DISCUSSION PAPER SERIES

DP18084

WHITE, MALE, AND ANGRY: A REPUTATION-BASED RATIONALE

Stephane Wolton

POLITICAL ECONOMY

ISSN 0265-8003

WHITE, MALE, AND ANGRY: A REPUTATION-BASED RATIONALE

Stephane Wolton

Discussion Paper DP18084
Published 17 April 2023
Submitted 04 April 2023

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

WHITE, MALE, AND ANGRY: A REPUTATION-BASED RATIONALE

Abstract

From the bottom to the top of society, many white men are angry. This note provides a reputational-based rationale for this anger. Individuals care about their social status (elite vs non-elite) and their reputation (others' perception of their ability). Citizens are also uncertain about of how one becomes a member of the elite. When new information reveal that the elite is biased in favor of white men, the reputation of all white men decreases causing a payoff loss. In contrast, policies meant to reduce inequalities in the access to the elite can be supported by some white men and opposed by some individuals from the other groups. I briefly discuss how to interpret reactions to recent events (such as #Metoo and Black Lives Matter) in light of these results.

JEL Classification: D80, J70, Z13

Keywords: Discrimination, Bias

Stephane Wolton - s.wolton@lse.ac.uk
London School of Economics and CEPR and CEPR

White, Male, and Angry: A Reputation-based Rationale

Stephane Wolton*

April 4, 2023

Abstract

From the bottom to the top of society, many white men are angry. This note provides a reputational-based rationale for this anger. Individuals care about their social status (elite vs non-elite) and their reputation (others' perception of their ability). Citizens are also uncertain of the conditions to become a member of the elite. When new information reveal that the elite is biased in favor of white men, the reputation of all white men decreases causing a payoff loss. In contrast, policies meant to reduce inequalities in the access to the elite can be supported by some white men and opposed by some individuals from the other group(s). I briefly discuss how to interpret reactions to recent events (such as #Metoo and Black Lives Matter) in light of these results.

*London School of Economics and Center for Economic Policy Research. Email: s.wolton@lse.ac.uk

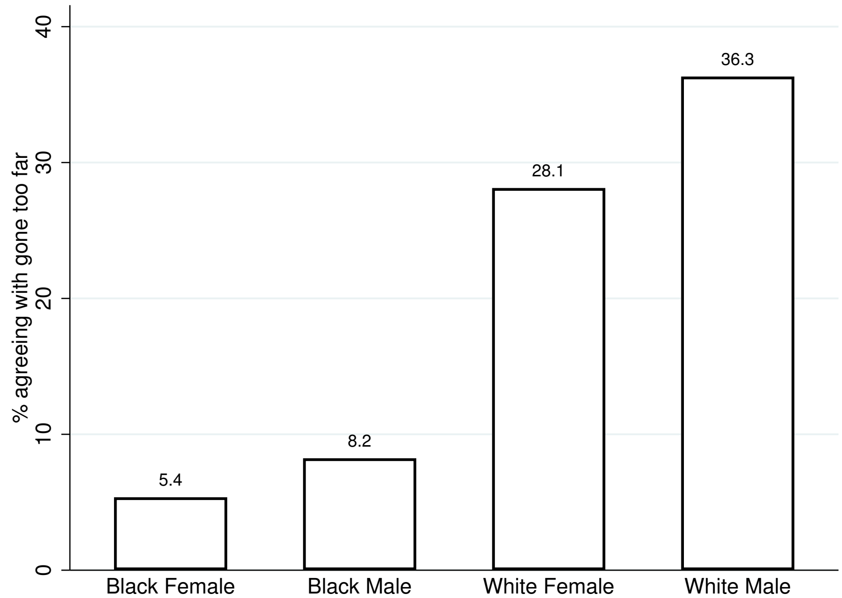
"I'm looking forward to a day when I am judged only on the content of my character" (comedian Nick Dixon tweeted after being turned down for a gig, Telegraph, 8 June 2021). Angry White Man. "There comes a point in time where you can't take anymore. It's like, enough is enough" (former American Express employee Nick Williams after allegedly being fired for being white, Fox News Digital, 13 June 2022). Angry White Man. "White male authors face another form of racism when it came to trying to break through as writers in TV, film, theater or publishing" (author James Paterson, Sunday Times, 13 June 2022). Angry White Man. "Middle-class white men are the most discriminated against in the television industry" (news presenter Jeremy Paxman, Telegraph, 24 August 2008). Angry White Man. Nick Dixon, Nick Williams, James Paterson, Jeremy Paxman are not alone. From the upper class to the working class, many white men are angry (Kimmel, 2017).

Why such anger? For some, recent changes are the cause of white men's irritation. Fifty-five percent of white Americans believe that there is discrimination against white people today (NPR, 24 October 2017), with this proportion increasing over time (The Conversation, 22 July 2022). This is not just a US phenomenon (as the quotes above also attest). In the United Kingdom, white men are four times more likely than black men to agree that attempts to give equal opportunity to ethnic minorities have gone too far as shown in Figure 1 (see also Online Appendix A, for similar regression results controlling for respondents' characteristics). White men then perceive themselves as "the new minority" (Gest, 2016).

Others disagree with this view. Scholars continuously document the advantages of white men. In the workplace, firms are as white as ever (Ray, 2019) with men earning more than women (Bonaccolto and Briel, 2022) and white men more than black men (Bayer and Charles, 2018). In everyday life, white people experience much less hardship than people of color (Kendall, 2012). Discourses about discrimination against women (e.g., #MeToo, Hilstrom, 2018) or about discrimination against African-Americans (e.g., Black Lives Matter, Taylor, 2016) can be re-interpreted as revealing, by contrapositive, *how easy* life has been for white men. As a result, white men in Western societies have lost some of their aura. "[T]hey'd lost some words that had real meaning to them: *honor, integrity, dignity*" (Kimmel, 2017: X, italics in the text).

This paper proposes a way to think about white men's anger using a stylised formal framework, which relies on three key assumptions. First, individuals are characterized by their group identity, their social status, and their ability. The society is divided between group D and group d, with

Figure 1: Attempts to give equal opportunity to ethnic minorities



Source: British Election Survey. Note: See Online Appendix A for more details.

D corresponding to the dominant group, in our context, white men. Social status corresponds to elite (upper class, college educated, wealthy) versus non-elite. Individuals' ability, in turn, affects their chances of joining the elite.

Second, the system is meritocratic within each identity group. Individuals with higher ability are more likely to succeed socially, though luck always plays a role in one's success. Yet, I assume that individuals may be uncertain about the bars individuals from each group have to pass to reach a high social status. Those bars need not be the same in each group.

Lastly, individuals care about their social status and their reputation, defined as others' perceptions of their ability. Reputation can matter for instrumental reason, better reputation yields better jobs, or intrinsic reason, for the prestige associated with being highly thought of (Origgi, 2019). Reputation in my model is closely connected to to Gidron and Hall's (2017) concept of 'subjective social standing' ("the level of social respect or esteem people believe is accorded to them within the social order" S67).

I study the consequences on individuals' payoffs of the two recent developments described above: (i) the arrival of new information about the difficulties of joining the elite for each group and (ii) facilitating the access to the elite for members of group d. I show that the first is likely to unify individuals along identity lines. To see why, suppose that it is revealed that citizens from group D only need to pass a low threshold to join the elite. This diminishes the accomplishments of those who have made it, reducing their reputation and triggering a payoff loss. It also exacerbates the failures of those who do not belong to the elite, also reducing their utility via worsened reputation. In turn, an easy bar for group D means a high bar for group d and so better reputation and higher payoff for all citizens with this identity.

Reducing inequalities in the chances of accessing the elite has, in turn, two contrasting effects. On the one hand, it lowers the chances that individuals from group D obtain a high social status. On the other hand, it increases the reputation of all group-D citizens (embellishing success, which is now harder to obtain, justifying failures, which are now more frequent). I explain how this dual impact can split group D into two. Individuals with very high ability and very low ability support policies helping the other group. It does not change their chances of joining the elite much, which remains relatively high and low, respectively, and improve their reputation. In turn, individuals with intermediary ability are the losers with the immediate loss in term of odds of reaching high status dominating any reputational gain. The reverse holds true for individuals from group d. Individuals with average ability tend to gain, individuals at the extremes tend to lose from the policy.

What is the source of white men's anger then? Is it a sense of status loss or opportunity loss? Evidence, so far, suggest it is both. Consistent with the idea that the aura of white men has decreased in recent years, the present irritation is shared by (some) white men from all social classes. Consistent with the removal of some privileges of white men, many are favorable to some form of positive action (and some members of minorities reject any form of affirmative action as shown in Figure 1 and Online Appendix A).[1]

Before turning to the model in the next section, I briefly connect my work to the most related formal literature, which looks at the role of identity in every day interaction. A long tradition considers how individuals use identity to form judgements about others (e.g., Phelps, 1972). This has led individuals from disadvantaged groups to seek to erase their identity and assimilate into the dominant group (Fang, 2001; Eguia, 2017). As a reaction, both members of the dominant group and those left behind in the disadvantaged group can "unite" to increase the cost of abandoning

[1]In an ironic twist, this paper got picked up by a blog leading many angry white men to comment on its content. Those comments illustrate how both channels discussed above can be the source of white men's anger. Some complained about their reduced reputation ("White American men went to sleep feeling as if they were on top of the world (having defeated nazi Germany and imperial Japanese and then winning the Cold War) and woke up only to feel that they were no longer welcomed in the country they built. How did we expect them to feel if not angry?") and others denounced their now lower chances of gaining high social position ("I think white men are being discriminated against so [anger] seems like a perfectly reasonable thing for us to think.").

one's original identity (Austen-Smith and Fryer Jr., 2005; Carvahlo, 2012; Schnakenberg, 2013). This form of defences can be subtle as McGee (2022) shows. Members of the dominant group can form negative beliefs about the distribution of ability of other individuals, stereotypes in McGee's words, in order to sustain their social dominance. My paper complements these important works with one twist. Even when the distribution of ability in all groups is known to be the same, differences in reputations can arise when individuals are uncertain about what it takes to join the elite.

As such, my work is also connected to Ashworth et al. (2023) who study whether women's underrepresentation in politics can be explained by differences in cost of entry or by voters' discriminatory behaviour. Like in the present work, many of the theoretical results in Ashworth et al. (2023) rely on differences in reputation. In their paper, the reputation is endogenous to male and female individuals' entry decision into politics, but there is perfect information about the different standards applied to men and women. In my manuscript, instead, reputation depends on an exogenous bar, which can be unknown to all actors. Further, I study the consequences of two policies: quotas and the revelation of new information. In particular, I highlight how revealing biases in the system can lead to a loss or gain in payoffs for all individuals with the same identity, independently of their social status.

A formal illustration of the argument

Baseline set-up

Take a large society with a mass of individuals/citizens. Individuals are characterized by their group identity, their social status, and their ability. On the identity dimension, a proportion α of citizens belong to the dominant group, denoted D, the rest $(1-\alpha)$ belongs to the disadvantaged group, denoted d. Regarding social status, a proportion e, commonly known, of the population belongs to the elite, with the rest being non-elite. Ability is an individual i's underlying type, denoted θ^i. This type is her private information. It is common knowledge that each citizen's ability is drawn independently and identically (i.i.d.) according to the cumulative distribution function (CDF) $F(\cdot)$, with associated probability density function (pdf) function $f(\cdot)$, over the interval $[0, \overline{\theta}]$.

Ability matters to reach an elite social status. So does luck. I capture luck by a random shock ϵ^i for each individual i. This shock is distributed i.i.d. according to the CDF $\Lambda(\cdot)$, with associated pdf $\lambda(\cdot)$, over the interval $[-\overline{\epsilon}, \overline{\epsilon}]$. No individual (whether i or other citizens) observes ϵ^i. Individual i from group g belongs to the elite if the sum of her ability and luck is above a threshold E_g for $g \in \{D, d\}$: $\theta^i + \epsilon^i \geq E_g$. Each citizen knows whether she is a member of the elite, whether another citizen is a member of the elite, and the way the system works. Individuals are, however, uncertain about the relevant threshold for each group. The common knowledge and shared prior is that $\widetilde{E}_D$ is distributed according to the CDF $\Gamma(\cdot)$, with associated pdf $\gamma(\cdot)$, over the interval $[\underline{E}_D, \overline{E}_D]$ with $0 < \underline{E}_D < \overline{E}_D < \overline{\theta}$.[2]

An individual i cares about her elite status $s^i \in \{0, 1\}$, with $s^i = 1$ denoting a member of the elite, and her reputation. Reputation consists of how other individuals *on average* perceive the ability of individual i given her social status. I denote it by $E_{-i}(\widetilde{\theta^i}|g^i, s^i)$. A citizen i's payoff is thus:

$$U^i(g^i, s^i) = s^i + \mathbb{E}_{-i}(\widetilde{\theta^i}|g^i, s^i)$$

The game, in turn, proceeds as follows. Nature determines each individual's ability θ^i according to the CDF $F(\cdot)$ over the interval $[0, \overline{\theta}]$ and each citizen's luck ϵ_i according to the CDF $\Lambda(\cdot)$ over the interval $[-\overline{\epsilon}, \overline{\epsilon}]$. Citizens in each group $g \in \{D, d\}$ with $\theta^i + \epsilon^i$ above the threshold E_g become elite members. Individuals form beliefs about others' abilities based on their social status and group membership ($\mathbb{E}_{-i}(\cdot)$). Payoffs are realized.

Before proceeding to the analysis, I impose a few restrictions on the model primitives. I assume that all pdfs (f, λ, γ) are continuous. I further assume that $\lambda(\cdot)$ is symmetric around 0, $\lambda'(\epsilon)$ is continuous and $\frac{\lambda'(\epsilon)}{\lambda(\epsilon)}$ is decreasing with ϵ. These last properties are satisfied by many common distributions such as the uniform distribution or the normal distribution (truncated or not). I also assume that luck plays a significant role. The least able individual has a chance to join the elite thanks to good luck and the ablest individual may fail to become an elite member due to bad luck. Formally, $\overline{\theta} < \overline{\epsilon}$. Further, the mass of individuals in each group is such that for each level of ability θ^i the full range of luck shocks is realized. I finally impose some restrictions on $\overline{E}_D$ to avoid corner solutions, see Online Appendix B for detail, where I also collect all proofs. In term of notation, in what follows, I distinguish between unknown quantities, denoted by $\widetilde{\cdot}$ and realised quantities, without tilde.

[2]The combination of E_D with the elite size e fully determines the value of E_d.

Preliminary observations

The only quantity of interest is the reputation of an individual. All individuals from the same group ($g \in \{D, d\}$) and with the same social status ($s \in \{0, 1\}$) have the same reputation since other citizens do not observe an individual's ability (θ^i) or luck (ϵ^i). I denote the reputation conditional on the distribution of $\widetilde{E}_D$ (Γ): $\theta_g^*(s; \Gamma) \equiv \mathbb{E}_{-i}(\widetilde{\theta}_i | s, g, \Gamma)$. The following lemma details some properties of individuals' reputations within each group:

Lemma 1. *Elite members have higher reputation than non-elite members:* $\theta_g^*(1; \Gamma) > \theta_g^*(1; \Gamma)$ *for all* $g \in \{d, D\}$ *and all* Γ.
Suppose that Γ_1 *first order stochastically dominate* Γ_2*, then:* $\theta_D^*(s; \Gamma_1) > \theta_D^*(s; \Gamma_2)$ *and* $\theta_d^*(s; \Gamma_1) < \theta_d^*(s; \Gamma_2)$ *for all* $s \in \{0, 1\}$.

The first point in Lemma 1 is straightforward. Given the meritocratic nature of the society (one belongs to elite if $\widetilde{\theta^i} + \widetilde{\epsilon^i} \geq \widetilde{E_g}$), abler individuals have greater chances of joining the elite. Hence, elite members have higher reputation than non-elite individuals. The second point comes from two observations. First, a higher threshold is associated with a higher reputation for an individual in the elite as well as a non-elite member. An elite member has made it despite the difficulties of succeeding and so must, in general, be of very high ability. An individual who failed to join the elite has good excuses for their lower status. Second, first order stochastic dominance implies that high realisations of the threshold are more likely under Γ_1 than under Γ_2. As a result, the reputation of all individuals from group D is higher under Γ_1 than under Γ_2.

What about group d? Recall that the size of the elite is e and commonly known. A high threshold for group D means that few members of the elite are from this group, which indicates that many members of the elite are from group d. For this to be true, it must be that the threshold for group d is relatively low, yielding a low reputation for all members of this identity group by the same reasoning as above. The reputation of the disadvantaged group (whatever their social status) is then higher under Γ_2 than Γ_1, the reverse than for the dominant group.

The effect of new information

To look at the effect of information, I assume that all individuals receive a public signal z distributed over the interval $[\underline{z}, \overline{z}]$ with CDF and associated pdf conditional on the E_D (the realized threshold for group D): $Z(\cdot|E_D)$ and $\zeta(\cdot|E_D)$, respectively. The conditional distributions satisfy the strict monotone likelihood ration property (MLRP): $\frac{\zeta(z|E_D^h)}{\zeta(z'|E_D^h)} > \frac{\zeta(z|E_D^l)}{\zeta(z'|E_D^l)}$ for all $z > z'$, $E_D^h > E_D^l$. This is a common assumption on information structure (see Milgrom, 1981), often used to understand the consequences of new information (e.g., Ashworth, Bueno de Mesquita, and Friedenberg, 2018). It states that a high E_D yields relatively more high than low signals than a low E_D.

The consequences of new information is summarized in the next proposition. To state it, I complement the notation introduced above by denoting $\theta_g^*(s;\Gamma|z)$ the reputation of individuals in group g and social group s after public signal $z \in [\underline{z}, \overline{z}]$ (recall $\theta_g^*(s;\Gamma)$ is the pre-signal reputation).

Proposition 1. *There exist a unique $z^0 \in (\underline{z}, \overline{z})$, such that for all $s \in \{0, 1\}$*

- $\theta_D^*(s;\Gamma|z^0) = \theta_D^*(s;\Gamma)$ *and* $\theta_d^*(s;\Gamma|z^0) = \theta_d^*(s;\Gamma)$*;*
- *For all $z > (<) z^0$, $\theta_D^*(s;\Gamma|z) > (<) \theta_D^*(s;\Gamma)$ and $\theta_d^*(s;\Gamma|z) < (>) \theta_d^*(s;\Gamma)$.*

A low signal ($z < z^0$) reveals that the elite is likely to be biased in favour of the dominant group. For example, it could indicate that the proportion of individuals from the dominated group belonging to the elite is low (and, therefore, the elite is dominated by group D). As a result, individuals realize that the bar E_D is low for group-D members. This reduces the reputation of elite member from this group, their successes are diminished, and the reputation of group-D non-elite individuals, their failures are exacerbated. As we have seen above, the reputation of individuals in the dominated group moves in the opposite direction than the reputation of members of the dominant group. Consequently, a low signal improves the reputation of *all members from the dominated group, no matter their social status.* The effect is obviously reverse for a high signal as it indicates that the system is likely to be biased against the dominant group.[3]

This subsection highlights that new information unify individuals from one group and polarize them with members of other groups. When the dominant group loses from the public signal, the dominated group gains, and vice versa. If one interpret recent events, such as #MeToo or Black Lives Matter, as information revelation on the extent of biases in society, Proposition 1 highlights a way to think about white men's rising anger.

[3]Formally, the MLRP property of the signal implies that for all $z > z'$, the CDF of the thresholds conditional on the signal ($\Gamma(\cdot|\cdot)$) is such that $\Gamma(\cdot|z)$ first order stochastically dominate $\Gamma(\cdot|z')$ (Milgrom, 1981). Using Lemma 1, this implies that $\theta_g^*(s;\Gamma|z)$ is strictly increasing with z. Further, it can then be shown that for sufficiently high signal z^t, $\Gamma(\cdot|z^t)$ first order stochastically dominates $\Gamma(\cdot)$, the prior distribution, so that $\theta_g^*(s;\Gamma|z^t) > \theta^*(s;\Gamma)$. Inversely, for sufficiently low signal z^b, $\theta_g^*(s;\Gamma|z^b) < \theta^*(s;\Gamma)$. Applying the theorem of intermediate values, we obtain Proposition 1.

Changing the entry conditions into the elite

Suppose (for simplicity) that there is no uncertainty about the system (E_D and E_d are known). Instead, consider the following policies meant to correct potential imbalances within the elite:

- The threshold for the dominant group is increased by $\Delta > 0$.
- The threshold for the disadvantaged group is decreased by $\delta > 0$.

To build some intuition on the effect of such reform, I focus on members of the dominant group. Absent reputation, all individuals from the dominant group would oppose the reform. An increase in the threshold to $E_D + \Delta$ reduces the chances that an individual from group D joins the elite. Multiplied by the payoff gain from being in the elite, this represents the loss from the policy. This is not, however, the only effect of such reform. By moving the bar upward, the policy also increases the reputation of all individuals from the dominant group. Thus, (some) members of the dominant group may support the reform.

To study whether this is the case, I look at a marginal change in the thresholds for both groups. This corresponds to *New York Times* journalist Anna Quindlen's observation that "[i]t is a system that once favored [a white male], and others like him. Now sometimes—just sometimes—it favors someone different" (cited in Kimmel, 2017, 219). Let $W_D(\theta^i, \Delta)$ and $W_d(\theta^i, \delta)$ be the expected utility of an individual i with ability θ^i (prior to her social status being determined) when the threshold increases by Δ and decreases by δ for group D and d respectively.

Proposition 2. *There exist θ^l_g, θ^h_g, $\theta^l_g < \theta^h_g$, unique if $\theta^j_h \in [0, \bar{\theta}]$ ($j \in \{l, h\}$), such that:*

- *In group D, for all individuals with $\theta^i \in [\theta^l_D, \theta^h_D]$, $\left.\frac{\partial W_D(\theta^i,\Delta)}{\partial \Delta}\right|_{\Delta=0} \leq 0$, for all individuals with $\theta^i \notin [\theta^l_D, \theta^h_D]$, $\left.\frac{\partial W_D(\theta^i,\Delta)}{\partial \Delta}\right|_{\Delta=0} > 0$.*
- *In group d, for all individuals with $\theta^i \in [\theta^l_d, \theta^h_d]$, $\left.\frac{\partial W_d(\theta^i,\delta)}{\partial \delta}\right|_{\delta=0} \geq 0$, for all individuals with $\theta^i \notin [\theta^l_d, \theta^h_d]$, $\left.\frac{\partial W_d(\theta^i,\delta)}{\partial \delta}\right|_{\delta=0} < 0$.*

Proposition 2 indicates who is more likely to win from small changes to the thresholds in each group. Absent further assumptions, the thresholds $\{\theta^l_g, \theta^h_g\}_{g \in \{d,D\}}$ are left undetermined. One particularly interesting case arises when $\left.\frac{\partial W_D(0,\Delta)}{\partial \Delta}\right|_{\Delta=0} > 0$, $\left.\frac{\partial W_D(\bar{\theta},\Delta)}{\partial \Delta}\right|_{\Delta=0} > 0$, and $\left.\frac{\partial W_D(\mathbb{E}(\theta^i),\Delta)}{\partial \Delta}\right|_{\Delta=0} < 0$. Then, the lowest ability members from group D have too little chances to join the elite to care about the direct effect of the policy change. They just benefit from the reputational gain. The individuals with the highest ability always have good odds to become elite members pre- or post-reform, so they also mostly enjoy the boost in their reputation. In contrast, individuals with average ability suffer from a change in the thresholds E_D. Within the dominant group, the "middle class" and the top and bottom of the distribution stand in opposition with each other. A similar patter can arise in the dominated group, but with different attitudes towards the policy change. Individuals in the middle see the greatest gain because they now can join the elite with higher probability; individuals with very low and very high abilities reject the reform due to their loss in reputation.

Is this case empirically relevant? Various empirical works suggest so. The literature on quotas for female candidates in politics have documented that mediocre men are the main losers and the quality of women selected to run decreases slightly, though the coefficients for women are generally non-statistically significant (Lassébie, 2020; Besley et al., 2017; Baltrunaite et al., 2014). Similarly, changes in the rules for selection to the National Academy of Science and the American Academy of Arts and Science have benefited women over men, with (likely) very good reasons for this pattern (Card et al., 2023). In all these cases, the main beneficiaries (losers) from the policy seem to be individuals with intermediate abilities in the dominated (dominant) group, just like hypothesized in the paragraph above. As such, it seems empirically plausible that reducing inequalities, unlike providing information, can generate splits within identity groups.

Conclusion

This paper suggests two possible rationales for the rising anger among white men when individuals care about their social status and their reputation (others' perception of their ability). I show that new information about biases in the system raise or decrease the reputation of all individuals with the same identity, no matter their social status. In contrast, change in privileges generally divide identity groups between winners and losers from the policy change.

These results can also serve to analyse recent political phenomena. White men, after all, have been the core constituency of Donald Trump (Pew, 9 August 2018 and 30 June 2021). This attraction to "the first white president" (Coates, 2017) is often seen under the prism of a mixture of cultural and economic backlashes (most recently, Baccini and Weymouth, 2021). Yet, how does culture matter? When voters care about their reputation, or their subjective social status (Gidron and Hall, 2017), cultural appeal, understood as information provision, can unite individuals along one identity dimension. All members from the same identity group react positively to signals favouring their group's narrative, such as a biased system against white men (contra Norris and Inglehart, 2019, who argue that only older cohorts are susceptible to populists' cultural messages, but consistent with Schäfer, 2022). The same cannot be guaranteed with policy changes as per Proposition 2.[4] Developing these observations into real insights about the appeal and strategies of populists requires a fully-fledged model of electoral competition. This is left for future research.

[4] See also Dewan and Wolton (2022) where symbolic policies (like the burqa ban) split the electorate along identity lines (natives versus non-natives) and social lines (native workers versus native employers).

References

Ashworth, Scott, Ethan Bueno de Mesquita, and Amanda Friedenberg. 2018. "Learning about voter rationality." American Journal of Political Science 62(1): 37-54.

Ashworth, Scott, Christopher R. Berry, and Ethan Bueno de Mesquita. 2023 "Modeling Theories of Women's Underrepresentation in Elections." American Journal of Political Science, forthcoming. Available here.

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

Baccini, Leonardo, and Stephen Weymouth. 2021. "Gone for good: Deindustrialization, white voter backlash, and US presidential voting." American Political Science Review 115(2): 550-567.

Baltrunaite, Audinga, Piera Bello, Alessandra Casarico, and Paola Profeta. 2014. "Gender quotas and the quality of politicians." Journal of Public Economics 118: 62-74.

Bayer, Patrick, and Kerwin Kofi Charles. 2018. "Divergent paths: A new perspective on earnings differences between black and white men since 1940." The Quarterly Journal of Economics 133(3): 1459-1501.

Besley, Timothy, Olle Folke, Torsten Persson, and Johanna Rickne. "Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden." American Economic Review 107(8): 2204-42.

Bonaccolto-Töpfer, Marina, and Stephanie Briel. 2022. "The gender pay gap revisited: Does machine learning offer new insights?." Labour Economics 78: 102223.

Card, David, Stefano DellaVigna, Patricia Funk, and Nagore Iriberri. 2023. "Gender gaps at the academies." Proceedings of the National Academy of Sciences 120(4): e2212421120.

Carvalho, Jean-Paul. 2013. "Veiling." The Quarterly Journal of Economics 128(1): 337-370.

Coates, Ta-Nehisi. 2017. "The first white president." *The Atlantic* 320(3): 74-87.

Dewan, Torun and Wolton, Stephane. 2022. "A Political Economy of Social Discrimination." CEPR Discussion Paper No. DP17291. Available at https://ssrn.com/abstract=4121476.

Eguia, Jon X. 2017. "Discrimination and assimilation at school." Journal of Public Economics 156: 48-58.

Fang, Hanming. 2001. "Social culture and economic performance." American Economic Review 91(4): 924-937.

Gest, Justin. 2016. *The new minority: White working class politics in an age of immigration and inequality.* Oxford, UK: Oxford University Press.

Gidron, Noam, and Peter A. Hall. 2017. "The politics of social status: Economic and cultural roots of the populist right." The British journal of sociology 68: S57-S84.

Hillstrom, Laurie Collier. 2018. *The #metoo movement.* Santa Barbara, CA: ABC-CLIO.

Kendall, Frances. 2012. *Understanding white privilege: Creating pathways to authentic relationships across race.* New York, NY: Routledge.

Kimmel, Michael. 2017. *Angry white men: American masculinity at the end of an era, 2nd edition.* New York, NY: Bold Type Books.

Lassébie, Julie. 2020. "Gender quotas and the selection of local politicians: Evidence from French municipal elections." European Journal of Political Economy 62: 101842.

McGee, Dan. 2022. "Stereotypes Under Group Competition," Working paper. Available at https://www.dropbox.com/ s/8ycxxs9r96ycp71/Stereotypes.

Milgrom, Paul R. 1981. "Good news and bad news: Representation theorems and applications." The Bell Journal of Economics 12(2): 380-391.

Norris, Pippa, and Ronald Inglehart. 2019. *Cultural backlash: Trump, Brexit, and authoritarian populism.* Cambridge, MA: Cambridge University Press, 2019.

Origgi, Gloria. 2019. *Reputation: What it is and why it Matters.* Princeton, NJ: Princeton University Press.

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

Ray, Victor. 2019. "Why So Many Organizations Stay White." Harvard Business Review, 19 November 2019. Available at https://hbr.org/2019/11/why-so-many-organizations-stay-white (last accessed 27 August 2022).

Schäfer, Armin. 2022. "Cultural Backlash? How (Not) to Explain the Rise of Authoritarian Populism." British Journal of Political Science 52(4): 1977–93.

Schnakenberg, Keith E. 2013. "Group Identity and Symbolic Political Behavior." Quarterly Journal of Political Science, 9: 137-167.

Taylor, Keeanga-Yamahtta. 2016. *From #BlackLivesMatter to black liberation.* Chicago, IL: Haymarket Books.

Online Appendix

(Not for publication)

A Survey evidence

Data and empirical specifications

Figure 1 and the analysis below use data from the British Election Study obtained from the UK Data service (the British Election Study Combined Wave 1-20 Internet Panel).[1] The main (dependent) variables used come from the goneTooFarGrid:

- blackEquality: Attempts to give equal opportunities to ethnic minorities,
- femaleEquality: Attempts to give equal opportunities to women,
- gayEquality: Attempts to give equal opportunities to gays and lesbians.

Respondents were asked whether the policies for each group above have "not gone nearly far enough," "not gone far enough," are "about right," have "gone too far," "gone much too far." Respondents were also allowed to answer "don't know." I dichotomize these variables with the new variables taking value one if the respondent has answered that a given attempt to give equal opportunity has gone too far or much too far and zero otherwise. The dummy variables are used throughout the analysis. The questions were asked in waves 1, 6, then 10 to 16 continuously (dates associated with each wave can be found in the documentation of the study available here). However, the data do not allow to distinguish waves for respondents who answered from wave 6 to wave 14.

The main explanatory variables are gender and ethnicity. For gender, I construct a dummy Male equal to one if the respondent is male (using the variable gender). For ethnicity, I construct a dummy White equal to one if the respondent reports being of white background (using the variable

[1] These data are safeguarded and require an account with UK Data Service to be obtained. However, the data can also be accessed from the British Election Study website.

p_ethnicity). I also construct a variable WhiteMale, which corresponds to the interaction of the two variables.

For the regression analysis, I use the following individual controls:

- a dummy equal to one if the respondent has finished university (using the variable p_education),
- a dummy equal to one if the respondent has finished high school (again using the variable p_education),
- a dummy equal to one if the respondent reports being of black ethnicity (again using the variable p_ethnicity),
- a dummy equal to one if the respondent owns their house, with or without mortgage (using the variable p_housing),
- a dummy equal to one if the respondent is married (using the variable p_marital),
- a dummy equal to one if the respondent is divorced (again using the variable p_marital),
- age (using variable age),
- fixed effects by income status by £5,000 or £10,000 increment (using p_gross_household),
- fixed effects by working status: full time, part time, etc. (using p_work_stat),
- fixed effects by type of work: private, public, non-profit (using p_job_sector).

The exact wording for each control variables can be found in the BES documentation (available here).

For the regression analysis, I run the following model (both as a linear probability model and a probit model):

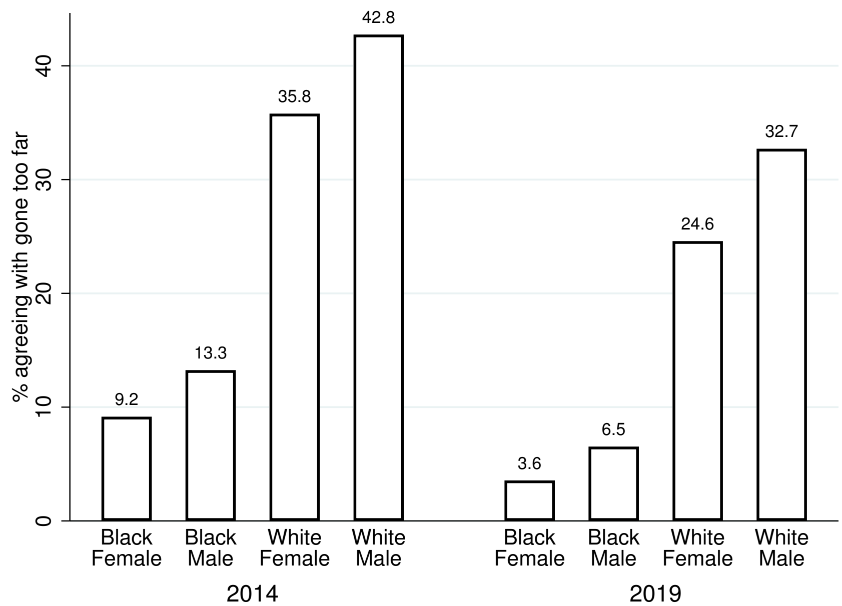
$$Y_{iW} = \alpha + \beta_1 White_{iW} + \beta_2 Male_{iW} + \beta_3 WhiteMale_{iW} + \gamma' X_{iW} + \delta_W + \epsilon_{iW}$$

Y_{iW} is a dummy equal to one if respondent i interviewed in wave W has answered that attempts to give equal opportunity to ethnic minorities/women/gay and lesbians have gone too far or much too far. White, Male, and WhiteMale are gender and ethnic dummies, X is the set of controls described above, δ_W is a fixed effect for each wave, and ϵ_{iW} are the standard errors clustered at the wave level.

Answers over time

While UK residents have become more tolerant of these policies over time, unlike the US, the gap between Black and White males' responses has hardly changed between 2014-15 (29.5%) and 2019 (26.2%) as shown in Figure A.1.

Figure A.1: Attempts to give equal opportunity to ethnic minorities



Note: Percentages come from the British Election Survey. The first set of bars correspond to answers given in wave 1 (in 2014). The second set of bars correspond to answers given in waves 15 to 17 (in 2019).

Regression results

Linear Probability Model

Table A.1 highlights that white male respondents complain the most about actions to favour equal opportunity to minority even after controlling, actually especially after controlling, for a bunch of respondents' characteristics such that income, education, age, or marital status. The differences are substantively significant, white males are around 11.5% more likely than non-white males (which include Asians and South Asian individuals, two large categories in the UK) and around 7% more likely than white women to agree that efforts to improve the fate of ethnic minorities have gone too far. This is relative to a mean of 30.5%.

The regression analysis also highlights how opposition to policies which favor disadvantaged groups depend on the group being made salient. Men (white or not) are the most likely to complain about policies for women or for gays and lesbians (the British Election Survey does not ask about LGBTQ+) because race is not highlighted in the questions.

Table A.1: Equal opportunity policies gone too far (White vs Non-White, Linear probability model)

	(1)	(2)	(3)	(4)	(5)	(6)
Equal opport. to	Minorities gone too far		Women gone too far		Lesbians-Gays gone too far	
White	0.042***	0.089***	0.027**	0.013	0.055***	-0.048***
	(0.006)	(0.000)	(0.044)	(0.200)	(0.003)	(0.007)
Male	0.071***	0.049***	0.080***	0.093***	0.124***	0.063***
	(0.000)	(0.009)	(0.000)	(0.000)	(0.000)	(0.000)
White*Male	0.011*	0.020*	-0.005	-0.019***	0.003	0.049***
	(0.061)	(0.055)	(0.552)	(0.002)	(0.691)	(0.005)
Individual controls		✓		✓		✓
N.obs	182732	138655	183262	138874	182732	138655

Notes: Dependent variables is a dummy equal to one if respondents agree to gone (much) too far. Robust standard errors are clustered by wave. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Equal opportunity policies gone too far (Black vs Non-Black, Linear probability model)

	(1)	(2)	(3)	(4)	(5)	(6)
Equal opport. to	Minorities gone too far		Women gone too far		Lesbians-Gays gone too far	
Black	-0.217***	-0.142***	-0.049*	-0.041**	0.026	0.087***
	(0.000)	(0.000)	(0.054)	(0.037)	(0.163)	(0.000)
Male	0.078***	0.067***	0.076***	0.075***	0.126***	0.109***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Black*Male	-0.049***	-0.053***	0.035**	0.032**	-0.038	-0.037
	(0.000)	(0.002)	(0.034)	(0.050)	(0.375)	(0.360)
Individual controls		✓		✓		✓
N.obs	182732	138655	183262	138874	182732	138655

Notes: Dependent variables is a dummy equal to one if respondents agree to gone (much) too far. Robust standard errors are clustered by wave. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Probit

The tables below show the analysis using a probit model (instead of a linear probability model). In the first and second columns of Table A.3, note that the intersection is not significant, yet due to the nature of the probit model, the probability a white man agrees with the statement that equal opportunity policies for minorities have gone too far is 42.2% against 33.5% for a white woman, 23.6% for a non-white man, and 15.4% for a non-white woman (using estimates from column (2)). There is no difference across ethnic group (white vs non-white), just based on gender—approx. 16.3% for men and 8% for women—when it comes to disapproving equal opportunity for women (using estimates from column (4)).

Table A.3: Equal opportunity policies gone too far (White vs Non-White, Probit regression)

	(1)	(2)	(3)	(4)	(5)	(6)
Equal opport. to	Minorities gone too far		Women gone too far		Lesbians-Gays gone too far	
White	0.131**	0.391***	0.165**	0.091	0.205***	-0.174**
	(0.013)	(0.000)	(0.049)	(0.192)	(0.009)	(0.016)
Male	0.215***	0.252**	0.418***	0.467***	0.425***	0.212***
	(0.001)	(0.015)	(0.000)	(0.000)	(0.000)	(0.001)
White*Male	0.015	-0.045	-0.077	-0.124**	-0.041	0.144**
	(0.351)	(0.300)	(0.138)	(0.031)	(0.142)	(0.018)
Individual controls		✓		✓		✓
N.obs	182732	138655	183262	138874	182732	138655

Notes: Dependent variables is a dummy equal to one if respondents agree to gone (much) too far. Robust standard errors are clustered by wave. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Equal opportunity policies gone too far (Black vs Non-Black, Probit regression)

	(1)	(2)	(3)	(4)	(5)	(6)
Equal opport. to	Minorities gone too far		Women gone too far		Lesbians-Gays gone too far	
Black	-1.001***	-0.785***	-0.364*	-0.327*	0.091	0.327***
	(0.002)	(0.003)	(0.083)	(0.068)	(0.195)	(0.001)
Male	0.222***	0.203***	0.357***	0.347***	0.392***	0.351***
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Black*Male	-0.001	-0.016	0.310**	0.307**	-0.124	-0.139
	(0.988)	(0.834)	(0.047)	(0.046)	(0.305)	(0.265)
Individual controls		✓		✓		✓
N.obs	182732	138655	183262	138874	182732	138655

Notes: Dependent variables is a dummy equal to one if respondents agree to gone (much) too far. Robust standard errors are clustered by wave. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Proofs

Before proving the formal result, I detail one of the assumptions mentioned in the text regarding the relationship between the thresholds and the elite size. Under the assumption that all luck shocks are realised for a given ability level, the thresholds for the dominant and dominated groups must satisfy (denoting $\mathbb{E}_\epsilon(\cdot)$ the expectation with respect to the luck shock):

$$e = \alpha\mathbb{E}_\epsilon\big(1 - F(E_D - \widetilde{\epsilon})\big) + (1-\alpha)\mathbb{E}_\epsilon\big(1 - F(E_d - \widetilde{\epsilon})\big) \tag{B.1}$$

I assume that the distribution of ability satisfies: $1 - e < \alpha\mathbb{E}_\epsilon(F(\overline{E}_D - \widetilde{\epsilon})) + (1-\alpha)\mathbb{E}_\epsilon(F(0 - \widetilde{\epsilon}))$ and $1 - e > \alpha\mathbb{E}_\epsilon(F(\underline{E}_D - \widetilde{\epsilon})) + (1-\alpha)\mathbb{E}_\epsilon(F(\overline{\theta} - \widetilde{\epsilon}))$. This assumption implies that the highest ($\overline{E}_d$) and lowest ($\underline{E}_d$) possible realisation of the threshold for group satisfy $\overline{E}_d < \overline{\theta}$ and $0 < \underline{E}_d$.

As noted in the text, reputation is the only interesting quantity here. In practice, each individual $j \neq i$ forms a different expectation about i's ability based on j's own ability, group, and social status since it affects j's belief about $\widetilde{E_D}$ and $\widetilde{E_d}$. However, because reputation takes the form of the average opinion of other individual about i's ability, we obtain by the law of iterated expectation:

$$\mathbb{E}_{-i}(\widetilde{\theta^i}|g^i, s^i) = \mathbb{E}_{g^j, s^j, \theta^j}\big(\mathbb{E}(\widetilde{\theta^i}|g^i, s^i, g^j, s^j, \theta^j)\big) = \mathbb{E}(\widetilde{\theta^i}|g^i, s^i) = \mathbb{E}(\widetilde{\theta^i}|\widetilde{\theta^i} + \widetilde{\epsilon^i} \geq \widetilde{E_g})$$

In other words, we can use the prior distributions for luck and threshold values rather than the distributions conditional on others' successes/failures and abilities.[2]

Using this observation, I can define the reputation of each social group for each identity group: $\theta^*_g(1;\Gamma)$. To do so, I slightly modify the notation in the main text. Let $\Gamma_D \equiv \Gamma$ and, using Equation B.1, Γ_d being defined by:

$$\Gamma_d(E) = 1 - \Gamma_D\left(v^{-1}\left(\frac{1 - e - (1-\alpha)v(E)}{\alpha}\right)\right), \tag{B.2}$$

with $v(E) = \mathbb{E}_\epsilon(F(E - \widetilde{\epsilon}))$ a strictly increasing function in E. Reputations take value:

$$\theta^*_g(1;\Gamma) = \int_{\underline{E}_g}^{\overline{E}_g}\int_{-\overline{\epsilon}}^{\overline{\epsilon}} \frac{\int_{\widetilde{E}-\widetilde{\epsilon}}^{\overline{\theta}} \widetilde{\theta} dF(\widetilde{\theta})}{1 - F(\widetilde{E} - \widetilde{\epsilon})} d\Lambda(\widetilde{\epsilon}) d\Gamma_g(\widetilde{E}) \tag{B.3}$$

$$\theta^*_g(1;\Gamma) = \int_{\underline{E}_g}^{\overline{E}_g}\int_{-\overline{\epsilon}}^{\overline{\epsilon}} \frac{\int_0^{\widetilde{E}-\widetilde{\epsilon}} \widetilde{\theta} dF(\widetilde{\theta})}{F(\widetilde{E} - \widetilde{\epsilon})} d\Lambda(\widetilde{\epsilon}) d\Gamma_g(\widetilde{E}) \tag{B.4}$$

[2] It is important that reputation is defined as *others'* perception of individual i's ability, not individual i's expectation of others' perception. In the latter case, $\mathbb{E}_{-i}(\cdot|\cdot)$ would also be a function of i's ability.

With this, we can easily prove Lemma 1.

Proof of Lemma 1

The first point comes from noticing that for any given realisation of ϵ and E, the properties of conditional expectations imply that:

$$\frac{\int_{E-\epsilon}^{\overline{\theta}} \widetilde{\theta} dF(\widetilde{\theta})}{1 - F(E-\epsilon)} > \frac{\int_{0}^{E-\epsilon} \widetilde{\theta} dF(\widetilde{\theta})}{F(E-\epsilon)}$$

Hence, we must have:

$$\int_{\underline{E}_g}^{\overline{E}_g} \int_{-\bar{\epsilon}}^{\bar{\epsilon}} \frac{\int_{\widetilde{E}-\widetilde{\epsilon}}^{\overline{\theta}} \widetilde{\theta} dF(\widetilde{\theta})}{1 - F(\widetilde{E}-\widetilde{\epsilon})} d\Lambda(\widetilde{\epsilon}) d\Gamma_g(\widetilde{E}) > \int_{\underline{E}_g}^{\overline{E}_g} \int_{-\bar{\epsilon}}^{\bar{\epsilon}} \frac{\int_{0}^{\widetilde{E}-\widetilde{\epsilon}} \widetilde{\theta} dF(\widetilde{\theta})}{F(\widetilde{E}-\widetilde{\epsilon})} d\Lambda(\widetilde{\epsilon}) d\Gamma_g(\widetilde{E})$$

To prove the second point, consider the function

$$H(E) = \int_{-\bar{\epsilon}}^{\bar{\epsilon}} \frac{\int_{E-\widetilde{\epsilon}}^{\overline{\theta}} \widetilde{\theta} dF(\widetilde{\theta})}{1 - F(E-\widetilde{\epsilon})} d\Lambda(\widetilde{\epsilon})$$

Notice that (after re-arranging):

$$H'(E) = \int_{-\bar{\epsilon}}^{\bar{\epsilon}} \frac{f(E-\widetilde{\epsilon})}{1 - F(E-\widetilde{\epsilon})} \left(\frac{\int_{E-\widetilde{\epsilon}}^{\overline{\theta}} \widetilde{\theta} dF(\widetilde{\theta})}{1 - F(E-\widetilde{\epsilon})} - (E - \widetilde{\epsilon}) \right) d\Lambda(\widetilde{\epsilon}) > 0$$

Hence, by definition of first-order stochastic dominance:

$$\theta_D^*(s; \Gamma^1) = \int_{\underline{E}_D}^{\overline{E}_D} H(\widetilde{E}) d\Gamma^1(\widetilde{E}) > \int_{\underline{E}_D}^{\overline{E}_D} H(\widetilde{E}) d\Gamma^2(\widetilde{E}) = \theta_D^*(s; \Gamma^2)$$

For group d, the result follows from Equation B.2. If Γ_D^1 first order stochastically dominates Γ_D^2, then the associated Γ_d^1 is first order stochastically dominated by Γ_d^2. □

Proposition 1

I only prove the result for the dominant group D. Using the same logic as for the proof of Lemma 1, the result can easily be extended to group d. The proof proceeds in four steps.

Step 1: notice that since the conditional pdfs of z satisfy the MLRP, for all $z > z'$, $\Gamma(\cdot|z)$ first order stochastically dominates $\Gamma(\cdot|z')$ (Milgrom, 1981). This means that $\theta^*_D(s;\Gamma|z)$ is strictly increasing with z using the second point of Lemma 1.

Step 2: $\Gamma(\cdot)$ (the prior distribution) first order stochastically dominates $\Gamma(\cdot|\underline{z})$. To see this, suppose it does not. First, suppose that there exists $E \in [\underline{E}_D, \overline{E}_D]$ such that $\Gamma(E|\underline{z}) > \Gamma(E)$. Now, since $\Gamma(\cdot|\underline{z})$ is first order stochastically dominated by $\Gamma(\cdot|z)$ for all $z > \underline{z}$, we must have $\Gamma(E|z) > \Gamma(E)$. Then, $\int_{\underline{z}}^{\overline{z}} \Gamma(E|\widetilde{z})dZ(\widetilde{z}) > \int_{\underline{z}}^{\overline{z}} \Gamma(E)dZ(\widetilde{z})$. By the law of total probabilities, $\int_{\underline{z}}^{\overline{z}} \Gamma(E|\widetilde{z})dZ(\widetilde{z}) = \Gamma(E)$. Since $\int_{\underline{z}}^{\overline{z}} \Gamma(E)dZ(\widetilde{z}) = \Gamma(E)$, we obtain $\Gamma(E) > \Gamma(E)$ a contradiction. Now, suppose that for all E, $\Gamma(E|\underline{z}) = \Gamma(E)$. Since $\Gamma(\cdot|\underline{z})$ is first order stochastically dominated by $\Gamma(\cdot|z)$ for all $z > \underline{z}$, there exists E' such as by the same reasoning as above, we obtain $\Gamma(E') > \Gamma(E')$, a contradiction.

Step 3: by the same reasoning, we can show that $\Gamma(\cdot)$ is first order stochastically dominated by $\Gamma(\cdot|\overline{z})$. Using this result and $\Gamma(\cdot)$ FOSD $\Gamma(\cdot|\underline{z})$, we obtain that $\theta^*_D(s;\Gamma|\underline{z}) < \theta^*_D(s;\Gamma) < \theta^*_D(s;\Gamma|\overline{z})$ (again using the second point of Lemma 1).

Step 4: Using the results from Step 1 ($\theta^*_D(s;\Gamma|z)$ strictly increasing in z) and from Step 3 ($\theta^*_D(s;\Gamma|\underline{z}) < \theta^*_D(s;\Gamma) < \theta^*_D(s;\Gamma|\overline{z})$) and the theorem of intermediate values, we obtain that there exists a unique z^0 such that $\theta^*_D(s;\Gamma) \geq (<) \theta^*_D(s;\Gamma|z)$ for all $z \geq (<) z^0$. □

Proof of Proposition 2

Consider an individual from the dominant group D with ability θ^i. Slightly amending notation, denote $\theta^*_D(s;\Gamma,\Delta)$ the reputation of group-D members with social status $s \in \{0,1\}$ after the threshold has been increased by Δ. The expected payoff of this individual is:

$$\begin{aligned} W_D(\theta^i,\Delta) =& \big(1-\Lambda(E_D+\Delta-\theta^i)\big)\big(1+\theta^*_D(1;\Gamma,\Delta)\big) \\ &+ \Lambda(E_D+\Delta-\theta^i)\big(0+\theta^*_D(0;\Gamma,\Delta)\big) \end{aligned} \tag{B.5}$$

The first term after the equal sign ($\big(1-\Lambda(\widetilde{E}+\Delta-\theta^i)\big)$) corresponds to the probability of joining the elite for an individual with ability θ^i: the luck choice must be high enough given the possible realisations of the threshold. The second term $(1+\theta_D^*(1;\Gamma,\Delta))$ corresponds to the payoff when in the elite. On the second line, the probability of missing the bar and the payoff when not in the elite.

Taking the derivative with respect to Δ, I obtain:

$$\begin{aligned}\frac{\partial W_D(\theta^i,\Delta)}{\partial\Delta} =& -\lambda(E_D+\Delta-\theta^i)\big(1+\theta_D^*(1;\Gamma,\Delta)-\theta_D^*(0;\Gamma,\Delta)\big)\\ &+big(1-\Lambda(E_D+\Delta-\theta^i)\big)\frac{\partial\theta_D^*(1;\Gamma,\Delta)}{\partial\Delta}\\ &+\Lambda(E_D+\Delta-\theta^i)\frac{\partial\theta_D^*(0;\Gamma,\Delta)}{\partial\Delta}\end{aligned}$$

Notice that using the proof of Lemma 1,

$$\begin{aligned}\frac{\partial\theta_D^*(1;\Gamma,\Delta)}{\partial\Delta} &= \int_{-\bar{\epsilon}}^{\bar{\epsilon}}\frac{f(E_D+\Delta-\widetilde{\epsilon})}{1-F(E_D+\Delta-\widetilde{\epsilon})}\left(\frac{\int_{E_D+\Delta-\widetilde{\epsilon}}^{\bar{\theta}}\widetilde{\theta}dF(\widetilde{\theta})}{1-F(E_D+\Delta-\widetilde{\epsilon})}-(E_D+\Delta-\widetilde{\epsilon})\right)d\Lambda(\widetilde{\epsilon})>0\\ \frac{\partial\theta_D^*(0;\Gamma,\Delta)}{\partial\Delta} &= \int_{-\bar{\epsilon}}^{\bar{\epsilon}}\frac{f(E_D+\Delta-\widetilde{\epsilon})}{F(E_D+\Delta-\widetilde{\epsilon})}\left((E+D+\Delta-\widetilde{\epsilon})-\frac{\int_0^{E_D+\Delta-\widetilde{\epsilon}}\widetilde{\theta}dF(\widetilde{\theta})}{F(E_D+\Delta-\widetilde{\epsilon})}\right)d\Lambda(\widetilde{\epsilon})>0\end{aligned}$$

Now consider how the derivative of $W_D(\theta^i,\Delta)$ wrt to Δ varies with ability θ^i:

$$\begin{aligned}\frac{\partial^2 W_D(\theta^i,\Delta)}{\partial\Delta\partial\theta^i} =& \lambda'(E_D+\Delta-\theta^i)\big(1+\theta_D^*(1;\Gamma,\Delta)-\theta_D^*(0;\Gamma,\Delta)\big)\\ &+\lambda(E_D+\Delta-\theta^i)\left(\frac{\partial\theta_D^*(0;\Gamma,\Delta)}{\partial\Delta}-\frac{\partial\theta_D^*(1;\Gamma,\Delta)}{\partial\Delta}\right)\end{aligned}$$

Rearranging, the sign of $\frac{\partial^2 W_D(\theta^i,\Delta)}{\partial\Delta\partial\theta^i}$ is the same as the sign of

$$\frac{\lambda'(E_D+\Delta-\theta^i)}{\lambda(E_D+\Delta-\theta^i)}+\frac{\frac{\partial\theta_D^*(0;\Gamma,\Delta)}{\partial\Delta}-\frac{\partial\theta_D^*(1;\Gamma,\Delta)}{\partial\Delta}}{1+\theta_D^*(1;\Gamma,\Delta)-\theta_D^*(0;\Gamma,\Delta)}$$

Since $\frac{\lambda'(\epsilon)}{\lambda(\epsilon)}$ is decreasing with ϵ by assumption, $\frac{\lambda'(\widetilde{E}+\Delta-\theta^i)}{\lambda(\widetilde{E}+\Delta-\theta^i)}$ is increasing with θ^i. As a result, there are three cases to consider:

(1) $\frac{\partial^2 W_D(\theta^i,\Delta)}{\partial\Delta\partial\theta^i}$ is negative for all θ^i;

(2) $\frac{\partial^2 W_D(\theta^i,\Delta)}{\partial\Delta\partial\theta^i}$ is positive for all θ^i;

(3) There exists θ^+ such that $\frac{\partial^2 W_D(\theta^i,\Delta)}{\partial\Delta\partial\theta^i}$ is strictly negative for all $\theta^i < \theta^+$ and positive for all $\theta^i > \theta^+$ (zero at $\theta^i = \theta^+$).

In all cases, we can have $\frac{\partial W_D(\theta^i,\Delta)}{\partial\Delta} < 0$ for all θ^i, in which cases pick $\theta_D^l < 0$ and $\overline{\theta} < \theta_D^h$, or $\frac{\partial W_D(\theta^i,\Delta)}{\partial\Delta} > 0$ for all θ^i, in which case pick $\overline{\theta} < \theta_D^l < \theta_D^h$. On top of this,

- In cases (1) and (3), if there exists a unique solution in $\theta^s \in [0,\overline{\theta}]$ to $\frac{\partial W_D(\theta^i,\Delta)}{\partial\Delta} = 0$ such that $\frac{\partial W_D(\theta^i,\Delta)}{\partial\Delta} < 0$ for all $\theta^i > \theta^s$, denote $\theta^s = \theta_D^l$ and pick $\theta_D^h > \overline{\theta}$.
- In cases (2) and (3), if there exists a unique solution in $\theta^s \in [0,\overline{\theta}]$ to $\frac{\partial W_D(\theta^i,\Delta)}{\partial\Delta} = 0$ such that $\frac{\partial W_D(\theta^i,\Delta)}{\partial\Delta} > 0$ for all $\theta^i > \theta^s$, denote $\theta^s = \theta_D^h$ and pick $\theta_D^l < 0$.
- In case (3), if there exists two solution in $\theta_1^s, \theta_2^s \in [0,\overline{\theta}]^2$ to $\frac{\partial W_D(\theta^i,\Delta)}{\partial\Delta} = 0$ denote $\theta_1^s = \theta_D^l$ and $\theta_2^s = \theta_D^h$.

This represents all possible cases. In all these cases, we have been able to define θ^l and θ^h satisfying the conditions of the proposition for the dominant group. We can apply a similar reasoning for the dominated group noting that δ has the opposite effect than Δ. □