcomes, no candidate would ever propose policies that make it salient. Indeed, recall that when indifferent, citizens oppose the implementation of such policy (citizens' electoral choice can be understood as a form of forward-looking behavior). Thus to understand why they are proposed, we look at the situation when the labour market is affected by policies that raise the salience of social identity.

What does a labour market shaped by social identity look like? Salient social identity generates a fully segregated labour market. Workers with majority social trait accept job offers from employers with the same identity. Likewise, employers with majority social trait only hire workers from their own majority group. Workers with majority social identity all face the same payoff from accepting a job offer. Hence, if one of them is willing to work for a firm owned by an employer with minority social trait, all would be willing to do so: the labour market would look identical to that under the non-discrimination equilibrium. Similarly, if majority firms hire minority workers while majority workers refuse to work with minority firms, then with probability a.e. one the proportion of either minority workers or majority employers would drop to zero over time. Hence, in the steady state, identity could no longer play a role. Given its consequences on hiring, we refer to equilibria in which identity serves the role of focal point in the labour market subgame as 'discrimination equilibria.'

How do employers behave in the labour market subgame in a discrimination equilibrium? As in the non-discrimination equilibrium, firms hire workers if and only if their productivity is above a certain threshold. However, employers with distinct identities use different hiring thresholds. When matched with a worker who shares the same social trait, an employer faces the familiar trade-off between profit today and the opportunity cost of missing out on a better match tomorrow. The opportunity cost, however, now depends on the proportion of workers with the same identity among the unemployed.

To see the effect of discrimination on labor market outcome,denote $V^f_{J,K}(\theta)$ the continuation value of an employer with physical trait $J \in \{M, m\}$ and social trait $K \in \{M, m\}$ when the position is filled (f) by an employee with productivity θ and similar social identity. Using this notation, and following the steps in our earlier construction (of value functions for a world absent discrimination), we obtain the following continuation value function:

$$V^f_{J,K}(\theta) = (1-\tau)(\theta - w) + (1-\delta)\beta V^f_{J,K}(\theta) + \delta\beta V^o_{J,K}. \tag{3}$$

Note that since we focus on the continuation value when the employee has similar social trait, the employer never changes social identity upon separation. This expression includes $V^o_{J,K}$ which denotes the continuation value should the employer's position becomes open (an event with probability δ). In a discrimination equilibrium, the continuation value depends upon the proportion of workers who share the employers identity. Correspondingly, we write the probability that an employer is matched with a majority unemployed worker as μ^M. Thus the relevant value functions for an employer with (physical and social) traits $J \in \{M, m\}$ and with an open position are given by:

$$V^o_{M,M} = 0 + \mu^M \beta E_\theta \max\{V^f_{M,M}(\theta), V^o_{M,M}\} + \beta(1-\mu^M)V^o_{M,M} \tag{4}$$

$$V^o_{m,m} = 0 + (1-\mu^M)\beta E_\theta \max\{V^f_{m,m}(\theta), V^o_{m,m}\} + \beta\mu^M V^o_{m,m} \tag{5}$$

Since a majority-trait employer only hires majority workers, his status can change (from open to filled position) only if he meets a majority worker (i.e., with probability μ^M). In turn, a minority-trait employer's position can become filled only if he is matched with a minority worker (i.e., with probability $1-\mu^M$).

As before, an employer uses a cutoff strategy: hires if and only if $\theta \geq \overline{\theta}^D_M$ ($\theta \geq \overline{\theta}^D_m$) for a majority (minority) employer. Since the pool of potential matches is not evenly balanced except in knife-edge cases, employers with different identities almost always use different thresholds. Because employers only hire workers with the same identity as their own, the pool of potential employees is always smaller with than without discrimination. As a consequence, both types of employer become more lenient in their hiring and so their hiring thresholds are strictly lower than when social identity is not salient. The proportion of unemployed with majority or minority traits, in turn, depends on the hiring thresholds adopted by firms. In equilibrium, unemployment rates for both types of workers and the hiring thresholds for both types of firms must be consistent with each other. Despite this complicated fixed point problem, we show that the equilibrium is unique. Lemma 2 details the key features of labour market strategies in this unique discrimination equilibrium.

Lemma 2. *When social identity is salient, in any discrimination equilibrium, labour market strategies satisfy:*

(i) A worker with a majority social identity only agrees to work with an employer with a majority social identity;

(ii) An employer with a majority social identity never hires a worker with a minority social identity;

(iii) There exists a unique pair of hiring thresholds $\theta_M^D(\alpha^W, \alpha^F), \theta_m^D(\alpha^W, \alpha^F) \in (w, \theta^{ND})^2$ *such that*

- *an employer with majority social identity hires a worker with majority social identity if and only if* $\theta \geq \theta_M^D(\alpha^W, \alpha^F)$*;*
- *an employer with minority social identity hires a worker with minority social identity if and only if* $\theta \geq \theta_m^D(\alpha^W, \alpha^F)$*.*

The hiring thresholds are a function of all parameters. For our purpose, their dependence on the proportions of workers with majority physical (and thus social) identity, α^W, and of employers with majority physical trait, α^F, are of key importance. Corollary 1 shows that the comparative statics with respect to these two parameters conform with straightforward intuition.

As the proportion of workers with majority identity increases, employers with majority traits are more likely to be matched with workers of the same type. As a result, they become more strict in their hiring decision: the hiring threshold strictly increases with α^W. In contrast, employers with minority identity are less likely to be matched with a worker from with minority social trait and so the hiring threshold of these firms decreases with α^W. In the limit, if all workers come from the majority, type-M employers are certain to be matched with workers of the same type and thus use the same hiring threshold as absent discrimination (θ^{ND}). Employers with minority trait are very unlikely to find anyone who accepts their job offer and so hire any profitable worker with identity m.

An increase in the proportion of employers with majority identity has the reverse effect on the hiring thresholds. A greater α^F implies that more workers with majority identity are already employed so the unemployed pool contains relatively more type-m workers. For employers with majority traits, the chances of obtaining a better match in the future decreases and they become more lenient in their employment decision. For employers with minority trait, the chances increase and they become more strict in their hiring decision.

Corollary 1. *In the unique discrimination equilibrium, the hiring thresholds $\theta_M^D(\alpha^W, \alpha^F)$ and $\theta_m^D(\alpha^W, \alpha^F)$ satisfy the following properties:*
(i) the hiring threshold of employers with majority social identity ($\theta_M^D(\alpha^W, \alpha^F)$) is strictly increasing with the proportion of workers with majority identity (α^W) and strictly decreasing with the proportion of employers with majority identity (α^F);
(ii) the hiring threshold of employers with minority social identity ($\theta_m^D(\alpha^W, \alpha^F)$) is strictly decreasing with the proportion of workers with majority identity (α^W) and strictly increasing with the proportion of employers with majority identity (α^F);
(iii) $\lim_{\alpha^W \to 1} \theta_M^D(\alpha^W, \alpha^F) = \theta^{ND}$ *and* $\lim_{\alpha^W \to 1} \theta_m^D(\alpha^W, \alpha^F) = w$.

Having characterized the labor market participants' strategies in the discrimination equilibrium in Lemma 2, henceforth the 'discrimination strategy,' we now turn to describing the conditions such that these strategies are mutual best responses. For this to be so, two conditions need to be satisfied: 1) a worker with majority trait refuses any job offer from an employer with minority identity and 2) an employer with majority identity always refuses to hire a worker with minority trait.

The first condition holds if a worker's present benefit of being employed by a type-m employer is higher than the future loss from being ostracized by employers with the majority trait (if the worker changes social identity). Being ostracised by majority employers is not always disadvantageous to the type-M worker. It depends critically on the proportion of these workers relative to the proportion of type-M employers. When there are few majority firms and many majority workers, (with discrimination) majority workers are likely to suffer unemployment. Then, paradoxically, workers with minority social identity experience relatively favorable labor market prospects.

It follows that workers with majority trait suffer a loss from being tainted by the minority only if their proportion relative to the percentage of type-M employers is not too high: formally, α^W is below a population threshold $\overline{\alpha}^W(\alpha^F)$; when not, a worker with majority trait gains from switching social identities and so refusing a job offer from a type-m employers is never an equilibrium strategy. In contrast, when $\alpha^W < \overline{\alpha}^W(\alpha^F)$, so that workers with majority identity have better employment prospects, the anticipated cost of being ostracised always dominates provided workers are sufficiently patient and the majority worker, then, turns down the opportunity to work for a type-m employer.

The second condition mirrors the first with respect to the employment decisions of type-M employers. It requires that the present benefit of hiring a type-M worker with maximum productivity ($\theta = 1$) is strictly lower than the loss induced from being ostracized by type-M workers in the future (if the employer changes social identity). Again, an employer does not always face a cost from belonging to the social minority. Indeed, if the proportion of employers with majority trait is very high compared to the percentage of workers from the same group, type-M firms are unlikely to be matched with workers with the same traits are most are already employed. Since most unemployed workers belong to the minority group, hiring is relatively easier for employers with minority social identity.

Thus employers with majority trait suffer a loss from being tainted by the minority only if their proportion relative to the percentage of type-M workers is not too high: formally, α^W is above a population threshold $\underline{\alpha}^W(\alpha^F)$. When this condition is not satisfied, an employer with majority trait gains from switching identity and refusing to hire a sufficiently productive type-m workers is never an equilibrium strategy. By contrast, when $\alpha^W > \underline{\alpha}^W(\alpha^F)$, so firms with majority traits have better prospects of filling their position, the future loss always dominates the present benefit of hiring a very productive type-m workers provided employers are sufficiently patient.

Bringing this reasoning together, we obtain a necessary condition for the existence of the discrimination equilibrium: the proportion of workers with majority identity relative to the percentage of employers with the same trait is intermediary (not too high and not too low).[9]

Proposition 1 summarizes our findings and Figure 1 illustrates them.

Proposition 1. *For all proportions of employers with majority identity $\alpha^F \in (1/2, 1)$, there exists two unique population thresholds $\underline{\alpha}^W(\alpha^F), \overline{\alpha}^W(\alpha^F) \in (1/2, \alpha^F) \times (\alpha^F, 1]$ such that if the proportion of workers with majority identity satisfies $\alpha^W \in (\underline{\alpha}^W(\alpha^F), \overline{\alpha}^W(\alpha^F))$, there exists $\overline{\beta} < 1$ such that for all $\beta \geq \overline{\beta}$, workers' and employers' discrimination strategies are mutual best response.*

The two population thresholds (as well as the lower bound on the discount factor $\overline{\beta}$) depend on all model parameters. In particular, they depend on the percentage of employers with majority trait α^F. As shown in Figure 1, both $\underline{\alpha}^W(\cdot)$ and $\overline{\alpha}^W(\cdot)$ increase with the proportion of employers exhibiting majority traits. The logic is similar to previous results. Higher α^F implies that unemployed workers with identity M are more likely to be matched with the right type of firms. Hence, these workers

[9] The lower and upper bounds are not mutually exclusive. Indeed, the population thresholds satisfy $\underline{\alpha}^W(\alpha^F) < \alpha^F < \overline{\alpha}^W(\alpha^F)$. Given the greater proportion of type-M employers and workers ($\alpha^F > 1/2$ and $\alpha^W > 1/2$), workers and employers with majority identity have greater chances of being matched together than their minority counterparts when $\alpha^W = \alpha^F$. A type-M workers then refuse a type-m employer's job offer and a type-M employers never hire a type-m workers as long as they are sufficiently patient.

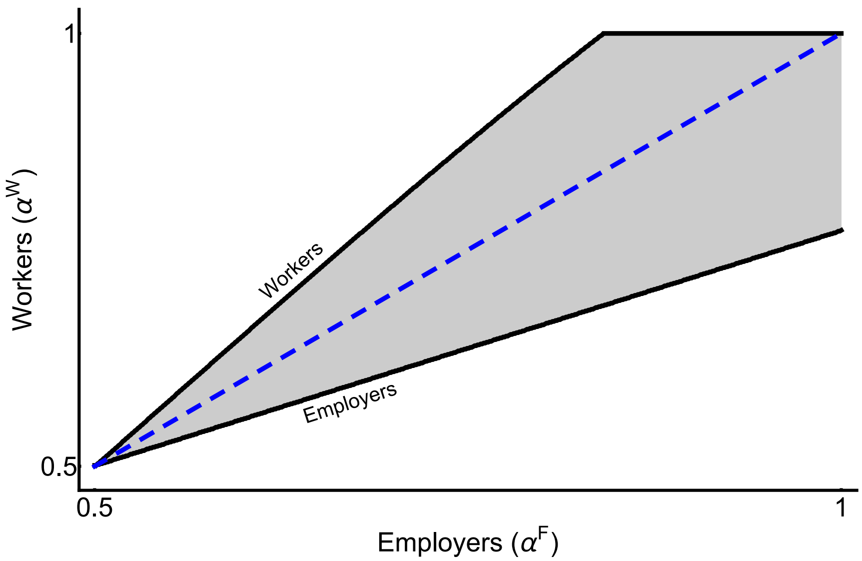

Figure 1: Labour market discrimination

The black plain curves labeled 'Employers' and 'Workers' depict $\underline{\alpha}^W(\alpha^F)$ and $\overline{\alpha}^W(\alpha^F)$, respectively. The dashed blue line is α^F. The shaded gray area correspond to proportions of type-M workers and employers such that discrimination strategies are mutual best responses (for appropriate β). Parameter values: $w = 0.3$, $\delta = 0.2$.

have greater incentives to refuse a job offer from an employer with minority trait so the upper bound $\overline{\alpha}^W(\alpha^F)$ increases. Higher α^F also implies that more type-M employers compete for the same pool of workers making it more attractive to hire a worker with minority trait so the lower bound $\underline{\alpha}^W(\alpha^F)$ increases as well. If employers with majority trait barely constitute a majority, type-M workers have very little incentive to reject a type-m employer's offer, and type-M employers little incentive not to hire a (sufficiently productive) type-m workers. Hence, both thresholds converge to 1/2 as firms become evenly split between majority and minority.

Corollary 2. *The thresholds $\underline{\alpha}^W(\alpha^F)$ and $\overline{\alpha}^W(\alpha^F)$ satisfy the following properties:*
(i) $\underline{\alpha}^W(\alpha^F)$ and $\overline{\alpha}^W(\alpha^F)$ are strictly increasing with α^F;
(ii) $\lim_{\alpha^F \to 1/2} \underline{\alpha}^W(\alpha^F) = 1/2 = \lim_{\alpha^F \to 1/2} \overline{\alpha}^W(\alpha^F)$;
(iii) There exists a unique $\widehat{\alpha}^F < 1$ such that $\overline{\alpha}^W(\alpha^F) = 1$ for all $\alpha^F \geq \widehat{\alpha}^F$;
(iv) $\lim_{a^F \to 1} \underline{\alpha}^W(\alpha^F) < 1$.

The discrimination strategies we describe in Lemma 2 generates some interesting empirical patterns when they constitute an equilibrium. First, workers with minority traits suffer from higher unemployment rates and longer spells of unemployment than their type-M counterparts. In turn, because in equilibrium employers with majority traits must find it easier to hire than their minority counterparts, type-M employers are more discerning in their hiring: their hiring threshold is strictly higher than that of type-m employers'. As a result, firms owned by employers with majority identity are more productive, on average.

Proposition 2 summarizes the observable labour market characteristics of the discrimination equilibrium, when it exists.

Proposition 2. *In the unique discrimination equilibrium,*
(i) the unemployment rate and duration of unemployment spell are strictly higher for workers with minority trait than workers with majority trait;
(ii) a firm's average productivity when producing is strictly higher when it is owned by an employer with majority trait than when it is owned by an employer with minority trait.

Although these outcomes do not stem from statistical discrimination, they are consistent with empirical observations that could inform reasoning that is akin to statistical discrimination: the higher output of majority owned firms is due to their employment of majority workers with high levels of productivity and the low productivity of minority owned firms is related to lower productivity amongst their minority employees. However, here, the observed correlation between identity, attributes and outcomes is entirely spurious. Differential firm outputs are not related to differences in productivity/human capital investments between groups in the general population. Instead they are generated by different endogenous employment thresholds deployed by majority and minority owned firms in equilibrium.

The results described above are robust to a change in the labor market. In Appendix A, we show that the necessary conditions that we characterize in Proposition 1 (and illustrate in Figure 1) remain unchanged in a labor market where employers with majority trait can hire for free a minority worker or where workers with majority trait can be offered a minority-owned firm's full profit when matched (see Proposition A.1 for details). The intuition is quite simple. The thresholds on α^W guarantee that the future cost for type-M employers and workers from being ostracized is strictly positive. If actors expect others to play a discrimination strategy, these costs remain the same for the majority no matter what the minority does since the labor market is fully segregated. Thus, we can always find a sufficiently high discount factor such that these future costs dominate any present benefit and discrimination strategies remain mutual best responses.

We conclude this section by highlighting our main finding thus far. Even when all types of workers and all types of employers are identical ex-ante, raising the salience of social identity sets in motion forces that lead to a fully segregated labour market. This, in turn, can lead to workers from different groups in society experiencing very different economic circumstances according to their shared identity. And it can lead to different patterns of production according to the identity of the firm owner.

That social identity can yield segregation has a long history. For example, in the Middle Ages, the Cagots suffered from total segregation. As Appiah (2018, p.28) describes, "the Cagots [of French and Spanish Pyrenées] were for a millennium treated as pariahs, relegated to disfavoured districts, even forced to use separate doors in churches, where they received the Communion water at the end of a stick." While it is hard to determine the cause of this discrimination, the fear of identity transmission seemed to have played a key role. Appiah continues (emphasis added) "*[b]ecause contact with the Cagots was contaminating*, they were severely punished for drinking from the same water basin as others, for farming, or even walking barefoot on the streets." More recently, in Austria, Muslims seem to have been fully separated from the rest of the population after the entry of the far-right party FPÖ into government and the implementation of several symbolic laws targeting this minority. As Professor Mahmud Yavuz describes, "Young people who want to help this country proper are cast aside. The gap is growing. Nobody wants to have contacts with us. Before neighbors used to come to open days at the mosque, not any more" (from news article "En Autriche, il ne fait pas bon tre musulman," Le Monde, May 17th 2019, authors' translation).

Recent studies also highlight how exogenous events or policies that arguably raise the salience of identity traits can have far-reaching consequences. Zussman (2013) documents that increased tension between Palestinians and Israeli breeds intolerance in the interactions with members of the out-group in the private used car market. Glover (2019) shows how the Charlie Hebdo attack in France in January 2015 significantly lowered employers' demand for Muslim workers and Muslim workers' job searching effort. Quite consistent with our theory, the effect is driven by communities exhibiting a lower ex-ante bias against minorities (as proxied by the far-right Front National's vote share). Abdelgadir and Fouka (2019) looks at the impact of the headscarf ban in 2004 on the achievements of minority pupils. Muslim girls see a decrease in their secondary educational attainment (lower completion rate, higher drop-out rates) and experience greater racism following the law.

While not about labor market outcomes, Abdelgadir and Fouka's findings stress that a symbolic policy like the headscarf ban can have adverse consequences for the minority as a whole, not just the small proportion of Muslim pupils directly targeted by the measure. They also show that the salience of social identity is not exogenous; it requires politicians to create an environment where identity matters.

And for this to be so requires that it is in politician's interests to do so. To determine when this condition is satisfied, we analyse the electoral competition stage of our framework.

5 Electoral competition

We have seen that when identity is salient to people's choices then labor discrimination that disadvantages minorities can occur in equilibrium. For policies that make identity salient to be introduced it must be that at least part of the citizenry favours outcomes associated with a segregated labour market. That is, some citizens must obtain a strictly higher ex-ante expected welfare (the adequate criterion since the electoral competition precedes the labour market) due to discrimination. The section of the population that stands to gain the most from labour market discrimination are those workers with majority traits. When will they demand the introduction of policy that facilitates discriminatory labour market outcomes in equilibrium?

Note that discrimination is not without costs even for majority workers. With discrimination they (the type-M workers) must refuse to work with type-m employers (Lemma 2). Without discrimination, they can enter working relationships with all types of employers. There are also fiscal consequences, since labour market discrimination induces lower transfers. Recall, from the previous section, that firms are less selective when it comes to hiring with discrimination (i.e., $\theta^{ND} > \theta^D_M > \theta^D_m$) and so firms with filled positions are less productive. Firms are also less likely to produce (indeed, it is because of the reduced opportunity to produce that firms are less demanding when hiring). These joint effects (lower productivity and lower proportion of firms producing) imply that the total amount available for taxation is lower.

Two conditions are thus necessary for workers to support the introduction of symbolic policies that alter the dynamics of the labor market. The loss from restricted employment opportunities (due to being less likely to be matched to an appropriate employer) must be strictly lower than the gain from the increased likelihood of being hired by an employer with majority identity. This condition is satisfied when the proportion of type-M firms α^F is large compared to the proportion of workers with majority trait α^W—i.e., α^W must be below a threshold denoted $\widehat{\alpha}^W(\alpha^F)$. The loss in term of tax transfers must also not be too large. This is so when transfers are not too sensitive to change in tax revenues from labour income. That is, the function $K(\cdot)$ does not increase too fast. Since $K(\cdot)$ is concave, this is equivalent to $K'(0)$ not being too large: an interpretation is that large inefficiency in redistribution and/or the existence of other tax revenues diminishes the importance of labour taxation (recall that we do not impose $K(0) = 0$).

Assuming (for now) that discrimination strategies are mutual best responses, we thus obtain the following necessary conditions for type-M workers to demand the introduction of symbolic policies such as the Burkini ban.

Proposition 3. *For all proportion of employers with majority identity $\alpha^F \in (1/2, 1)$, there exists a unique population threshold $\widehat{\alpha}^W(\alpha^F) \in [1/2, \alpha^F)$ such that if the proportion of workers with majority identity satisfies $\alpha^W \in (1/2, \widehat{\alpha}^W(\alpha^F))$ (possibly an empty interval), there exists an upper bound $\overline{K} > 0$ such that workers with majority identity demand the symbolic policy whenever the marginal effect of labour income taxation on transfers satisfies $K'(0) < \overline{K}$.*

One property of the population threshold $\widehat{\alpha}^W(\alpha^F)$ is worth highlighting. The demand for social discrimination is positive only if the proportion of workers with majority identity is *strictly lower* than the proportion of employers with the same identity (i.e., $\widehat{\alpha}^W(\alpha^F) < \alpha^F$). Hence, social discrimination is possible only if the minority is poorly integrated economically in this sense.

The intuition for this condition again involves the loss due to reduced employment opportunities: with discrimination, workers with majority traits no longer compete with type-m workers for positions offered by employers with majority identity but they also abandon the opportunity to work with type-m employers. If the proportion of types is the same among workers and employers, the ratio of jobs to workers is the same (one-to-one) with or without work discrimination. Type-M workers are, however, less likely to be matched with a firm that will employ them (since they exclude employers with minority identity). Employers' adjusted hiring practices (hiring threshold θ_M^D instead of θ^{ND}) are not sufficient to compensate for such lower employment prospects. Although firms take into account the reduced likelihood of being matched with type-M workers, they also consider the possibility of meeting a more productive adequate employee tomorrow. A necessary condition for workers with majority trait to benefit from discrimination is, thus, that the jobs to workers ratio becomes favorable ($\alpha^W < \alpha^F$) *and* employers with majority identity become sufficiently lenient in their hiring decisions (recall from Corollary 1 that θ_M^D decreases with α^F).

The population threshold $\widehat{\alpha}^W(\alpha^F)$ exhibits other important properties. Intuitively, the upper bound on the demand for social discrimination is less stringent as the proportion of type-M employers increases. Since then workers with majority identity are more likely to be matched with the right type of employer, they benefit more from the labour market opportunities with discrimination. In the limit when all employers exhibit majority traits, type-M worker always gain income from work discrimination since they face reduced competition for open positions. We thus obtain:

Corollary 3. *The population threshold $\widehat{\alpha}^W(\alpha^F)$ satisfies the following properties:*
(i) $\widehat{\alpha}^W(\alpha^F)$ is increasing with α^F, strictly if $\widehat{\alpha}^W(\alpha^F) > 1/2$;
(ii) $\lim_{\alpha^F \to 1} \widehat{\alpha}^W(\alpha^F) = 1$.

We have established conditions such that a majority of workers benefit from workplace discrimination and so will demand policies that facilitate such discrimination. Next we turn to the supply side and consider candidates' behavior. It is useful for this purpose to define the following two quantities: Let τ^{ND} be the preferred tax rate of workers in the absence of work discrimination; and, in turn, let τ_J^D be the preferred tax rate of workers with identity $J \in \{M, m\}$ with discrimination. Notice that employers do not receive any transfer so their preferred tax rate is always 0.

Observe that candidates' problem is two-dimensional: they must choose a tax rate and a position on the symbolic dimension. While multi-dimensional electoral competition models are often intractable, we can take advantage of the binary nature of the symbolic dimension (offer or not the symbolic policy) and the shared interest of each separate group of voters (majority workers, minority workers, majority citizen employers, minority citizen firm owners) to establish the following properties of equilibrium platforms. In equilibrium, candidates propose one of two platforms. Candidates either do not offer the symbolic policy and propose a tax rate of τ^{ND} ($(0, \tau^{ND})$) or they promise the symbolic policy with a tax rate of τ_M^D ($(1, \tau_M^D)$).

Lemma 3. *In equilibrium, candidates converge to the same platform. Further, for all $J \in \{A, B\}$, candidate J's platform satisfies either $(d_J, \tau_J) = (0, \tau^{ND})$ or $(d_J, \tau_J) = (1, \tau_M^D)$.*

The classic dynamics of spatial politics explains this result. Any candidate proposing a platform $(0, \tau)$, with $\tau \neq \tau^{ND}$, faces certain defeat. Indeed, if his opponent offers $(0, \tau^{ND})$, he gathers the votes of all workers, a plurality of the citizenry since $1 > f$, and wins the election. Hence, by the usual logic, any candidate who is elected proposing no symbolic policy must also be proposing workers' preferred tax rate. Similarly, any candidate $J \in \{A, B\}$ proposing a platform $(1, \tau)$ with $\tau \neq \tau_M^D$, is certain to lose. If $\tau > \tau_M^D$, his opponent by offering $(1, \tau_M^D)$ forms a winning electoral coalition consisting of workers with majority traits and all employers. If $\tau < \tau_M^D$, his opponent by offering $(1, \tau_M^D)$ forms a winning electoral coalition consisting of all workers. Indeed, since with social identity salient, workers with minority identity experience worse labour market outcomes than their majority counterparts (Proposition 2), they favour a higher tax rate than τ_M^D and so would vote for $(1, \tau_M^D)$ over any platform $(1, \tau)$ with $\tau < \tau_M^D$. Hence, by the usual logic, any candidate who is elected proposing the symbolic policy must also be offering the tax rate preferred by type-M workers.

We are now ready to determine when the existence of the discrimination equilibrium, or equivalently conditions under which both candidates propose the symbolic policy. Two conditions need to be satisfied: 1) discrimination strategies are mutual best responses and 2) a plurality of the citizenry demands social discrimination; that is, prefers $(1, \tau_M^D)$ to $(0, \tau^{ND})$ by Lemma 3. To satisfy jointly these two conditions, it is necessary that the proportion of type-M workers relative to the proportion of type-M employers be intermediary: $\alpha^W \in (\underline{\alpha}^W(\alpha^F), \widehat{\alpha}^W(\alpha^F))$ (see Propositions 1 and 3)). In particular, note that the upper bound is determined by the political demand. Indeed, for discrimination to be a best response, type-M workers compare their labour situation as an advantaged majority compared to a discriminated minority. When considering the demand for discrimination, type-M workers compare their ex-ante welfare (including lower transfer) with and without discrimination. The second intuitively is a more stringent condition.

A positive demand for social discrimination is not sufficient, demand must be sufficiently large so that candidates offer the symbolic policy. If type-M workers are a sufficient mass ($\alpha^W > (1+f)/2$), this is guaranteed. If not, it must be that employers with majority identity side with workers with the same trait. Majority firm owners indeed face a trade-off between accepting lower productivity, but also lower taxation with discrimination. While we cannot exclude the possibility that employers with majority identity favors the symbolic policy, we cannot prove it either. The reason is that the hiring thresholds described in Lemmas 1 and 2 are only implicitly defined making it hard to compare employers' expected payoffs across equilibria.

The properties of the population thresholds $\underline{\alpha}^W(\alpha^F)$ (Corollary 1) and $\widehat{\alpha}^W(\alpha^F)$ (Corollary 3) guarantee that all necessary conditions can be satisfied simultaneously whenever the proportion of type-M employers (α^F) is sufficiently large. We thus obtain the following proposition which summarizes (sufficient) conditions on the population and other parameters for the symbolic policy to be implemented and social discrimination to be observed in equilibrium.

Proposition 4. *The equilibrium is the discrimination equilibrium with both candidates A and B offering the symbolic policy $d = 1$ and a tax rate of $\tau = \tau_M^D$ when:*

(a) the proportion of employers with majority identity satisfies $\alpha^F \in (\underline{\alpha}^F, 1)$ for some threshold $\underline{\alpha}^F \in (1/2, 1)$;

(b) the proportion of workers with majority identity satisfies $\alpha^W \in (\underline{\alpha}^W(\alpha^F), \widehat{\alpha}^W(\alpha^F))$*;*

(c) Workers and employers are sufficiently patient: $\beta > \overline{\beta}$*;*

(d) The marginal effect of labour income taxation on transfers is sufficient low: $K'(0) < \overline{K}$*.*

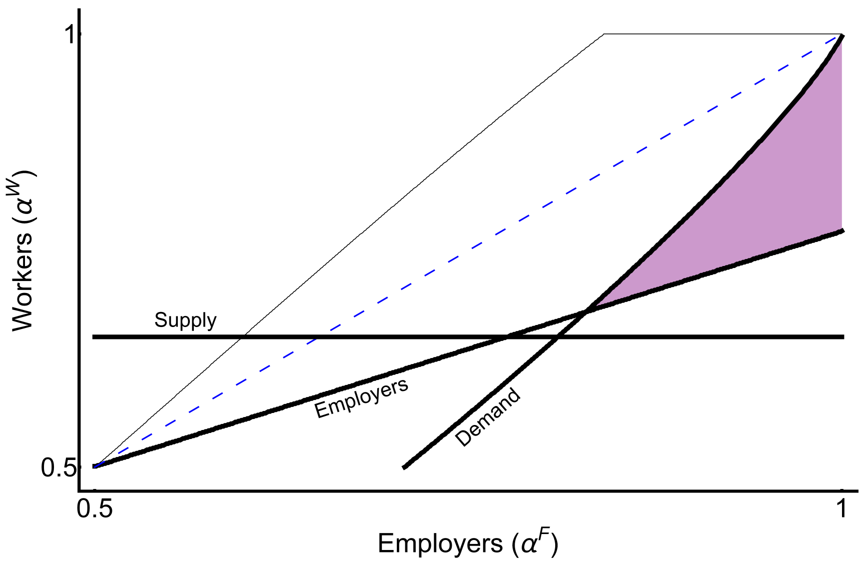

Figure 2: Labour market discrimination

The black plain curve labeled 'Employers' depicts $\underline{\alpha}^W(\alpha^F)$. The curve labeled 'Demand' depicts $\widehat{\alpha}^W(\alpha^F)$. The horizontal line labeled 'Supply' depicts $\frac{1+f}{2}$. The shaded purple area correspond to proportions of type-M workers and employers such that discrimination is an equilibrium outcome (for appropriate β and $K'(0)$). Parameter values: $w = 0.3$, $\delta = 0.2$, $f = 0.3$.

Figure 2 illustrates the conditions detailed in Proposition 4. In equilibrium, symbolic policies are introduced only if the proportion of minority workers is sizeable (so there is a benefit from labour discrimination), but not too important (so it cannot block discrimination through democratic means). Further, it is necessary that the minority is poorly integrated economically (i.e., $\alpha^F > \underline{\alpha}^F$ and $\alpha^W < \alpha^F$). While we cannot directly test whether these conditions are satisfied, a look at demographic statistics suggests that they are not completely unrealistic. The Muslim population in Europe is estimated to amount to 8.8% of the Austrian population (Austrian census), 6% of Belgians, 5% in Germany, 4.7% in Great Britain (Dancygier, 2017, 12 footnote 33). There are few statistics available on firms owned by employers with minority identity. Nevertheless, there seems to be agreement that many minorities are under-represented in several European countries. For example, self-employment is lower for Turks than for natives in Germany European Commission (2008). In Great Britain, the reverse holds true with a greater proportion of Pakistani and Indian minorities than White British choosing self-employment Ram and Jones (2008). This does not falsify our findings—far from it–since Great Britain is an exception in that the U.K. has not (yet) experienced policies of the form of burqa or hijab bans.

Our last set of results details the consequences of symbolic policies. To do so, we compare equilibrium outcomes when identity is salient and social discrimination occurs to those when it is not. This comparison is adequate for two complementary reasons. First, countries are in either one regime or the other and so such a contrast corresponds to testable empirical predictions. Second, when identity is not salient, observable quantities arising in the unique equilibrium do not depend on the share of workers and employers with majority traits they only depend on the distribution of skills which is kept constant throughout). As a consequence, equilibrium outcomes when identity is not salient provide the correct counterfactual: they correspond to what would have happened were the conditions in Proposition 4 met but symbolic policies not implemented.

The next proposition establishes the large economic consequences of symbolic policies such as those mentioned in our introductory remarks. First, as a result of the discrimination that follows their introduction overall production is lower as firms are less likely to be matched with workers they can hire. Second, the consequent discrimination increases unemployment rate of minorities who are now excluded from most jobs. Third, even though they are orthogonal to symbolic policy, redistributive transfers are strictly lower with than without discrimination. This follows from the the lower production which leads to less income available from taxation, but not only. Workers with majority traits, unlike their minority counterparts, experience better labour market prospects under discrimination. Thus, taxation is relatively more costly for them. This implies that the change in tax rates τ_M^D does not fully compensate the lower tax base. Higher labour market income can even induce workers with majority identity to demand lower taxation when symbolic policies are introduced than they would in their absence (e.g., if transfers are not too sensitive to labour income taxation at the margin).

Proposition 5. *Compared to the unique equilibrium when symbolic policies are not introduced and there is no discrimination, in the discrimination equilibrium*

(i) Total production is strictly lower;

(ii) The employment rate of workers with majority social identity is strictly higher,

(ii') The employment rate of workers with minority social identity is strictly lower;

(iii) Redistributive transfers are weakly lower, strictly if $\tau^{ND} > 0$;

(iv) Further, if the elasticity of the marginal effect of labour income taxation on transfers is less than one in absolute values, then the tax rate is also weakly lower, strictly if $\tau^{ND} > 0$.

Proposition 5 highlights that workers with majority traits are generally the only beneficiaries of discrimination (though as mentioned above, we cannot exclude that type-M employers may also gain from it). Social discrimination is thus best understood as a transfer from the rest of the population to the native working class majority. Workers with minority identity are, in turn, severely negatively affected. They see their employment prospects reduced and their non-labour income diminished. That is, even though transfers are uniformly redistributed to the working population, workers with minority traits get less assistance when they need it the most.

This negative relationship between symbolic policies raising the salience of identity traits and redistribution has received some anecdotal and empirical support. Anecdotally, during the French presidential campaign, Marine Le Pen played both the identity card (France for French first) and promised a 10% tax cut. Empirically, Alt and Iversen (2017) show that labor market segmentation along ethnic lines is a strong predictor of lower support for redistribution.

6 Summary and discussion

Discussing the causes of ethnic conflict, Fearon and Laitin (2000, 857) remark that "political elites use violence to construct antagonistic ethnic identities, which in turn favour more violence" with the masses following. Our paper shows that this phenomenon is not limited to weakly institutionalized environments. In democracies, politicians use social discrimination to construct antagonistic social identities, which in turn favour labor market discrimination. They do so when such policies are supported by the native working class (who form a majority) who can benefit from these labour effects. The labor market discrimination we document is not due to overt prejudice. In our theoretical framework, there is no room for taste-based discrimination. The labor market discrimination we describe is not caused by difference in human capital of workers. All workers are ex-ante identical so there is no room for statistical discrimination. Rather, labor market discrimination is sustained by anticipations about what others would do if a majority worker/employer interacts with a minority employer/worker. It is rooted in expectations about others' behaviors rather than preferences or beliefs (for experimental evidence on the importance of expectations, see Daskalova, 2018). As such, the discrimination our paper analyses is best understood as social discrimination.

This form of discrimination has dramatic consequences. In equilibrium, the labor market is segregated. Large differences emerge in the employment patterns of minority and majority workers and in the productivity of majority and minority owned firms. These labour market effects are magnified by fiscal policies that further erode minority welfare: when they need it most, minority workers unable to find employment in the segregated market receive lower redistributive transfers even though revenues from taxation are uniformly redistributed. The economic impact of policies that evoke discrimination have negative consequences for the economy as a whole. The segregated labour market reduces the likelihood that firms find adequate workers and overall production. While majority workers are beneficiaries of policy that invoke discrimination, via better employment prospects, this labour gain translates into lower demand for redistribution that further hurts minorities. In short, the annoyance that minorities might face upon the introduction of policies such as a burka or beef ban is nothing compared to the broader impact of such policies that simultaneously erode their career prospects and their safety net.

Our predictions regarding when social discrimination arises have consequences for evaluating natives' reactions to minorities or to immigration. Rather than looking at the size of the minority to determine whether the majority becomes more tolerant via contact or radicalises due to threat, our paper suggests that it is critical to take into account the composition of the minority. When the minority is mostly composed of workers, then discrimination becomes more likely; when it is made of employers, then the risk of discrimination recedes (for qualitative evidence consistent with this prediction, see Dancygier, 2010). Further, even the concept of threat may need revision. Natives or majority members may favour policies targeting the minority because of the gain from future job prospects associated with a segregated labor market.

Social discrimination does not just have redistributive consequences, it also reduces the size of the whole pie. This has two important implications. First, it suggests that economic arguments, such as the detrimental impact on growth of some discriminatory policies, are unlikely to be persuasive (e.g., the failure of the so-called 'Project Fear' in the Brexit referendum, an event which, arguably, changed the perception of Eastern European migrants, e.g. Rzepnikowska, 2019). Our paper shows that the majority of voters (workers with majority traits) may find the overall adverse economic consequences unimportant as long as the policy benefits them personally. Second, the shrinking of the pie suggests a normative rationale to restrict the possibility of implementing policies raising the salience of identity for discriminatory purposes.

Our paper provides an analysis of the social discrimination that arises due symbolic policies. What then are the policies that can counteract social discrimination? A clear (though not simple) policy recommendation to limit the attractiveness of symbolic policies or any policy triggering labor market segregation via social discrimination. As noted in Proposition 3, there is a demand for social discrimination only if the minority is poorly integrated economically. Favouring entrepreneurship by migrants and people from ethnic minorities does not just represent an economic opportunity European Commission (2008). It is also a defence against policies that have the potential to hurt the economy. Another solution would be to impose quotas *within firms* for minority workers. This form of affirmative action in the workplace may be beneficial not so much because it improves the employment prospect of minority workers, but because it reduces the risk that the labor market is segregated. As such, our theoretical framework suggests a novel argument for policies imposing some diversity in the workplace. A probably even less popular policy to fight social discrimination would be to open the country to immigration and, especially, naturalisation. Increasing the share of minority workers makes it harder to sustain discrimination strategies as mutual best responses (see Proposition 1 and Figure 1) and decreases the electoral incentives to propose symbolic policies.

More generally, our paper offers a word of caution for policy-makers seeking to reduce the possibility of discrimination against minorities. The effectiveness of a policy much depends on the type of discrimination politicians wish to target. Besides well-known forms of discrimination–such as taste-based or statistical discrimination–our paper highlights an additional one, namely social discrimination. Policies effective in fighting one type of discrimination may reinforce another. Take the diversity-compliant label to firms with diverse workforce proposed by Adida, Laitin, and Valfort (2016, 157). This measure may be effective in combatting statistical discrimination, while having perverse consequences when it comes to social discrimination (raising the salience of identity or increasing the probability of identity transmission). Similarly, workplace affirmative action may help limit social and statistical discrimination, but might raise resentment and thus worsens discrimination that is fundamentally taste-based.

The last observation suggests that identifying the source of discrimination is of primary importance for policy purposes.[10] For example, the empirical patterns resulting from social discrimination we describe in Proposition 2 are inconsistent with statistical discrimination against workers with minority traits (firms owned by minority employers should be more productive since they could hire majority and minority workers), taste-based discrimination by employers (same effect) or taste-based discrimination by workers with majority traits (minority workers would have better employment prospect being hired

[10]Fershtman and Gneezy (2001) and Gneezy, List, and Price (2012) use, respectively, laboratory and field experiments to distinguish between statistical and taste-based discriminations and find little evidence in favor of the latter. Evidence that is consistent with statistical discrimination is also presented by Glover, Pallais, and Pariente (2017) who analyzed the performance of cashiers distinguished by majority or minority identity in a French grocery store chain. They show that manager bias negatively affects minority job performance and when working with unbiased managers, minority workers perform substantially better than non-minorities.

by majority and minority employers). On the other hand, the observables we uncover are not inconsistent with all citizens with majority traits being prejudiced against the minority as a whole. Another interesting avenue for future research is to consider how social discrimination affects integration: does it lead to more separation (e.g., Carvalho, 2012; Bisin, Patacchini, Verdier, and Zenou, 2016; Fouka, 2019a) or more assimilation (e.g., adoption of the majority names as in Fouka, 2019b)?

References

Abdelgadir, A., and V. Fouka (2019): "Secular Policies and Muslim Integration in the West: The Effects of the French Headscarf Ban," *Stanford Working Paper*.

Adida, C. L., D. D. Laitin, and M.-A. Valfort (2016): *Why Muslim integration fails in Christian-heritage societies*. Harvard University Press.

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

Akerlof, G. A., and R. E. Kranton (2000): "Economics and identity," *The Quarterly Journal of Economics*, 115(3), 715–753.

——— (2010): "Identity Economics: How Identities Shape Our Work, Wages, and Well-Being," *Princeton, NJ: Princeton University Press*.

Alt, J., and T. Iversen (2017): "Inequality, Labor Market Segmentation, and Preferences for Redistribution," *American Journal of Political Science*, 61(1), 21–36.

Appiah, K. A. (2018): *The Lies That Bind: Rethinking Identity*. Profile Books.

Arrow, K., et al. (1973): "The theory of discrimination," *Discrimination in labor markets*, 3(10), 3–33.

Austen-Smith, D., and R. G. Fryer Jr (2005): "An economic analysis of acting white," *The Quarterly Journal of Economics*, 120(2), 551–583.

Basu, K. (2005): "Racial conflict and the malignancy of identity," *The Journal of Economic Inequality*, 3(3), 221–241.

Becker, G. S. (2010): *The economics of discrimination, 2nd edition*. University of Chicago press.

Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou (2016): "Bend it like Beckham: Ethnic identity and integration," *European Economic Review*, 90, 146–164.

Bramoullé, Y., and S. Goyal (2016): "Favoritism," *Journal of Development Economics*, 122, 16–27.

Carvalho, J.-P. (2012): "Veiling," *The Quarterly Journal of Economics*, 128(1), 337–370.

Cavounidis, C., and K. Lang (2015): "Discrimination and worker evaluation," *National Bureau of Economic Research Working Paper*, No. 21612.

Choy, J. P. (2017): "Social Division with Endogenous Hierarchy," *The Economic Journal*, 128(615), 2711–2742.

Dancygier, R. M. (2010): *Immigration and conflict in Europe*. Cambridge University Press.

——— (2017): *Dilemmas of inclusion: Muslims in European politics*. Princeton University Press.

Daskalova, V. (2018): "Discrimination, social identity, and coordination: An experiment," *Games and Economic Behavior*, 107, 238–252.

Eeckhout, J. (2006): "Minorities and endogenous segregation," *The Review of Economic Studies*, 73(1), 31–53.

Eguia, J. X. (2017): "Discrimination and assimilation at school," *Journal of Public Economics*, 156, 48–58.

European Commission (2008): "Conclusions and Recommendations of the European Commission's Network "Ethnic Minority Businesses"," .

Fearon, J. D., and D. D. Laitin (2000): "Violence and the Social Construction of Identity," *International Organization*, 4(54), 845–877.

Fernandez, R., and G. Levy (2008): "Diversity and redistribution," *Journal of public economics*, 92(5-6), 925–943.

Fershtman, C., and U. Gneezy (2001): "Discrimination in a segmented society: An experimental approach," *The Quarterly Journal of Economics*, 116(1), 351–377.

Fouka, V. (2019a): "Backlash: the unintended effects of language prohibition in US schools after World War I," *Review of Economic Studies*, forthcoming.

——— (2019b): "How do immigrants respond to discrimination? The case of Germans in the US during World War I," *American Political Science Review*, 113(2), 405–422.

Glover, D. (2019): "Job Search and Intermediation under Discrimination: Evidence from Terrorist Attacks in France," *Chaire Securisation des Parcours Professionels Working Paper*, (2019-02).

GLOVER, D., A. PALLAIS, AND W. PARIENTE (2017): "Discrimination as a self-fulfilling prophecy: Evidence from French grocery stores," *The Quarterly Journal of Economics*, p. qjx006.

GNEEZY, U., J. LIST, AND M. K. PRICE (2012): "Toward an understanding of why people discriminate: Evidence from a series of natural field experiments," Discussion paper.

GROSSMAN, G. M., AND E. HELPMAN (2018): "Identity Politics and Trade Policy," *National Bureau of Economic Research Working Paper*, No. 25348.

HARBAUGH, R., AND T. TO (2014): "Opportunistic discrimination," *European Economic Review*, 66, 192–204.

HUBER, J. D. (2017): *Exclusion by elections: inequality, ethnic identity, and democracy*. Cambridge University Press.

KAMPHORST, J. J., AND O. H. SWANK (2016): "Don't demotivate, discriminate," *American Economic Journal: Microeconomics*, 8(1), 140–65.

KRASA, S., AND M. POLBORN (2014): "Social ideology and taxes in a differentiated candidates framework," *American Economic Review*, 104(1), 308–22.

LANDA, D., AND D. DUELL (2015): "Social identity and electoral accountability," *American Journal of Political Science*, 59(3), 671–689.

LEVY, G. (2004): "A model of political parties," *Journal of Economic theory*, 115(2), 250–277.

MATAKOS, K., AND D. XEFTERIS (2017): "Divide and rule: redistribution in a model with differentiated candidates," *Economic Theory*, 63(4), 867–902.

——— (2019): "Class vs. Identity: The Effect of Candidates Race on the Inequality-Redistribution Nexus," *Working Paper*.

PENN, E. M. (2008): "Citizenship versus ethnicity: The role of institutions in shaping identity choice," *The Journal of Politics*, 70(4), 956–973.

PESKI, M., AND B. SZENTES (2013): "Spontaneous discrimination," *American Economic Review*, 103(6), 2412–36.

Phelps, E. S. (1972): "The statistical theory of racism and sexism," *The american economic review*, pp. 659–661.

Ram, M., and T. Jones (2008): *Ethnic Minorities in Business.* SERTeam.

Ramachandran, R., and C. Rauh (2013): "Discrimination without Taste-How Discrimination can spillover and persist," *Working Paper.*

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

Rzepnikowska, A. (2019): "Racism and xenophobia experienced by Polish migrants in the UK before and after Brexit vote," *Journal of Ethnic and Migration Studies*, 45(1), 61–77.

Schnakenberg, K. E. (2013): "Group identity and symbolic political behavior," *Schnakenberg, Keith E." Group Identity and Symbolic Political Behavior." Quarterly Journal of Political Science*, 9(2014), 137–167.

Shayo, M. (2009): "A model of social identity with an application to political economy: Nation, class, and redistribution," *American Political science review*, 103(2), 147–174.

Zussman, A. (2013): "Ethnic Conflict and Segregation in the Marketplace: Evidence from Israel," *Unpublished manuscript, Hebrew University.*

Proofs

A Proofs of Section 4

Proof of Lemma 1

An employer's continuation value when his position is filled is given by Equation 1. Equation 1 directly indicates that $V^f(\theta)$ is strictly increasing with θ. Hence, there exists a threshold $\theta^{ND} \in [w, 1]$ such that $\max\{V^f(\theta), V^o\} = V^f(\theta)$ for all $\theta \geq \theta^{ND}$. Using the reasoning above, we use Equation 2 to rewrite V^o as:

$$
\begin{aligned}
V^o =& 0 + \beta E_S\big(V^f(\theta)\big|\theta \geq \theta^{ND}\big)(1 - S(\theta^{ND})) + \beta V^o \times S(\theta^{ND}) \\
\Leftrightarrow \quad V^o =& \frac{\beta E_S\big(V^f(\theta)\big|\theta \geq \theta^{ND}\big)(1 - S(\theta^{ND}))}{1 - \beta S(\theta^{ND})}
\end{aligned} \tag{A.1}
$$

Now,

$$
\begin{aligned}
E_S\big(V^f(\theta)\big|\theta \geq \theta^{ND}\big)(1 - S(\theta^{ND})) =& \int_{\theta^{ND}}^{1} V^f(\theta) dS(\theta) \\
=& \frac{\int_{\theta^{ND}}^{1}(1-\tau)(\theta - w) dS(\theta) + (1 - S(\theta^{ND}))\beta\delta V^o}{1 - \beta(1-\delta)}
\end{aligned}
$$

Plugging this result into Equation A.1 and rearranging, we obtain:

$$
V^o = \frac{\beta \int_{\theta^{ND}}^{1}(1-\tau)(\theta - w) dS(\theta)}{(1-\beta)(1 - \beta(1-\delta) + \beta(1 - S(\theta^{ND})))} \tag{A.2}
$$

Notice further that by definition—assuming for now θ^{ND} is interior—$V^f(\theta^{ND}) = V^o$. So θ^{ND} is the solution to:

$$
\begin{aligned}
V^f(\theta^{ND}) =& V^o \\
\Leftrightarrow \quad (1-\tau)(\theta^{ND} - w) =& (1-\beta)V^o
\end{aligned}
$$

The second equality comes from Equation 1. Now using Equation A.2, we obtain that θ^{ND} is the solution to:

$$
\theta^{ND} - w = \frac{\beta \int_{\theta^{ND}}^{1}(\theta - w) dS(\theta)}{1 - \beta(1-\delta) + \beta(1 - S(\theta^{ND}))} \tag{A.3}
$$

Using Equation A.3, one can check that a threshold exists and is indeed always interior. $\theta^{ND} = w$ would require $0 \geq \frac{\beta \int_w^1 (\theta - w) dS(\theta)}{1-\beta(1-\delta)+\beta(1-S(w))}$ which is of course impossible. $\theta^{ND} = 1$ would require $1 - w \leq 0$, again a contradiction. We now show that θ^{ND} is unique.

Simple rearrangement of Equation A.3 given the employer's threshold strategy yields:

$$\begin{aligned} \theta^{ND} - w =& \frac{\beta\Big(E_\theta\big(\theta|\theta \geq \theta^{ND}\big) - w\Big)(1-\theta^{ND})}{1-\beta(1-\delta)+\beta(1-\theta^{ND})} \\ \Leftrightarrow \quad \frac{\theta^{ND} - w}{\frac{1+\theta^{ND}}{2} - w} =& \frac{\beta(1-\theta^{ND})}{1-\beta(1-\delta)+\beta(1-\theta^{ND})} \end{aligned} \tag{A.4}$$

It can be checked that the left-hand side of Equation A.4 is strictly increasing with θ^{ND}, whereas the right-hand side is strictly decreasing. Hence, θ^{ND} is unique. □

The next preliminary lemmas set the ground for the proof of Lemma 2. Throughout, we assume that discrimination equilibria exist. We prove existence in Proposition 1. In the proofs, we use the shorthand majority (minority) worker to denote a worker with majority (minority) social identity.

Lemma A.1. *In any work discrimination equilibrium,*

(i) A worker with a majority social identity only agrees to work with an employer with a majority social identity;

(ii) An employer with a majority social identity never hires a worker with a minority social identity.

Proof. Under our assumptions (exogenous wage and match-specific productivity), all majority workers are identical. Hence, if one majority worker agrees to work with a minority employer, all must also prefer to do so. Since, in turn, a minority worker always agrees to work with both a majority and minority employers, all workers work with all employers. Hence, a firm has no reason to discriminate. This proves point (i).

For point (ii), note that give point (i), if (ii) does not hold then the proportion of minority workers or the proportion of majority employers go to 0 in equilibrium (since we only consider the steady state and no majority worker works with minority firm so becomes a minority member or changes the identity of minority firm by (i)), and there is no longer any work discrimination.[11] □

We now detail the properties of a discrimination assessment. That is, we describe the best responses of workers and employers in the labor market subgame when other players play a discrimination strategy

[11] As we will also see below, demand for discrimination requires that the share of majority workers and firms be interior.

(i.e., strategies satisfy the conditions of Lemma A.1). We then determine conditions such that all labor market actors' discrimination strategies are mutual best response.

Denote μ^M the proportion of majority citizen among unemployed workers. Employers when taking their employment decision takes this proportion as given. Lemma A.2 then shows the hiring best response takes the form of threshold strategy. We denote $V^f_{J,K}(\theta)$ the continuation value of an employer with physical trait $J \in \{M, m\}$ and social trait $K \in \{M, m\}$ when the position is filled (f) with an employee with productivity θ. In turn, $V^o_{J,K}$ denotes the continuation value for a similar employer when his position is open.

Lemma A.2. *In a work discrimination assessment, there exists unique* $\overline{\theta}_M(\mu^M), \overline{\theta}_m(\mu^M) \in (w, 1)^2$ *such that the employers' best responses as a function of* μ^M *satisfy a majority (minority) employer hires a majority (minority) worker if and only if* $\theta \geq \overline{\theta}_M(\mu^M)$ $(\theta \geq \overline{\theta}_m(\mu^M))$.

Further, $\overline{\theta}_M(\mu^M)$ *is strictly increasing with* μ^M *and* $\overline{\theta}_m(\mu^M)$ *is strictly decreasing with* μ^M.

Proof. A majority (minority) employer's continuation value when his position is open is given by Equation 4 (Equation 5). Observe from the equations in the main text that the key difference between Equation 2 and Equation 4 is that a majority employer needs to be matched with a majority worker before considering hiring. In turn, for Equation 5, a minority firm needs to be matched with a minority worker to have a chance to hire.

As before, an employer uses a cutoff strategy: hires if and only if $\theta \geq \overline{\theta}_M$ $(\theta \geq \overline{\theta}_m)$ for a majority (minority) employer. By the same computations as in the proof of Lemma 1, the majority cutoff is interior and defined by the following equation:

$$\frac{\overline{\theta}_M - w}{\frac{1+\overline{\theta}_M}{2} - w} = \frac{\beta\mu^M(1 - \overline{\theta}_M)}{1 - \beta + \beta[\delta + \mu^M(1 - \overline{\theta}_M)]} \tag{A.5}$$

As for Lemma 1, the threshold is unique. By the Implicit Function Theorem (for which all conditions are satisfied), $\overline{\theta}_M(\mu^M)$ is strictly increasing with μ^M (the right-hand side is strictly increasing with μ^M and strictly decreasing with $\overline{\theta}_M$, the left-hand side is strictly increasing with $\overline{\theta}_M$).

Similarly, the minority cutoff is interior and defined by the following equation:

$$\frac{\overline{\theta}_m - w}{\frac{1+\overline{\theta}_m}{2} - w} = \frac{\beta(1 - \mu^M)(1 - \overline{\theta}_m)}{1 - \beta + \beta[\delta + (1 - \mu^M)(1 - \overline{\theta}_m)]} \tag{A.6}$$

By the Implicit Function Theorem, $\overline{\theta}_m(\mu^M)$ is strictly decreasing with μ^M. □

Lemma A.2 determines the employers' strategy for some proportion of majority worker among unemployed, μ^M. In equilibrium, however, μ^M is endogenous to the employers' strategy. The next lemmas show that there exists a unique pair of equilibrium hiring thresholds θ_M, θ_m. That is, we want to ensure that the problem does not blow out. To determine these equilibrium thresholds, we first look at the property of μ^M as a function of the two thresholds, taken as exogenous, as well as some other comparative statics.

Lemma A.3. *In a work discrimination assessment, there exists a unique proportion of majority workers $\mu^M(\theta_M, \theta_m)$ among the unemployed. Further $\mu^M(\cdot)$ is strictly increasing with θ_M and α^W and strictly decreasing with θ_m and α^F.*

Proof. Before looking at the proportion of majority workers among unemployed, it is useful to look at the mass of positions filled in majority firms—denoted F_M^f—and in minority firms—denoted F_m^f. Since we consider only the labour market steady state, we obtain that F_M^f and F_m^f must satisfy, respectively:

$$F_M^f = (1-\delta)F_M^f + \mu^M(1-S(\theta_M))(\alpha^F - F_M^f)$$
$$F_m^f = (1-\delta)F_m^f + (1-\mu^M)(1-S(\theta_m))((1-\alpha^F) - F_m^f)$$

On the right-hand side, the first term is the mass of firms which started with a filled position at the beginning of the period and remains with a filled position at the end; the second term is the mass of firms that fills their position this period. The equality with the left-side indicates that the mass for both majority and minority employers is constant over time in the steady state.
Rearranging, we obtain:

$$\frac{F_M^f(\mu^M)}{\alpha^F} = \frac{\mu^M(1-S(\theta_M))}{\mu^M(1-S(\theta_M))+\delta} \tag{A.7}$$
$$\frac{F_m^f(\mu^M)}{1-\alpha^F} = \frac{(1-\mu^M)(1-S(\theta_m))}{(1-\mu^M)(1-S(\theta_m))+\delta} \tag{A.8}$$

Observe that the mass of majority (minority) workers equals the mass of majority (minority) filled positions in a work discrimination equilibrium (since the labour market is completely segregated by Lemma A.1). Hence, we can write the share of majority workers among the unemployed as:

$$\mu^M = \frac{\alpha^W - F_M^f(\mu^M)}{\alpha^W - F_M^f(\mu^M) + \big((1-\alpha^W) - F_m^f(\mu^M)\big)} \tag{A.9}$$

To see that a solution to Equation A.9 exists, notice that when $\mu^M = 0$, the right-hand side Equation A.9 is strictly positive, whereas when $\mu^M = 1$, the right-hand side is strictly less than 1. By continuity, a solution exists. To prove uniqueness, notice that using Equation A.7 and Equation A.8, $F_M^f(\mu^M)$ is strictly increasing with μ^M and $F_m^f(\mu^M)$ is strictly decreasing with μ^M (fixing the thresholds). Define the function $H(x, y) = \frac{\alpha^W - x}{\alpha^W - x + \left((1-\alpha^W)-y\right)}$. The function is strictly decreasing in x and strictly increasing in y. Putting the results together, the right-hand side of Equation A.9 is strictly decreasing with μ^M. The remaining properties follow from the Implicit Function Theorem since the right-hand side is strictly increasing in θ_M, strictly decreasing in θ_m, strictly increasing in α^W, and strictly decreasing in α^F. □

Our next result establishes that the probability of hiring a majority employee is decreasing with the strictness of majority employers' hiring threshold, even if there is a greater proportion of majority workers among the unemployed as per the previous lemma.

Lemma A.4. *In a work discrimination assessment,* $\mu^M(\theta_M, \theta_m)(1 - S(\theta_M))$ *is strictly decreasing with* θ_M*;* $(1 - \mu^M(\theta_M, \theta_m))(1 - S(\theta_m))$ *is strictly decreasing with* θ_m*.*

Proof. We only prove the result for the majority, a similar reasoning holds for the minority. The proof proceeds by contradiction. Suppose $\mu^M(\theta_M, \theta_m)(1 - S(\theta_M))$ is weakly increasing with θ_M. By definition of $F_M^f(\mu^M)$ (Equation A.7), $F_M^f(\cdot)$ is weakly increasing with θ_M. This implies that $\mu^M(\theta_M, \theta_m)$ is weakly decreasing with θ_M (by Equation A.9 and the Implicit Function Theorem). Consequently, $\mu^M(\theta_M, \theta_m)(1 - S(\theta_M))$ is strictly decreasing with θ_M, a contradiction. □

The next Lemma proves the existence of the equilibrium hiring threshold (assuming existence of work discrimination equilibrium) making use of the preliminary lemmas above and the Bouwer's fixed point theorem.

Lemma A.5. *In any work discrimination equilibrium, the employers' strategy is characterized by a pair of thresholds* $\widehat{\theta}_M, \widehat{\theta}_m \in (w, 1)^2$ *such that a majority (minority) employer hires a majority (minority) worker if and only if* $\theta \geq \widehat{\theta}_M$ *(*$\theta \geq \widehat{\theta}_m$*).*

Proof. The equilibrium thresholds are defined as the solution to:

$$\frac{\theta_M - w}{\frac{1+\theta_M}{2} - w} = \frac{\beta \mu^M(\theta_M, \theta_m)(1 - \theta_M)}{1 - \beta + \beta[\delta + \mu^M(\theta_M, \theta_m)(1 - \theta_M)]} \tag{A.10}$$

$$\frac{\theta_m - w}{\frac{1+\theta_m}{2} - w} = \frac{\beta(1 - \mu^M(\theta_M, \theta_m))(1 - \theta_m)}{1 - \beta + \beta[\delta + (1 - \mu^M(\theta_M, \theta_m))(1 - \theta_m)]} \tag{A.11}$$

Fixing a $\theta_m \in [w, 1]$, repeating the same reasoning as in the proof of Lemma 1, it can be checked that there exists a unique $\tilde{\theta}_M(\theta_m) \in (w, 1)$ solving Equation A.10. Further, since $\mu^W(\theta_M, \theta_m)$ is continuous in θ_m, the solution $\tilde{\theta}_M(\theta_m)$ is also continuous in θ_m. Similarly, fixing a $\theta_M \in [w, 1]$, there exists a unique $\tilde{\theta}_m(\theta_M) \in (w, 1)$ solving Equation A.11, and $\tilde{\theta}_m(\theta_M)$ is continuous in θ_M.

We can now define the two-dimensional function $\Theta(\theta_M, \theta_m) = (\tilde{\theta}_M(\theta_m), \tilde{\theta}_m(\theta_M))$. Using the reasoning above, this function is continuous and mapping the convex set $[w, 1]$ into itself. Hence, applying Bouwer fixed point theorem, a fixed point exists. Thus, there exists equilibrium hiring thresholds as claimed. □

Our next lemma proves that the pair of threshold is unique and so is the employers' strategy.

Lemma A.6. *In the work discrimination equilibrium, the employers' strategy is characterized by a unique pair of equilibrium thresholds $\theta_M^D, \theta_m^D \in (w, 1)^2$ such that a majority (minority) employer hires a majority (minority) worker if and only if $\theta \geq \theta_M^D$ ($\theta \geq \theta_m^D$).*

Proof. Consider the following function

$$\mathcal{D}(\theta_M) = \frac{\theta_M - w}{\frac{1+\theta_M}{2} - w} - \frac{\beta\mu^M(\theta_M, \tilde{\theta}_m(\theta_M))(1-\theta_M)}{1-\beta+\beta[\delta+\mu^M(\theta_M, \tilde{\theta}_m(\theta_M))(1-\theta_M)]}, \tag{A.12}$$

with $\tilde{\theta}_m(\theta_M)$ the unique solution to Equation A.11. Observe that any solution to $\mathcal{D}(\theta_M) = 0$ is a fixed point for the hiring problem. We cannot prove uniqueness directly since by the Implicit Function Theorem, $\tilde{\theta}_m(\theta_M)$ is increasing with θ_M as $\mu^M(\theta_M, \theta_m)$ is increasing with θ_M so the total derivative of $\mu^M(\theta_M, \tilde{\theta}_m(\theta_M))(1-\theta_M)$ can no longer be signed. We instead prove uniqueness by contradiction. Suppose there are $K > 1$ solutions to $\mathcal{D}(\theta_M) = 0$ which we index by k—$\widehat{\theta}_M^{(k)}$—and order such that $\widehat{\theta}_M^{(1)} < \widehat{\theta}_M^{(2)} < ... < \widehat{\theta}_M^{(K)}$ for all $\alpha^W \in (1/2.1)$.

As a first step, notice that as $\alpha^W \to 1$, then $\mu^M(\theta_M, \theta_m) \to 1$ for all $\theta_M, \theta_m \in [w, 1]^2$. Hence, it must be that for all $k \in \{1, ..., K\}$, $\lim_{\alpha^W \to 1} \widehat{\theta}_M^{(k)} = \theta^{ND}$ (since Equation A.10 become Equation A.4).

By the usual reasoning $\mathcal{D}(w) < 0$ and $\mathcal{D}(1) > 0$. Further, any odd solution (i.e., k odd) to $\mathcal{D}(\theta_M) = 0$ satisfies $\mathcal{D}'(\widehat{\theta}_M^{(k)}) > 0$ and any even solution (k even) satisfies $\mathcal{D}'(\widehat{\theta}_M^{(k)}) < 0$. Totally differentiating $\mathcal{D}(\cdot)$ with respect to α^W, we obtain: $\frac{d\mathcal{D}(\widehat{\theta}^{(k)})}{d\alpha^W} = \mathcal{D}'(\widehat{\theta}_M^{(k)})\frac{\partial\widehat{\theta}_M^{(k)}}{\partial\alpha^W} + \frac{\partial\mathcal{D}(\widehat{\theta}_M^{(k)})}{\partial\alpha^W}$. Observe that $\frac{\partial\mathcal{D}(\widehat{\theta}_M^{(k)})}{\partial\alpha^W} \propto -\left(\frac{\partial\mu^M((\widehat{\theta}_M^{(k)}, \tilde{\theta}_m(\widehat{\theta}_M^{(k)}))}{\partial\alpha^W} + \frac{\partial\mu^M(\widehat{\theta}_M^{(k)}, \tilde{\theta}_m((\widehat{\theta}_M^{(k)}))}{\partial\theta_m}\frac{\partial\tilde{\theta}_m((\widehat{\theta}_M^{(k)})}{\partial\alpha^W}\right)$. By Lemma A.3, we know that for all θ_M, θ_m, $\frac{\partial\mu^M(\theta_M,\theta_m)}{\partial\alpha^W} > 0$ and $\frac{\partial\mu^M(\theta_M,\theta_m)}{\partial\theta_m} < 0$. Further, by the Implicit Function Theorem and using Lemma A.2, $\frac{\partial\tilde{\theta}_m(\theta_M)}{\partial\alpha^W} < 0$ for all θ_M. Hence, $\frac{\partial\mathcal{D}(\widehat{\theta}_M^{(k)})}{\partial\alpha^W} < 0$.

Using both results above and the Implicit Function Theorem, this implies that $\widehat{\theta}_M^{(k)}$ is strictly increasing with α^W if k is odd, and strictly decreasing if k is even.[12] This, in turn, implies that $\widehat{\theta}_M^{(3)} < \widehat{\theta}_M^{(2)}$ for all $\alpha^W \in (1/2, 1)$ (since $\widehat{\theta}_M^{(2)}$ decreases with α^W, it converges to θ^{ND}; since $\widehat{\theta}_M^{(3)}$ increases with α^W, it converges to θ^{ND} from below). This contradicts $\widehat{\theta}_M^{(2)} < \widehat{\theta}_M^{(3)}$.

Consequently, the solution to $\mathcal{D}(\theta_M) = 0$ is unique and so are the pair of equilibrium thresholds. □

Proof of Lemma 2

The proof follows directly from Lemma A.1 and Lemma A.6. □

Proof of Corollary 1

Point (i). The proof that $\theta_M^D(\alpha^W, \alpha^F)$ is strictly increasing with α^W follows directly from the proof of Lemma A.6. Recalling that the function $\mathcal{D}(\theta_M)$ is defined in Equation A.12 and using the same reasoning as in Lemma A.6, we obtain (ignoring arguments for simplicity) that (a) $\frac{\partial \theta_M^D}{\partial \alpha^F}$ has the opposite sign than $\frac{\partial \mathcal{D}(\theta_M^D)}{\partial \alpha^F}$ (since $\mathcal{D}'(\theta_M^D) > 0$) and (b) $\frac{\partial \mathcal{D}(\theta_M^D)}{\partial \alpha^F}$ has the opposite sign than $\frac{\partial \mu^M((\theta_M^D, \tilde{\theta}_m(\theta_M^D))}{\partial \alpha^F} + \frac{\partial \mu^M(\theta_M^D, \tilde{\theta}_m(\theta_M^D))}{\partial \theta_m} \frac{\partial \tilde{\theta}_m(\theta_M^D)}{\partial \alpha^F}$. By Lemma A.3, $\frac{\partial \mu^M(\theta_M^D, \tilde{\theta}_m(\theta_M^D))}{\partial \theta_m} < 0$ and $\frac{\partial \mu^M((\theta_M^D, \tilde{\theta}_m(\theta_M^D))}{\partial \alpha^F} < 0$. Using Lemma A.2, this last inequality implies $\frac{\partial \tilde{\theta}_m(\theta_M^D)}{\partial \alpha^F} > 0$ since $\tilde{\theta}_m$ depends on α^F only through μ^M. Hence, $\frac{\partial \mu^M((\theta_M^D, \tilde{\theta}_m(\theta_M^D))}{\partial \alpha^F} + \frac{\partial \mu^M(\theta_M^D, \tilde{\theta}_m(\theta_M^D))}{\partial \theta_m} \frac{\partial \tilde{\theta}_m(\theta_M^D)}{\partial \alpha^F}$ is strictly negative and $\frac{\partial \theta_M^D(\alpha^W, \alpha^F)}{\partial \alpha^F} < 0$.

Point (ii). For point (ii), let's define

$$\mathcal{F}(\theta_m) = \frac{\theta_m - w}{\frac{1+\theta_m}{2} - w} - \frac{\beta(1 - \mu^M(\tilde{\theta}_M(\theta_m), \theta_m))(1 - \theta_m)}{1 - \beta + \beta[\delta + (1 - \mu^M(\tilde{\theta}_M(\theta_m), \theta_m))(1 - \theta_m)]}, \tag{A.13}$$

with $\tilde{\theta}_M(\theta_m)$ the unique solution to Equation A.10. Note that by Lemma A.6, there exists a unique solution to $\mathcal{F}(\theta_m) = 0$ and this solution is θ_m^D (ignoring arguments). Since $\mathcal{F}(w) < 0$ and $\mathcal{F}(1) > 0$, we obtain that $\mathcal{F}'(\theta_m^D) > 0$. As for point (i), for $\alpha \in \{\alpha^W, \alpha^F\}$, $\frac{\partial \theta_m^D}{\partial \alpha}$ has the opposite sign than $\frac{\partial \mathcal{F}(\theta_m^D)}{\partial \alpha}$. Further, using Equation A.13, $\frac{\partial \mathcal{F}(\theta_m^D)}{\partial \alpha}$ has the same sign as $\frac{\partial \mu^M(\tilde{\theta}_M(\theta_m^D), \theta_m^D)}{\partial \alpha} + \frac{\partial \mu^M(\tilde{\theta}_M(\theta_m^D), \theta_m^D)}{\partial \theta_M} \frac{\partial \tilde{\theta}_M(\theta_m^D)}{\partial \alpha}$. By Lemma A.3, $\frac{\partial \mu^M(\tilde{\theta}_M(\theta_m^D), \theta_m^D)}{\partial \alpha^W} > 0$. $\frac{\partial \mu^M(\tilde{\theta}_M(\theta_m^D), \theta_m^D)}{\partial \alpha^F} < 0$, and $\frac{\partial \mu^M(\tilde{\theta}_M(\theta_m^D), \theta_m^D)}{\partial \theta_M} > 0$. Further, $\frac{\partial \tilde{\theta}_M(\theta_m^D)}{\partial \alpha^W} > 0$ and $\frac{\partial \tilde{\theta}_M(\theta_m^D)}{\partial \alpha^F} < 0$ since $\tilde{\theta}_M(\theta_m)$ only depends on α^W and α^F through μ^M. Hence, we obtain that $\frac{\partial \theta_m^D}{\partial \alpha^W} < 0$ and $\frac{\partial \theta_m^D}{\partial \alpha^F} > 0$ as claimed.

Point (iii). To see this last point, note that $\mu^M \xrightarrow{\alpha^W \to 1} 1$ by Equation A.9 (given $F_m^f \leq 1 - \alpha^W$). Hence, Equation A.10 becomes equivalent to Equation A.4 when $\alpha^W \to 1$, and the unique solution is $\tilde{\theta}_M(\theta_m) = \theta^{ND}$ for all θ_m. In turn, Equation A.11 becomes $\frac{\theta_m - w}{\frac{\theta_m + 1}{2} - w} = 0$ as $\alpha^W \to 1$, and the unique

[12] Note that this comparative statics also implies that the ordering we postulate cannot change with α^W for $k = 2$. If $\widehat{\theta}_M^{(2)} < \widehat{\theta}_M^{(3)}$ for some α^1, this is also true for all $\alpha^W > \alpha^1$.

solution is $\tilde{\theta}_m(\theta_M) = w$ for all θ_M. $\square$

We now turn to proving Proposition 1. The proof requires several preliminary lemmas on the properties of several equilibrium objects. First, we establish some properties of the equilibrium threshold.

Lemma A.7. *For all $\alpha^F \in (1/2, 1)$, there exists a unique $\underline{\alpha}^W(\alpha^F) \in (1/2, 1)$ such that $\theta_M^D(\alpha^W, \alpha^F) \geq \theta_m^D(\alpha^W, \alpha^F)$ for all $\alpha^W \geq \underline{\alpha}^W(\alpha^F)$.*
The threshold $\underline{\alpha}^W(\alpha^F)$ satisfies the following properties:
(i) $\underline{\alpha}^W(\alpha^F)$ is strictly increasing with α^F;
(ii) $\underline{\alpha}^W(1/2) = 1/2$;
(iii) $\lim_{\alpha^F \to 1} \underline{\alpha}^W(\alpha^F) < 1$.

Proof. From Corollary 1, we already know that (a) $\theta_M^D(\alpha^W, \alpha^F)$ ($\theta_m^D(\alpha^W, \alpha^F)$) is strictly increasing (decreasing) in α^W, (b) $\theta_M^D(\alpha^W, \alpha^F)$ ($\theta_m^D(\alpha^W, \alpha^F)$) is strictly decreasing (increasing) in α^F, and (c) for all α^F, $\lim_{\alpha^W \to 1} \theta_M^D(\alpha^W, \alpha^F) = \theta^{ND} > \lim_{\alpha^W \to 1} \theta_m^D(\alpha^W, \alpha^F) = w$. Further, given equilibrium uniqueness, it must be (slightly abusing notation) that (d) if $\alpha^W = \alpha^F = 1/2$, then $\theta_m^D(1/2, 1/2) = \theta_M^D(1/2, 1/2)$. Property (d) together with property (b) implies that for all $\alpha^F > 1/2$, $\theta_m^D(1/2, \alpha^F) > \theta_M^D(1/2, \alpha^F)$. This result coupled with properties (a) and (c) implies that there exists a unique $\underline{\alpha}^W(\alpha^F) \in (1/2, 1)$ such that $\theta_M^D(\alpha^W, \alpha^F) \geq \theta_m^D(\alpha^W, \alpha^F)$ for all $\alpha^W \geq \underline{\alpha}^W(\alpha^F)$.
For the rest, point (i) follows from the Implicit Function Theorem using properties (a) and (b). Point (ii) is a direct consequence of properties (d) and (a). Point (iii) follows from (c) holding for all α^F. $\square$

We now study how the probability of hiring a majority worker (a minority worker) for majority firm (minority firm) with openings depends on the proportion of majority type in the population. For ease of exposition, we ignore arguments in the equilibrium threshold.

Lemma A.8. *In a work discrimination equilibrium,*
$\mu^M(\theta_M^D, \theta_m^D)(1 - \theta_M^D)$ is strictly increasing with α^W and strictly decreasing with α^F;
$(1 - \mu^M(\theta_M^D, \theta_m^D))(1 - \theta_m^D)$ is strictly decreasing with α^W and strictly increasing with α^F.

Proof. Recall that the hiring thresholds satisfy (ignoring arguments)

$$\frac{\theta_M^D - w}{\frac{1+\theta_M^D}{2} - w} = \frac{\beta \mu^M(\theta_M^D, \theta_m^D)(1 - \theta_M^D)}{1 - \beta + \beta[\delta + \mu^M(\theta_M^D, \theta_m^D)(1 - \theta_M^D)]} \tag{A.14}$$

$$\frac{\theta_m^D - w}{\frac{1+\theta_m^D}{2} - w} = \frac{\beta(1 - \mu^M(\theta_M^D, \theta_m^D))(1 - \theta_m^D)}{1 - \beta + \beta[\delta + (1 - \mu^M(\theta_M^D, \theta_m^D))(1 - \theta_m^D)]} \tag{A.15}$$

It must then be that $\frac{\partial \mu^M(\theta_M^D,\theta_m^D)(1-\theta_M^D)}{\partial \alpha}$ has the same sign as $\frac{\partial \theta_M^D}{\partial \alpha}$ and $\frac{\partial (1-\mu^M(\theta_M^D,\theta_m^D))(1-\theta_m^D)}{\partial \alpha}$ has the same sign as $\frac{\partial \theta_m^D}{\partial \alpha}$ for $\alpha \in \{\alpha^W, \alpha^F\}$. The claim then follows from Corollary 1. $\square$

For our next result, we introduce the proportion of majority employer among firms with opening which we denote $\kappa^F(\theta_M, \theta_m)$. We detail the properties of this equilibrium object.

Lemma A.9. *The proportion of majority employers among firms with openings is:*

$$\kappa^F(\theta_M^D, \theta_m^D) = \frac{\alpha^F}{\alpha^F + (1-\alpha^F)\frac{\mu^M(\theta_M^D,\theta_m^D)(1-\theta_M^D)+\delta}{(1-\mu^M(\theta_M^D,\theta_m^D))(1-\theta_m^D)+\delta}}. \tag{A.16}$$

$\kappa^F(\theta_M^D, \theta_m^D)$ is strictly decreasing with α^W and strictly increasing with α^F.

Proof. The mass of majority employers with filled position is F_M^f defined in Equation A.7. The mass of minority employers with filled position is F_m^f defined in Equation A.8. Hence, the proportion of majority employers among firms with open position is

$$\begin{aligned}\kappa^F(\theta_M^D, \theta_m^D) =& \frac{\alpha^F - F_M^f}{\alpha^F - F_M^f + [(1-\alpha^F) - F_m^f]} \\ =& \frac{\alpha^F}{\alpha^F + (1-\alpha^F)\frac{\mu^M(\theta_M^D,\theta_m^D)(1-\theta_M^D)+\delta}{(1-\mu^M(\theta_M^D,\theta_m^D))(1-\theta_m^D)+\delta}}\end{aligned} \tag{A.17}$$

Observe that $\kappa^F(\theta_M^D, \theta_m^D)$ depends on α^W only through $\frac{\mu^M(\theta_M^D,\theta_m^D)(1-\theta_M^D)+\delta}{(1-\mu^M(\theta_M^D,\theta_m^D))(1-\theta_m^D)+\delta}$. By Lemma A.8, this ratio is strictly increasing with α^W. Hence, $\kappa^F(\cdot)$ is strictly decreasing with α^W.
In turn, $\frac{\partial \kappa^F(\theta_M^D,\theta_m^D)}{\partial \alpha^F}$ has the same sign as

$$\begin{aligned}\Delta =& \alpha^F + (1-\alpha^F)\frac{\mu^M(\theta_M^D,\theta_m^D)(1-\theta_M^D)+\delta}{(1-\mu^M(\theta_M^D,\theta_m^D))(1-\theta_m^D)+\delta} \\ & - \alpha^F\left(1 - \frac{\mu^M(\theta_M^D,\theta_m^D)(1-\theta_M^D)+\delta}{(1-\mu^M(\theta_M^D,\theta_m^D))(1-\theta_m^D)+\delta} + (1-\alpha^F)\frac{\partial \frac{\mu^M(\theta_M^D,\theta_m^D)(1-\theta_M^D)+\delta}{(1-\mu^M(\theta_M^D,\theta_m^D))(1-\theta_m^D)+\delta}}{\partial \alpha^F}\right) \\ =& \frac{\mu^M(\theta_M^D,\theta_m^D)(1-\theta_M^D)+\delta}{(1-\mu^M(\theta_M^D,\theta_m^D))(1-\theta_m^D)+\delta} - \alpha^F(1-\alpha^F)\frac{\partial \frac{\mu^M(\theta_M^D,\theta_m^D)(1-\theta_M^D)+\delta}{(1-\mu^M(\theta_M^D,\theta_m^D))(1-\theta_m^D)+\delta}}{\partial \alpha^F}\end{aligned}$$

By Lemma A.3, $\frac{\mu^M(\theta_M^D,\theta_m^D)(1-\theta_M^D)+\delta}{(1-\mu^M(\theta_M^D,\theta_m^D))(1-\theta_m^D)+\delta}$ is strictly decreasing with α^F so $\Delta > 0$. Hence, $\kappa^F(\cdot)$ is strictly increasing with α^F. $\square$

Observe that when $\alpha^F = \alpha^W$, then $\kappa^F(\theta_M^D, \theta_m^D) = \mu^M(\theta_M^D, \theta_m^D)$ (compare Equation A.9 and Equation A.17). We make use of the properties below.

The next corollary proves useful for the rest of the proofs.

Corollary A.1. *$\kappa^F(\theta_M^D, \theta_m^D)(1-\theta_M^D)$ is strictly decreasing with α^W and strictly increasing with α^F; $(1-\kappa^F(\theta_M^D, \theta_m^D))(1-\theta_m^D)$ is strictly increasing with α^W and strictly decreasing with α^F.*

Proof. Using Corollary 1 and Lemma A.9, we obtain that $\kappa^F(\theta_M^D, \theta_m^D)$ and $(1-\theta_M^D)$ have the same comparative statics with respect to $\alpha \in \{a^W, \alpha^F\}$ and similarly for $(1-\kappa^F(\theta_M^D, \theta_m^D))$ and $(1-\theta_m^D)$. The claim follows directly from Lemma A.9. □

We now prove that the probability that an unemployed majority worker is hired by a majority firm ($\kappa_F(1-\theta_M)$) is strictly higher than the probability that an unemployed minority worker is hired by a minority firm ($(1-\kappa^F)(1-\theta_m)$ if and only if the proportion of majority-owner firm is sufficiently large.

Lemma A.10. *For all $\alpha^F \in (1/2, 1)$, there exists a unique $\overline{\alpha}^W(\alpha^F) \in (1/2, 1)$ such that*

$$\kappa^F(\theta_M^D(\alpha^W,\alpha^F), \theta_m^D(\alpha^W,\alpha^F))(1-\theta_M^D(\alpha^W,\alpha^F)) \geq (1-\kappa^F(\theta_M^D(\alpha^W,\alpha^F), \theta_m^D(\alpha^W,\alpha^F)))(1-\theta_m^D(\alpha^W,\alpha^F))$$

if and only if $\alpha^W \leq \overline{\alpha}^W(\alpha^F)$.
The thresholds $\overline{\alpha}^W(\alpha^F)$ satisfies:
(i) $\overline{\alpha}^W(\alpha^F)$ is increasing with α^F (strictly if interior);
(ii) $\overline{\alpha}^W(1/2) = 1/2$;
(iii) There exists $\widehat{\alpha}^F < 1$ such that $\overline{\alpha}^W(\alpha^F) = 1$ for all $\alpha^F \geq \widehat{\alpha}^F$.

Proof. From Corollary A.1, we know that (ignoring arguments whenever possible): (a) $\kappa^F(\theta_M^D, \theta_m^D)(1-\theta_M^D)$ ($(1-\kappa^F(\theta_M^D, \theta_m^D))(1-\theta_m^D)$) is strictly decreasing (increasing) with α^W and (b) $\kappa^F(\theta_M^D, \theta_m^D)(1-\theta_M^D)$ ($(1-\kappa^F(\theta_M^D, \theta_m^D))(1-\theta_m^D)$) is strictly increasing (decreasing) with α^F. Given the definition of F_m^f (Equation A.8), we also have (c) for all α^W, $\lim_{\alpha^F \to 1} \kappa^F(\theta_M^D, \theta_m^D)(1-\theta_M^D) > \lim_{\alpha^F \to 1}(1-\kappa^F(\theta_M^D, \theta_m^D))(1-\theta_m^D) = 0$. Further, by symmetry and $\mu^M = \kappa^F$ when $\alpha^F = \alpha^W$, we obtain (d) if $\alpha^W = \alpha^F = 1/2$, then $\kappa^F(\theta_M^D, \theta_m^D)(1-\theta_M^D) = (1-\kappa^F(\theta_M^D, \theta_m^D))(1-\theta_m^D)$.

Property (d) together with property (a) implies that for all if $\alpha^F = 1/2$, for all $\alpha^W > 1/2$, then $\kappa^F(\theta_M^D(\alpha^W), \theta_m^D(\alpha^W))(1-\theta_M^D(\alpha^W)) < (1-\kappa^F(\theta_M^D(\alpha^W), \theta_m^D(\alpha^W)))(1-\theta_m^D(\alpha^W))$. This result coupled with properties (b) and (c) implies that there exists a unique $\underline{\alpha}^F(\alpha^W)$ such that for all, $\kappa^F(\theta_M^D(\alpha^W), \theta_m^D(\alpha^W))(1-\theta_M^D(\alpha^W)) \geq (1-\kappa^F(\theta_M^D(\alpha^W), \theta_m^D(\alpha^W)))(1-\theta_m^D(\alpha^W))$ if and only if $\alpha^F \geq \underline{\alpha}^F(\alpha^W)$.

Notice that by the Implicit Function Theorem and properties (a) and (b), $\underline{\alpha}^F(\alpha^W)$, if interior, is strictly increasing with α^W. Further, by property (c), $\lim_{a^W \to 1} \underline{\alpha}^F(\alpha^W) < 1$.

To finish the proof, define $\overline{\alpha}^W(\alpha^F) = \min\{(\underline{\alpha}^F)^{-1}(\alpha^F), 1\}$. The reasoning above proves that $\overline{\alpha}^W(\alpha^F)$ is unique for all α^F, that $\kappa^F(\theta_M^D(\alpha^W), \theta_m^D(\alpha^W))(1 - \theta_M^D(\alpha^W)) \geq (1 - \kappa^F(\theta_M^D(\alpha^W), \theta_m^D(\alpha^W)))(1 - \theta_m^D(\alpha^W))$ if and only if $\alpha^W \leq \overline{\alpha}^W(\alpha^F)$, that there exists $\widehat{\alpha}^F < 1$ such that $\overline{\alpha}^W(\alpha^F) = 1$ for all $\alpha^F \geq \widehat{\alpha}^F$. Finally, $\overline{\alpha}^W(1/2) = 1/2$ comes from properties (a) and (c) above. □

We now turn to findings conditions such that an employer never wants to hire a minority worker (Lemma A.11) and a worker never wants to work for a minority employer (Lemma A.12).

Lemma A.11. *For all $\alpha^F \in (1/2, 1)$, if $\alpha^W > \underline{\alpha}^W(\alpha^F)$, there exists $\overline{\beta}^F < 1$ such that for all $\beta \geq \overline{\beta}^F$, in a work discrimination assessment, when his position is open, a majority employer's best response is to never hire a minority worker.*

Proof. To prove the lemma, we introduce the following function $V_{J,K}^f(\theta; \neg K)$ which corresponds to the continuation value of a physical type J, social trait $K \in \{M, m\}$ employer with position filled by a worker of productivity θ and social trait $\neg K$. (Note that the continuation value we have used so far is implicitly $V_{J,K}^f(\theta; K)$, we have omitted the second argument for ease of exposition).

From the proof of Lemma 2, we know that $V_{M,M}^o = V_{M,M}^f(\theta_M^D)$ so that $V_{M,M}^o = (1 - \tau)\frac{\theta_M^D - w}{1 - \beta}$ (using Equation 3). By a similar reasoning, $V_{m,m}^o = \frac{\theta_m^D - w}{1 - \beta}$. Further, since once a majority employer is tainted, in a work discrimination assessment, he can only hire minority worker (i.e., he must behave like a minority employer), we obtain: $V_{M,m}^o = \frac{\theta_m^D - w}{1 - \beta}$.

With these preliminary results, we can now turn to the proof of the claim. We just need to look at one-shot deviation. Further, it can easily be checked that if a majority employer prefers not to hire a minority worker with productivity 1, he never hires a minority worker. We thus obtain the following continuation value if the majority employer deviates and hires a productivity-1 minority worker:

$$V_{M,M}^f(1; m) = 1 - w + \beta(1 - \delta)V_{M,M}^f(1; m) + \beta\delta\left[\rho V_{M,m}^o + (1 - \rho)V_{M,M}^o\right] \tag{A.18}$$

The employer gets $1 - w$ today by hiring the minority worker. If the match breaks (probability δ), his identity changes with probability ρ and his continuation value is similar to a minority employer. With the complementary probability, he is not tainted and his continuation value is the same as a majority employer.

A majority employer's best response is to not hire a productivity-1 minority worker if and only if:

$V^f_{M,M}(1;m) \leq V^o_{M,M}$. Using Equation A.18, this is equivalent to

$$1 - w + \beta\delta\rho(V^o_{M,m} - V^o_{M,M}) \leq (1-\beta)V^o_{M,M}$$

Plugging the values for the continuation value above and rearranging, the inequality is equivalent to:

$$1 - \theta^D_M \leq \frac{\beta}{1-\beta}\delta\rho\big(\theta^D_M - \theta^D_m\big) \tag{A.19}$$

Since the left-hand side is strictly positive, a necessary condition for this inequality to be satisfied is $\theta^D_M > \theta^D_m$. By Lemma A.8, this is equivalent to $\alpha^W > \underline{\alpha}^W(\alpha^F)$ for all $\alpha^F \in (1/2, 1)$. If $\alpha^W > \underline{\alpha}^W(\alpha^F)$ holds, then the right-hand side goes to ∞ as $\beta \to 1$, whereas the left-hand side is always bounded. Hence, there exists $\overline{\beta}^F$ such that the claim holds.[13] $\square$

Lemma A.12. *For all $\alpha^F \in (1/2, 1)$, if $\alpha^W < \overline{\alpha}^W(\alpha^F)$, there exists $\overline{\beta}^W < 1$ such that for all $\beta \geq \overline{\beta}^W$, in a work discrimination assessment, when unemployed, a majority worker's best response is to reject any minority firm's work offer.*

Proof. For all $J, K \in \{M, m\}^2$, denote a type-(J,K) worker's continuation value when employed with a social type $L \in \{M, m\}$ employer $W^e_{J,K}(L)$ and his continuation value when unemployed$W^u_{J,K}$.
In a work discrimination assessment, when playing the prescribed strategy (never works with a minority employer), a minority worker's continuation values are (again, we prove existence by construction and we ignore arguments for ease of exposition):

$$W^e_{M,M}(M) = (1-\tau)w + T(\tau) + \beta(1-\delta)W^e_{M,M}(M) + \beta\delta W^u_{M,M} \tag{A.20}$$

$$W^u_{M,M} = 0 + T(\tau) + \beta\kappa^F(1-\theta^D_M)W^e_{M,M}(M) + \beta\big(1 - \kappa^F(1-\theta^D_M)\big)W^u_{M,M} \tag{A.21}$$

From Equation A.20 and Equation A.21, we obtain:

$$W^e_{M,M}(M) = \frac{T(\tau)}{1-\beta} + \frac{(1-\beta) + \beta\kappa^F(1-\theta^D_M)}{1-\beta(1-\delta) + \beta\kappa^F(1-\theta^D_M)}\frac{(1-\tau)w}{1-\beta} \tag{A.22}$$

$$W^u_{M,M} = \frac{T(\tau)}{1-\beta} + \frac{\beta\kappa^F(1-\theta^D_M)}{1-\beta(1-\delta) + \beta\kappa^F(1-\theta^D_M)}\frac{(1-\tau)w}{1-\beta} \tag{A.23}$$

[13] Note that we cannot prove uniqueness since we do not know the comparative statics of the hiring thresholds with respect to β.

Similarly, recalling that in a work discrimination assessment she only gets hired by minority employer, a minority worker's continuation values are:

$$
\begin{aligned}
W^e_{m,m}(m) =& (1-\tau)w + T(\tau) + \beta(1-\delta)W^e_{m,m}(m) + \beta\delta W^u_{m,m} && \text{(A.24)}\\
W^u_{m,m} =& 0 + T(\tau) + \beta(1-\kappa^F)(1-\theta^D_m)W^e_{m,m}(m) + \beta\big(1-(1-\kappa^F)(1-\theta^D_m)\big)W^u_{m,m} && \text{(A.25)}
\end{aligned}
$$

After rearranging, this is equivalent to:

$$
\begin{aligned}
W^e_{m,m}(m) =& \frac{T(\tau)}{1-\beta} + \frac{(1-\beta)+\beta(1-\kappa^F)(1-\theta^D_m)}{1-\beta(1-\delta)+\beta(1-\kappa^F)(1-\theta^D_m)}\frac{(1-\tau)w}{1-\beta} && \text{(A.26)}\\
W^u_{m,m} =& \frac{T(\tau)}{1-\beta} + \frac{\beta(1-\kappa^F)(1-\theta^D_m)}{1-\beta(1-\delta)+\beta(1-\kappa^F)(1-\theta^D_m)}\frac{(1-\tau)w}{1-\beta} && \text{(A.27)}
\end{aligned}
$$

In turn, when a majority worker deviates and works with a minority employer, his continuation value is:

$$
W^e_{M,M}(m) = (1-\tau)w + T(\tau) + \beta(1-\delta)W^e_{M,M}(m) + \beta\delta\Big[\gamma W^u_{M,m} + (1-\gamma)W^u_{M,M}\Big] \tag{A.28}
$$

A majority worker becomes a socially minority worker with probability γ upon break of the relationship. Note that $W^u_{M,m} = W^u_{m,m}$ since a tainted majority worker (i.e., with majority physical type and minority social type) is never hired by a majority employer.

A majority worker's best response when unemployed and matched with a minority worker is to refuse a job offer if and only if: $W^e_{M,M}(m) \leq W^u_{M,M}$. Using Equation A.28, this is equivalent to:

$$
(1-\tau)w + T(\tau) + \beta\delta\gamma\big(W^u_{M,m} - W^u_{M,M}\big) \leq (1-\beta)W^u_{M,M}
$$

Using Equation A.23 and Equation A.27, this inequality is equivalent to after rearranging:

$$
\begin{aligned}
&\frac{1-\beta(1-\delta)}{1-\beta(1-\delta)+\beta(1-\kappa^F)(1-\theta^D_m)}\\
&\qquad \leq \frac{\beta}{1-\beta}\delta\gamma\left(\frac{\beta\kappa^F(1-\theta^D_M)}{1-\beta(1-\delta)+\beta\kappa^F(1-\theta^D_M)} - \frac{\beta(1-\kappa^F)(1-\theta^D_m)}{1-\beta(1-\delta)+\beta(1-\kappa^F)(1-\theta^D_m)}\right)
\end{aligned} \tag{A.29}
$$

The left-hand side is strictly positive and bounded. A necessary condition for (A.29) to hold is thus that the right-hand side is strictly positive. Notice that the function $H(x) = \frac{x}{1-\beta(1-\delta)+x}$ is strictly increasing with x. Consequently, the right-hand side of (A.29) is strictly positive if and only if $\kappa^F(1-\theta^D_M) > (1-\kappa^F)(1-\theta^D_m)$. Using Lemma A.10, this is the case for all α^F if and only if $\alpha^W < \overline{\alpha}^W(\alpha^F)$. When this condition is satisfied, the left-hand side goes to infinity as $\beta \to 1$ so there always exists $\overline{\beta}^W < 1$ such that (A.29) holds for all $\beta \geq \overline{\beta}^W$ as claimed. $\square$

We are now ready to prove the proposition.

Proof of Proposition 1

Using Lemmas A.8, A.10, A.11, and A.12, we know existence and uniqueness of the thresholds. There only remains to show that the interval $(\underline{\alpha}^W(\alpha^F), \overline{\alpha}^W(\alpha^F))$ is not empty for all $\alpha^F \in (1/2, 1)$. Once this is proved, we just need to define $\overline{\beta} = \max\{\overline{\beta}^F, \overline{\beta}^W\} < 1$ such that the claim holds.

Suppose that $\alpha^W < \overline{\alpha}^W(\alpha^F)$ so that $\kappa^F(1-\theta_M^D) > (1-\kappa^F)(1-\theta_m^D)$ (Lemma A.10). Recall that for $\alpha^F = \alpha^W$, $\mu^M = \kappa^F$. Further, comparing Equation A.9 and Equation A.16, it can easily be checked that $\mu^M > \kappa^F$ for all $\alpha^W > \alpha^F$. Hence, for all $\alpha^W \geq \alpha^F$, $\kappa^F(1-\theta_M^D) > (1-\kappa^F)(1-\theta_m^D)$ implies $\mu^M(1-\theta_M^D) > (1-\mu^M)(1-\theta_m^D)$. Using Equation A.14 and Equation A.15, this implies that $\theta_M^D > \theta_m^D$. Hence, $\alpha^W < \overline{\alpha}^W(\alpha^F) \;\Rightarrow\; \underline{\alpha}^W(\alpha^F) < \alpha^F$ (recall that θ_M^D is strictly increasing with α^W and θ_m^D strictly decreasing). Since, by construction, $\underline{\alpha}^W(\alpha^F)$ does not depend on α^W, this implies $\underline{\alpha}^W(\alpha^F) < \alpha^F$.

Suppose now that $\alpha^W > \underline{\alpha}^W(\alpha^F)$ so $\theta_M^D > \theta_m^D$. This implies that $\mu^M(1-\theta_M^D) > (1-\mu^M)(1-\theta_m^D)$. For all $\alpha^F \geq \alpha^W$, $\kappa^F \geq \mu^M$ so $\mu^M(1-\theta_M^D) > (1-\mu^M)(1-\theta_m^D) \Rightarrow \kappa^F(1-\theta_M^D) > (1-\kappa^F)(1-\theta_m^D)$. Using Lemmas A.9 and A.10, this implies that $\alpha^W > \underline{\alpha}^W(\alpha^F) \;\Rightarrow\; \overline{\alpha}^W(\alpha^F) > \alpha^F$. Since, by construction, $\overline{\alpha}^W(\alpha^F)$ does not depend on α^W, this implies $\overline{\alpha}^W(\alpha^F) > \alpha^F$.

Putting the two previous points together, we obtain that for all α^F, $\underline{\alpha}^W(\alpha^F) < \alpha^F < \overline{\alpha}^W(\alpha^F)$, which complete the claim. $\square$

Proof of Corollary 2

The claim has been proved in Lemma A.8 for $\underline{\alpha}^W(\alpha^F)$ and in Lemma A.10 for $\overline{\alpha}^W(\alpha^F)$. $\square$

Proof of Proposition 2

Point (i). Denote u_J the mass of workers with identity $J \in \{M, m\}$ who are unemployed. Given that majority workers only work with majority firm and minority workers with minority firm, in steady state, we have (ignoring arguments):

$$\begin{aligned}
u_M =& u_M(1-\kappa^F(1-\theta^D_M)) + \delta(\alpha^W - u_M) \\
u_m =& u_m(1-(1-\kappa^F)(1-\theta^D_m)) + \delta((1-\alpha^W) - u_m)
\end{aligned}$$

This is equivalent to:

$$\frac{u_M}{\alpha^W} = \frac{\delta}{\delta + \kappa^F(1-\theta^D_M)} \tag{A.30}$$

$$\frac{u_m}{1-\alpha^W} = \frac{\delta}{\delta + (1-\kappa^F)(1-\theta^D_m)} \tag{A.31}$$

The first term on the right-hand side (RHS) corresponds to unemployed workers who do not gain employment (either because they are matched with an employer of a different type or the match is not productive enough). The second corresponds to the mass of employed workers who lose employment. In steady state, the mass of unemployed workers remain constant.

By Lemma A.12, in the work discrimination equilibrium, $\kappa^F(1-\theta^D_M) > (1-\kappa^F)(1-\theta^D_m)$. Hence, we directly obtain that the unemployment rate of majority workers is strictly lower than the unemployment rate of minority workers from Equation A.30 and Equation A.31.

For the duration of unemployment spells, notice that the probability that a majority worker finds a job each period is $\kappa^F(1-\theta^D_M)$. For a minority worker, it is $(1-\kappa^F)(1-\theta^D_m)$. Again, the claim holds using Lemma A.12.

Point (ii). The average productivity of a majority firm producing is $\frac{1+\theta^D_M}{2} > \frac{1+\theta^D_m}{2}$. By Lemma A.11, in the work discrimination equilibrium $\theta^D_M > \theta^D_m$ so point (ii) holds. □

Robustness to change in wage offers

In this subsection, we discuss the robustness of our results to some changes in the hiring rule. More specifically, we assume here that (1) a minority worker (and only a minority worker) can propose to work for free for a majority firm and (2) a minority firm can offer its whole profit to a majority worker to entice him to enter an employment relationship. We call this setting the 'amended labor market.' We obtain that this change to the labor market has no effect on the necessary proportion of majority workers for discrimination strategies to be mutual best response.

To understand this result, note that our amended labor market only changes the present benefit of deviating. The future cost of hiring a minority worker (for majority employers) and the future cost of working for a minority firm (for majority workers) remain unchanged in this setting. Both actors still expect to be ostracized by the relevant majority types if they become tainted in the future. As long as these costs are positive, when α^W is intermediate, we can always find a discount factor so that the future cost is so high that it overweights all present benefits. Because our amended labor market increases the present benefits, the discount factor needs to be higher than in the baseline model. However, the basic logic of the model holds in this amended labor market. Proposition 1 holds among unchanged.

Proposition A.1. *In the amended labor market, for all proportions of employers with majority identity* $\alpha^F \in (1/2, 1)$*, there exists two unique population thresholds* $\underline{\alpha}^W(\alpha^F), \overline{\alpha}^W(\alpha^F) \in (1/2, \alpha^F) \times (\alpha^F, 1]$ *such that if the proportion of workers with majority identity satisfies* $\alpha^W \in (\underline{\alpha}^W(\alpha^F), \overline{\alpha}^W(\alpha^F))$*, there exists* $\overline{\beta}^a \in (\overline{\beta}, 1)$ *such that for all* $\beta \geq \overline{\beta}^a$*, workers' and employers' discrimination strategies are mutual best response.*

Proof. In the amended labor market, the same logic as in the baseline applies on-path when other actors play a discrimination strategy. Hence, all results up to Lemma A.11 hold.

Now regarding Lemma A.11, the continuation values on-path are still Equation 3-Equation 5. The key difference now is the value from hiring a minority who is willing to work for zero salary.[14] That is, the continuation value from deviation is (again we only need to check for the highest productivity case):

$$V^f_{M,M}(1;m) = 1 + \beta(1-\delta)V^f_{M,M}(1;m) + \beta\delta\Big[\rho V^o_{M,m} + (1-\rho)V^o_{M,M}\Big]$$

Again, discrimination is a best response if and only if $V^f_{M,M}(1;m) \leq V^o_{M,M}$. Similar algebra as in the proof of Lemma A.11 yields that this condition is equivalent to:

$$\begin{aligned} & 1 + \beta\delta\rho(V^o_{M,m} - V^o_{M,M}) \leq (1-\beta)V^o_{M,M} \\ \Leftrightarrow \quad & 1 + w - \theta^D_M \leq \frac{\beta}{1-\beta}\delta\rho\big(\theta^D_M - \theta^D_m\big) \end{aligned}$$

A necessary condition is still that $\theta^D_M > \theta^D_m$ or, equivalently, $\alpha^W > \underline{\alpha}^W(\alpha^F)$. When this holds, we can define $\overline{\beta}^{Fa} < 1$ such that the condition holds for $\beta \geq \overline{\beta}^{Fa}$. Note that since the condition above is more stringent than Equation A.19, we have $\overline{\beta}^{Fa} > \overline{\beta}^F$, with $\overline{\beta}^F$ defined in Lemma A.11.

[14]Note that we do not check that the minority worker would accept a zero salary offer. However, this is inconsequential for the proof.

Regarding Lemma A.12, the continuation values on-path are still Equation A.22-Equation A.25. The key difference now is the value from working with a minority firm who is willing to offer a salary of θ.[15] That is, the continuation value from deviation is (we only need to check for the highest productivity case):

$$W^e_{M,M}(m) = (1-\tau) + T(\tau) + \beta(1-\delta)W^e_{M,M}(m) + \beta\delta\Big[\gamma W^u_{M,m} + (1-\gamma)W^u_{M,M}\Big]$$

Again, discrimination is a best response if and only if $W^e_{M,M}(m) \leq W^u_{M,M}$. Similar algebra as in the proof of Lemma A.12 yields that this condition is equivalent to:

$$\begin{aligned}
& (1-\tau) + T(\tau) + \beta\delta\gamma\big(W^u_{M,m} - W^u_{M,M}\big) \leq (1-\beta)W^u_{M,M} \\
\Leftrightarrow \quad & \frac{1}{w} - \frac{\beta(1-\kappa^F)(1-\theta^D_m)}{1-\beta(1-\delta)+\beta(1-\kappa^F)(1-\theta^D_m)} \\
& \qquad \leq \frac{\beta}{1-\beta}\delta\gamma\left(\frac{\beta\kappa^F(1-\theta^D_M)}{1-\beta(1-\delta)+\beta\kappa^F(1-\theta^D_M)} - \frac{\beta(1-\kappa^F)(1-\theta^D_m)}{1-\beta(1-\delta)+\beta(1-\kappa^F)(1-\theta^D_m)}\right)
\end{aligned}$$

The necessary condition, thus, remains the same as in Lemma A.12: $\alpha^W < \overline{\alpha}^W(\alpha^F)$. When this holds, we can define $\overline{\beta}^{Wa} < 1$ such that the condition holds for $\beta \geq \overline{\beta}^{Wa}$. Note that since the condition above is more stringent than Equation A.29, we have $\overline{\beta}^{Wa} > \overline{\beta}^F$, with $\overline{\beta}^W$ defined in Lemma A.12.

We can then use the reasoning in the proof of Proposition 1 to prove the claim. □

B Proof of Section 5

Before proving Proposition 3, we first compare some important functions in a discrimination assessment and in an assessment without discrimination. We first determine the tax revenue when in the labour market subgame, the equilibrium exhibits work discrimination ($R^D(\tau)$) or does not exhibit work discrimination ($R^{ND}(\tau)$) as a function of the tax rate τ.

[15] Note again that we do not check that the minority employer would be willing to offer a salary of one. However, this is inconsequential for the proof.

Lemma B.1. *For all $\tau \in [0,1]$, the tax revenue with and without work discrimination satisfies, respectively:*

$$R^D(\tau) = \tau \times \left(\alpha^F \frac{\mu^M(1-\theta_M^D)}{\mu^M(1-\theta_M^D)+\delta} \frac{1+\theta_M^D}{2} + (1-\alpha^F) \frac{(1-\mu^M)(1-\theta_m^D)}{(1-\mu^M)(1-\theta_m^D)+\delta} \frac{1+\theta_m^D}{2} \right) \tag{B.1}$$

$$R^{ND}(\tau) = \tau \times \left(\frac{1-\theta^{ND}}{(1-\theta^{ND})+\delta} \frac{1+\theta^{ND}}{2} \right) \tag{B.2}$$

Proof. In a work discrimination, the average production of a majority firm that produces is:

$$Pr(Maj\ employer) \times Pr(Produces|Maj\ employer) \times E(\theta|Produces,\ Maj\ employer),$$

with $Pr(Maj\ employer) = \alpha^F$, $Pr(Produces|Maj\ employer) = \frac{F_M^f}{\alpha^F} = \frac{\mu^M(1-\theta_M^D)}{\mu^M(1-\theta_M^D)+\delta}$ from Equation A.7, and $E(\theta|Produces,\ Maj\ employer) = E(\theta|\theta \geq \theta_M^D) = \frac{1+\theta_M^D}{2}$. Using the fact that we have a mass of employers, we have that the total production by majority employers is: $\alpha^F \frac{\mu^M(1-\theta_M^D)}{\mu^M(1-\theta_M^D)+\delta} \frac{1+\theta_M^D}{2}$. Given both workers and employers are taxed, this represents the total revenues available from taxation from pair of worker and employer with majority social identity. A similar reasoning for the minority yields $R^D(\alpha)$. In turn, using a similar reasoning as in Lemma A.3, it can be checked that the mass and proportion of firms with filled positions absent work discrimination is $\frac{1-\theta^D}{1-\theta^D+\delta}$. Hence, the total revenues available for taxation are then: $\frac{1-\theta^{ND}}{(1-\theta^{ND})+\delta} \frac{1+\theta^{ND}}{2}$ as claimed. □

Corollary B.1. *For all $\tau \in (0,1]$, $R^D(\tau) < R^{ND}(\tau)$.*

Proof. We can rewrite $(R^D(\tau) - R^{ND}(\tau))/\tau$ as

$$\alpha^F \left(\frac{\mu^M(1-\theta_M^D)}{\mu^M(1-\theta_M^D)+\delta} \frac{1+\theta_M^D}{2} - \frac{1-\theta^{ND}}{(1-\theta^{ND})+\delta} \frac{1+\theta^{ND}}{2} \right)$$
$$+ (1-\alpha^F) \left(\frac{(1-\mu^M)(1-\theta_m^D)}{(1-\mu^M)(1-\theta_m^D)+\delta} \frac{1+\theta_m^D}{2} - \frac{1-\theta^{ND}}{(1-\theta^{ND})+\delta} \frac{1+\theta^{ND}}{2} \right)$$

From Corollary 1, we know that θ_M^D is strictly increasing with α^W and $\theta_M^D \xrightarrow{\alpha^W \to 1} \theta^{ND}$. Hence, $\theta_M^D < \theta^{ND}$. Further, using Equation A.14 and Equation A.4, this implies that $\mu^M(1-\theta_M^D) < 1 - \theta^{ND}$. Hence, the parenthesis on the first line is negative. Further, in a work discrimination equilibrium, $\theta_M^D > \theta_m^D$ (see Lemma A.11) which yields $\mu^M(1-\theta_M^D) > (1-\mu^M)(1-\theta_m^D)$ (from Equation A.14 and Equation A.15) so the parenthesis in the second line is also negative. Hence $R^D(\tau) < R^{ND}(\tau)$. □

Denote the transfers: $T^D(\tau) = K\big(R^D(\tau)\big)$ with discrimination and $T^{ND}(\tau) = K\big(R^{ND}(\tau)\big)$ without. Denote as well $\mathcal{W}_J^D(\tau)$ and $\mathcal{W}_J^{ND}(\tau)$ the expected welfare of a worker with social identity $J \in \{M, m\}$ when the labour market equilibrim exhibits discrimination (subscript D) or does not exhibit discrimination (subscript ND) and the level of taxation is τ. Recall that we assume that the labour market adjusts instantly.

Lemma B.2. *For all $\tau \in [0,1]$, the expected welfare of a majority worker with and without work discrimination is, respectively:*

$$\mathcal{W}_M^D(\tau) = (1-\tau)\frac{w}{1-\beta} \times \frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta} + \frac{T^D(\tau)}{1-\beta} \tag{B.3}$$

$$\mathcal{W}_M^{ND}(\tau) = (1-\tau)\frac{w}{1-\beta} \times \frac{(1-\theta^{ND})}{(1-\theta^{ND})+\delta} + \frac{T^{ND}(\tau)}{1-\beta} \tag{B.4}$$

Proof. From an ex-ante perspective (i.e., before the labour market starts), a majority worker has a probability $\frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta}$ to be employed. Hence, his expected welfare in a work discrimination equilibrium is:

$$\mathcal{W}_M^D(\tau) = \frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta} \times W_{M,M}^e(M) + \frac{\delta}{\kappa^F(1-\theta_M^D)+\delta} W_{M,M}^u$$

Using Equation A.22 and Equation A.23, we obtain:

$$\begin{aligned}
\mathcal{W}_M^D(\tau) &= (1-\tau)\frac{w}{1-\beta}\Bigg[\frac{(1-\beta)+\beta\kappa^F(1-\theta_M^D)}{1-\beta(1-\delta)+\beta\kappa^F(1-\theta_M^D)} \times \frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta} \\
&\qquad + \frac{\beta\kappa^F(1-\theta_M^D)}{1-\beta(1-\delta)+\beta\kappa^F(1-\theta_M^D)} \times \frac{\delta}{\kappa^F(1-\theta_M^D)+\delta}\Bigg] + \frac{T^D(\tau)}{1-\beta} \\
&= (1-\tau)\frac{w}{1-\beta}\frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta}\frac{(1-\beta)+\beta\kappa^F(1-\theta_M^D)+\beta\delta}{1-\beta(1-\delta)+\beta\kappa^F(1-\theta_M^D)} + \frac{T^D(\tau)}{1-\beta} \\
&= (1-\tau)\frac{w}{1-\beta}\frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta} + \frac{T^D(\tau)}{1-\beta}
\end{aligned}$$

A similar reasoning holds for an equilibrium without work discrimination noting that the probability a worker is employed is then $\frac{1-\theta^{ND}}{1-\theta^{ND}+\delta}$. □

Lemma B.3. *For all $\alpha^F \in (1/2, 1)$, there exists a unique $\widehat{\alpha}^W(\alpha^F) \in [1/2, \alpha^F)$ such that if $\alpha^W \in (1/2, \widehat{\alpha}^W(\alpha^F))$ (possibly an empty interval), then $\frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta} > \frac{(1-\theta^{ND})}{(1-\theta^{ND})+\delta}$.*
The threshold $\widehat{\alpha}^W(\alpha^F)$ is strictly increasing with α^F and satisfies $\lim_{\alpha^F \to 1} \widehat{\alpha}^W(\alpha^F) = 1$.

Proof. From Corollary A.1, we know that $\kappa^F(1-\theta_M^D)$ is strictly increasing with α^F. Recall that at $\alpha^W = \alpha^F$, $\kappa^F = \mu^M$. From Corollary B.1, this implies that $\kappa^F(1-\theta_M^D) < 1-\theta^{ND}$ and so $\frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta} < \frac{(1-\theta^{ND})}{(1-\theta^{ND})+\delta}$. Notice now that for all $\alpha^W \in (1/2,1)$, $\lim_{\alpha^F\to 1} \kappa^F(1-\theta_M^D) = \lim_{\alpha^F\to 1} 1-\theta_M^D > 1-\theta^{ND}$ (see Corollary B.1) so $\lim_{\alpha^F\to 1} \frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta} > \frac{(1-\theta^{ND})}{(1-\theta^{ND})+\delta}$. All results together imply that for all $\alpha^W \in (1/2,1)$, there exists a unique $\underline{\underline{\alpha}}^F(\alpha^W) \in (\alpha^W, 1)$ such that $\frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta} > \frac{(1-\theta^{ND})}{(1-\theta^{ND})+\delta}$ if and only if $\alpha^F > \underline{\underline{\alpha}}^F(\alpha^W)$. Notice further that since $\kappa^F(1-\theta_M^D)$ is strictly decreasing with α^W (Corollary A.1) and θ^{ND} obviously does not depend on α^W, it must be that $\underline{\underline{\alpha}}^F(\underline{\alpha}^W)$ is strictly increasing with α^W. Finally note that (slightly abusing notation) since $\theta_M^D < \theta^{ND}$ for all $\alpha^W < 1$ and $\lim_{\alpha^W\to 1} \theta_M^D = \theta^{ND}$ (Corollary 1), then $\underline{\underline{\alpha}}^F(\alpha^W) < 1$ for all $\alpha^W < 1$ and $\lim_{\alpha^W\to 1} \underline{\underline{\alpha}}^F(\alpha^W) = 1$.
Defining $\widehat{\alpha}^W(\cdot) = (\underline{\underline{\alpha}}^F)^{-1}(\cdot)$ yields the result. In particular, $\underline{\underline{\alpha}}^F(\alpha^W) > \alpha^W$ implies $\widehat{\alpha}^W(\alpha^F) < \alpha^F$. $\square$

For the next proposition, it is useful to denote

$$
\begin{aligned}
A^D =& \alpha^F \frac{\mu^M(1-\theta_M^D)}{\mu^M(1-\theta_M^D)+\delta}\frac{1+\theta_M^D}{2} + (1-\alpha^F)\frac{(1-\mu^M)(1-\theta_m^D)}{(1-\mu^M)(1-\theta_m^D)+\delta}\frac{1+\theta_m^D}{2} \\
B^D =& \frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta} w \\
\sigma^D =& \frac{B^D}{A^D} \\
A^{ND} =& \frac{1-\theta^{ND}}{(1-\theta^{ND})+\delta}\frac{1+\theta^{ND}}{2} \\
B^{ND} =& \frac{(1-\theta^{ND})}{(1-\theta^{ND})+\delta} w \\
\sigma^{ND} =& \frac{B^{ND}}{A^{ND}}
\end{aligned}
$$

Proof of Proposition 3

In what follows, we assume that the tax rate is always strictly positive. The proof can easily be extended to the case when it is zero under work discrimination or both with and without work discrimination. Absent work discrimination, all workers are identical and they form a majority of the citizenry ($f < 1$). Hence, the tax rate, denoted τ^{ND} maximizes $\mathcal{W}_M^{ND}(\tau)$. Hence, the tax rate satisfies using the notation above:

$$A^{ND} K'\big(A^{ND}\tau^{ND}\big) \geq B^{ND} \tag{B.5}$$

Define $k(r) = (K')^{-1}(r)$. We obtain that

$$\tau^{ND} = \min\left\{\frac{1}{A^{ND}}k(\sigma^{ND}), 1\right\} \tag{B.6}$$

If the social discriminatory policy is implemented and a work discrimination labour equilibrium follows, we claim and prove in the proof of Proposition 4 that the tax rate preferred by majority workers is implemented. Hence, in this case, we obtain that the tax rate, denoted τ_M^D, satisfies:

$$A^D K'\big(A^D \tau^D\big) \geq B^D \tag{B.7}$$

Or equivalently,

$$\tau^D = \min\left\{\frac{1}{A^D}k(\sigma^D), 1\right\} \tag{B.8}$$

Denote now $\mathcal{W}_M^D := \mathcal{W}_M^D(\tau_M^D)$ a majority worker's expected equilibrium welfare with work discrimination. Similarly denote $\mathcal{W}_M^{ND} := \mathcal{W}_M^{ND}(\tau^{ND})$ a majority worker's expected equilibrium welfare without work discrimination. We obtain:

$$\begin{aligned}
\mathcal{W}_M^D =& (1-\tau_M^D)\frac{w}{1-\beta} \times \frac{\kappa^F(1-\theta_M^D)}{\kappa^F(1-\theta_M^D)+\delta} + \frac{T^D(\tau_M^D)}{1-\beta} \\
\mathcal{W}_M^{ND} =& (1-\tau^{ND})\frac{w}{1-\beta} \times \frac{(1-\theta^{ND})}{(1-\theta^{ND})+\delta} + \frac{T^{ND}(\tau^{ND})}{1-\beta}
\end{aligned}$$

From Corollary B.1, we know that $T^D(\tau) < T^{ND}(\tau)$. Hence, if $\tau_M^D = 1$, we necessarily have $\mathcal{W}_M^D < \mathcal{W}_M^{ND}$. Indeed, for all $\tau^{ND} \leq 1$, then $\mathcal{W}_M^{ND}(\tau^{ND}) \geq \mathcal{W}_M^{ND}(1) = \frac{T^{ND}(1)}{1-\beta} > \frac{T^D(1)}{1-\beta} = \mathcal{W}_M^D(\tau_M^D)$. Since the function $K(\cdot)$ is strictly concave, there exists a unique K_1 such that $\tau_M^D < 1$ for all $K'(0) < K_1$.
Using Equation B.6 and Equation B.8 and the notation above, we obtain:

$$(1-\beta)\mathcal{W}_M^D = B^D - \sigma^D k(\sigma^D) + K\big(k(\sigma^D)\big) \tag{B.9}$$

$$(1-\beta)\mathcal{W}_M^{ND} = B^{ND} - \sigma^{ND} k(\sigma^{ND}) + K\big(k(\sigma^{ND})\big) \tag{B.10}$$

It proves useful to define $Z(\sigma) = \sigma k(\sigma) - K\big(k(\sigma)\big)$. Notice that given $k(r) = (K')^{-1}(r)$, $Z'(\sigma) = k(\sigma) > 0$.
We first show that $B^D > B^{ND}$ is a necessary condition for $\mathcal{W}_M^D > \mathcal{W}_M^{ND}$. Suppose $B^D \leq B^{ND}$. The Envelop Theorem then yields $\frac{d\mathcal{W}^{ND}}{dB^D} = (1-\tau_M^D) > 0$. So $\mathcal{W}_M^D \leq B^{ND} - Z\big(\frac{B^{ND}}{A^D}\big)$. From Corollary B.1, $A^D < A^{ND}$. From above, $-Z(r)$ is strictly decreasing so $\mathcal{W}_M^D \leq B^{ND} - Z\big(\frac{B^{ND}}{A^D}\big) < B^{ND} - Z\big(\frac{B^{ND}}{A^{ND}}\big) =$

$\mathcal{W}_M^{ND}$.

The demand for discrimination is thus positive only if $B^D > B^{ND}$. Using Lemma B.3, a necessary condition is thus that $\alpha^W \in (1/2, \widehat{\alpha}^W(\alpha^F))$.

We now show a sufficient condition on $K'(0)$. Supposing $\alpha^W \in (1/2, \widehat{\alpha}^W(\alpha^F))$ holds, $\mathcal{W}_M^D > \mathcal{W}_M^{ND}$ is equivalent to:

$$B^D - B^{ND} > Z(\sigma^D) - Z(\sigma^{ND}),$$

with $\sigma^D > \sigma^{ND}$. From above, $Z'(\sigma) = k(\sigma)$. Given the definition of $k(\cdot)$ ($k(\cdot) = (K')^{-1}(\cdot)$), $k(\cdot)$ is strictly decreasing ($K(\cdot)$ is strictly concave) so $Z'(\sigma) < k(0)$. Hence, a sufficient condition for positive demand for discrimination is:

$$B^D - B^{ND} > k(0)(\sigma^D - \sigma^{ND})$$

This means that there exists $\overline{k} > 0$ such that if $k(0) < \overline{k}$, then the demand for discrimination is strictly positive. Note that $k(0) < \overline{k} \Leftrightarrow K'(k(0)) = 0 > K'(\overline{k}) \Leftrightarrow K'(0) < K'(K'(\overline{k}))$. Thus define, $K_0 = K'(K'(\overline{k}))$. To finish the claim, define $\overline{K} = \min\{K_1, K_0\} > 0$. □

Before proving Proposition 4, we first establish that minority workers do not demand the symbolic policy when majority workers do (necessary condition $\alpha^W < \alpha^F$).

Lemma B.4. *Suppose $\alpha^W < \alpha^F$, then minority workers' expected welfare is strictly higher under $(d, \tau) = (0, \tau^{ND})$ than under $(1, \tau)$ for all $\tau \in [0, 1]$.*

Proof. By the exact same reasoning as in Lemma B.2, we obtain that minority workers' expected welfare with and without work discrimination are, respectively:

$$\mathcal{W}_m^D(\tau) = (1-\tau)\frac{w}{1-\beta} \times \frac{(1-\kappa^F)(1-\theta_m^D)}{(1-\kappa^F)(1-\theta_m^D)+\delta} + \frac{T^D(\tau)}{1-\beta} \tag{B.11}$$

$$\mathcal{W}_m^{ND}(\tau) = (1-\tau)\frac{w}{1-\beta} \times \frac{(1-\theta^{ND})}{(1-\theta^{ND})+\delta} + \frac{T^{ND}(\tau)}{1-\beta} \tag{B.12}$$

Denote

$$\begin{aligned}
A^D =& \alpha^F \frac{\mu^M(1-\theta_M^D)}{\mu^M(1-\theta_M^D)+\delta}\frac{1+\theta_M^D}{2} + (1-\alpha^F)\frac{(1-\mu^M)(1-\theta_m^D)}{(1-\mu^M)(1-\theta_m^D)+\delta}\frac{1+\theta_m^D}{2} \\
B_m^D =& \frac{(1-\kappa^F)(1-\theta_m^D)}{(1-\kappa^F)(1-\theta_m^D)+\delta} w \\
\sigma_m^D =& \frac{B_m^D}{A^D}
\end{aligned}$$

Minority workers' preferred tax rate with work discrimination, denoted τ_m^D then satisfies using a similar reasoning as in the proof of Proposition 3:

$$\tau_m^D = \min\left\{\frac{1}{A^D}k(\sigma_m^D), 1\right\}$$

From Proposition 2, $B_m^D < B_M^D$ so $\tau_m^D \geq \tau_M^D$ with strict inequality whenever $\tau_M^D \in (0,1)$.
To prove the result, it is sufficient to show that $\mathcal{W}_m^D := \mathcal{W}_m^D(\tau_m^D) < \mathcal{W}_m^{ND} = \mathcal{W}_m^{ND}(\tau^{ND})$. If $\tau_m^D = 1$, a similar reasoning as in the proof of Proposition 3 yields that $\mathcal{W}_m^D < \mathcal{W}_m^{ND}$. If $\tau_m^D < 1$, we just need to show that $B_m^D \leq B^{ND}$ which is equivalent to show that $(1-\kappa^F)(1-\theta_m^D) \leq (1-\theta^{ND})$. Recall that $\mu^M = \kappa^F$ at $\alpha^W = \alpha^F$ and $\mu^M < \kappa^F$ for all $\alpha^W < \alpha^F$ (Lemma A.8 and Corollary A.1). Hence, for all $\alpha^W > \alpha^F$, $(1-\mu^M)(1-\theta_m^D) > (1-\kappa^F)(1-\theta_m^D)$. We know that discrimination strategies are mutual best response only if $(1-\mu^M)(1-\theta_m^D) < \mu^M(1-\theta_M^D)$. Further, $\mu^M(1-\theta_M^D) < 1-\theta^{ND}$ for all $\alpha^W \in (1/2, 1)$ (Lemma B.3). Putting all results together, $(1-\kappa^F)(1-\theta_m^D) \leq (1-\theta^{ND})$ so $\mathcal{W}_m^D < \mathcal{W}_m^{ND}$. □

We can now turn to candidates' behavior.

Proof of Lemma 3

Candidates are office-motivated and so always propose the platform preferred by a majority of voters. Recall that from our assumption on indifference, if a majority workers prefer the symbolic policy, all other citizens prefer $d = 0$ and $\alpha^W = \frac{1+f}{2}$ (so votes would be split), candidates propose $d = 0$ then.
For the second part of the lemma, observe that the preferred tax rate of employers is zero since they do not receive any transfer. We show that no candidate offers $(1, \tau)$ with $\tau \neq \tau_M^D$ in equilibrium. Suppose that candidate $-J$ proposes $(1, \tau_{-J})$ with $\tau_{-J} \neq \tau_M^D$. J can then win the election for sure by offering $(1, \tau_M^D)$. Indeed, if $\tau_{-J} > \tau_M^D$, both majority and minority employers vote for J and so do majority workers. If $\tau_{-J} < \tau_M^D$, then all workers vote for J since the minority's preferred tax rate under social

discrimination τ_m^D satisfies $\tau_m^D > \tau_M^D$ (see Lemma B.4). Hence, by the usual reasoning, no candidate offers platform $(1, \tau)$ with $\tau \neq \tau_M^D$.

A similar reasoning yields that $(0, \tau)$ with $\tau \neq \tau^{ND}$ would be defeated by $(0, \tau^{ND})$. Hence, in equilibrium, candidates propose one of the two platforms detailed in the text of the Lemma. □

Proof of Proposition 4

Using Lemma 3, two conditions need to be satisfied for the symbolic policy to be proposed and implemented in equilibrium:

(i) if identity is salient, then discrimination strategies are mutual best response (remember, citizens oppose the symbolic policy when indifferent);

(ii) a plurality of the citizenry prefers $(d, \tau) = (1, \tau_M^D)$ to $(0, \tau^{ND})$.

Necessary conditions for (i) to hold are for all α^F, $\alpha^W \in (\underline{\alpha}^W(\alpha^F), \overline{\alpha}(\alpha^F))$ with the thresholds derived in Proposition 1. Condition (ii) requires that (a) $\alpha^W < \widehat{\alpha}^W(\alpha^F)$ for all α^F (among others) with the threshold derived in Proposition 3 hold and (b) majority workers and their potential allies form a majority of the population. Denote α^{pl} the proportion of majority workers necessary such that (b) holds if and only if $\alpha^W > \alpha^{pl}$. If majority employers strictly prefer $(1, \tau_M^D)$ to $(0, \tau^{ND})$ then $\alpha^{pl} = 1/2$.[16] Otherwise, $\alpha^{pl} = \frac{1+f}{2}$.

We now show that the set $(\underline{\alpha}^W(\alpha^F), \overline{\alpha}(\alpha^F)) \cap (1/2, \widehat{\alpha}^W(\alpha^F)) \cap (\alpha^{pl}, 1)$ is non empty. First, note from Propositions 1 and 3 that $\widehat{\alpha}(\alpha^F) < \alpha^F < \overline{\alpha}(\alpha^F)$. Hence, we need to show that there exists $\alpha^F \in (1/2, 1)$ such that $\widehat{\alpha}^W(\alpha^F) > \max\{\alpha^{pl}, \underline{\alpha}^W(\alpha^F)\}$. From Corollary 2, recall that $\lim_{\alpha^F \to 1} \underline{\alpha}^W(\alpha^F) < 1$. From Lemma B.3, $\lim_{\alpha^F \to 1} \widehat{\alpha}^W(\alpha^F) = 1$. Since both thresholds are continuous and $\alpha^{pl} < 1$, there exists $\underline{\alpha}^F \in (1/2, 1)$ such that for all $\alpha^F > \underline{\alpha}^F$, then $\widehat{\alpha}^W(\alpha^F) > \max\{\alpha^{pl}, \underline{\alpha}^W(\alpha^F)\}$. (Note that there can be other sets where all conditions are satisfied.)

Putting all results together, for all α^F and α^W such that $\alpha^F > \underline{\alpha}^F$, $\alpha^W \in (\min\{a^{pl}, \underline{\alpha}^W(\alpha^F)\}, \widehat{\alpha}^W(\alpha^F))$, $\beta > \overline{\beta}$, and $K'(0) < \overline{K}$, then: (i) when social identity is salient, discrimination strategies are mutual best response, (ii) a plurality of the citizenry strictly prefer $(1, \tau_M^D)$ to $(0, \tau^{ND})$. Since candidates are office-motivated, both candidates converge to $(1, \tau_M^D)$ then and the unique equilibrium is the discrimination equilibrium. □

[16] By Proposition 2, majority employers are always more sympathetic to the symbolic policy than minority employers. Due to the multiple moving parts, we are not able to determine conditions such that majority employers prefer $(1, \tau_M^D)$ to $(0, \tau^{ND})$.

Proof of Proposition 5

As discussed in the main text, all non-discriminatory equilibria are similar (at least in term of comparative statics with respect to α^W and α^F).

Point (i). This follows directly from Corollary B.1.

Point (ii). This follows directly from Lemma B.3.

Point (ii'). The result is proved in the proof of Lemma B.4 given that social discrimination to occur requires $\alpha^W < \widehat{\alpha}^W(\alpha^F) < \alpha^F$.

Point (iii). A discrimination equilibrium exists only if $\tau_M^D < 1$. Further, using Equation B.5, the tax rate is zero under no discrimination if and only if $K'(0) \leq \frac{B^{ND}}{A^{ND}}$. Similarly, using Equation B.7, the tax rate is zero under discrimination if and only if $K'(0) \leq \frac{B^D}{A^D}$. From Lemma B.1, $A^{ND} > A^D$. From the proof of Proposition 3, for discrimination to occur, then $B^{ND} < B^D$. Hence, $\tau^{ND} = 0 \Rightarrow \tau^D = 0$ and transfers are equal then. Suppose $\tau_M^D = 0 < \tau^{ND}$, then transfers are strictly lower under discrimination. Suppose finally that $\tau_M^D > 0$. Then total transfers are $T^{ND} = K(A^{ND}\tau^{ND})$ under no discrimination and they are determined by $K'(A^{ND}\tau^{ND}) = \frac{B^{ND}}{A^{ND}}$. Total transfers are $T^D = K(A^D\tau^D)$ under discrimination and they are determined by $K'(A^D\tau^D) = \frac{B^D}{A^D}$. Since $\frac{B^{ND}}{A^{ND}} < \frac{B^{ND}}{A^{ND}}$ and $K'(\cdot)$ is strictly decreasing, we obtain $A^D\tau^D < A^{ND}\tau^{ND}$. Since $K(\cdot)$ is strictly increasing, this implies $T^D < T^{ND}$.

Point (iv). The result follows directly when $\tau^{ND} = 0$, $\tau^{ND} = 1$ (since a discrimination equilibrium implies $\tau_M^D < 1$ by Proposition 3), or $\tau^{ND} > 0$ and $\tau_M^D = 0$. Suppose thus that $\tau_M^D > 0$ and $\tau^{ND} < 1$. Recall that, in this case, the tax rate is the solution to (ignoring superscripts): $AK(A\tau) = B$.

To prove the result, it is useful to define the function $E(A) = AK'(A\tau)$. $E'(A) = K'(A\tau) + \tau AK''(A\tau) = K'(A\tau) \times \left(1 + \frac{\tau A}{K'(\tau A)} \frac{\partial K'(A\tau)}{\partial A\tau}\right)$. Under the assumption that $\varepsilon_{K',r} = \frac{r}{K'(r)} \frac{\partial K'(r)}{\partial r} > -1$, note that $K'(A\tau)\left(1 + \frac{A\tau}{K'(A\tau)} \frac{\partial K'(A\tau)}{\partial A\tau}\right) > 0$ and the function $E(A)$ satisfies $E'(A) > 0$.

Suppose now, by way of contradiction, that $\tau_M^D \geq \tau^{ND}$. This implies that $A^D K'(A^D\tau_M^D) \leq A^D K'(A^D\tau^{ND})$ since $K(\cdot)$ is strictly concave. We know that $A^D < A^{ND}$ (Corollary B.1), so, under the assumption, $A^D K'(A^D\tau^{ND}) < A^{ND}K'(A^{ND}\tau^{ND})$. This implies, using equation determining the tax rate that, $B^D = A^D K'(A^D\tau_M^D) < A^{ND}K'(A^{ND}\tau^{ND}) = B^{ND}$. However, by Proposition 3, in the discrimination equilibrium (specifically, for parameter values such that the equilibrium is the discrimination equilibrium), $B^D > B^{ND}$. Hence, we have reached a contradiction. $\square$