

RESEARCH ARTICLE

Markets and communities: the social cost of the meritocracy

Jean-Paul Carvalho

University of Oxford, Oxford, UK

Corresponding author. Email: jean-paul.carvalho@economics.ox.ac.uk

(Received 1 October 2020; revised 29 May 2021; accepted 30 May 2021; first published online 7 July 2021)

Abstract

Critiques of the meritocracy have centered on its narrow definition and biased assessment of merit, its stigmatization of the unsuccessful, and excessive competition. This paper identifies a different mechanism that could have pernicious social and political consequences. Economic mobility sorts people based on certain ‘productive’ traits, separating them into classes, and thus alters social externalities. This sorting–separation–externalities mechanism can produce between-class polarization in social outcomes (e.g. alcoholism and drug abuse) and worsen aggregate outcomes over all classes, consistent with rising ‘deaths of despair’ in the United States (Case and Deaton, 2020, *Deaths of Despair and the Future of Capitalism*. Princeton University Press). When traits are endogenous, transition out of a caste-based society produces an initial burst of economic mobility which dissipates over time. Thus, a dynamic meritocratic society devolves into a static class-based society. I set out an alternative model called the ‘experimental society’, which is less susceptible to these problems.

Keywords: Austrian economics; cultural evolution; meritocracy; social economics; social mobility; time preferences

JEL codes: Z13; P00

1. Introduction

To understand contemporary social problems in the United States, one should start with the transition to intensive economic growth *circa* 1,800, beginning with Britain and the Low Countries, followed by the rest of Western Europe, the United States, Canada, Australia, and New Zealand. The growth take-off around the Industrial Revolution boosted the emerging economic, military, and political power of the Western world *vis-à-vis* other power centers in the Islamic World and Asia (Jones, 2003; Kuran, 2010). The causes of this Great Divergence are much debated (e.g. Hoffman, 2017; McCloskey, 2010), with the dominant theory today being the institutional view of North (1981, 1990), Acemoglu and Robinson (2005, 2012), and others. There is consensus, however, that the transition to intensive growth was real and not a narrow phenomenon of rising real incomes, but a more holistic form of economic development in which literacy, life expectancy, heights, caloric intake, and much else rose in lockstep with real GDP (Fogel, 2004a; Weil, 2014). Fogel (2004b) has described this as a technophysio revolution. For example, the poor in the modern period are taller, more literate, and live more than twice as long as the rich in the pre-modern period (Clark, 2008). There was also major political and social changes with democratic revolutions (Acemoglu and Robinson, 2005), state formation (Tilly, 1992), growing social tolerance (Johnson and Koyama, 2019), and the ascent of the bourgeoisie at the expense of traditional hereditary elites (McCloskey, 2010).

The social disruption caused by this Great Transformation did not go unnoticed. *The Communist Manifesto* (Marx and Engels, 1992) described the severing of local feudal and patriarchal obligations, leaving ‘no other nexus between man and man than naked self-interest’, leading to ‘naked, shameless,

© The Author(s), 2021. Published by Cambridge University Press on behalf of Millennium Economics Ltd.

direct, brutal exploitation’ (p. 5), and exposing workers to ‘all the vicissitudes of competition, to all the fluctuations of the market’ (p. 9). The conversion of human beings to workers in the industrial class system was thought to produce various social and psychological effects which Marx and Engels (2009) referred to as ‘alienation’. Emile Durkheim proposed a different theory of alienation in which rapid social change, especially industrialization and rural–urban migration, brought about a state of normlessness that he called *anomie* (Durkheim, 2002). This is a concept with wide-ranging explanatory power. For example, Egyptian jihadists in the 1970s were often rural migrants to the peripheral suburbs of Cairo who had been dislocated from their family and village networks (Ibrahim, 1980). Karl Polanyi describes the process of modern economic development as a singular event insofar as it tore apart traditional communities and created a ‘market society’: ‘Instead of the economy being embedded in social relations, social relations are embedded in the economic system’ (Polanyi, 2001: 60). Over the 20th century, these concerns were the impetus for both socialist revolutions and the development of social democratic models around the welfare state, mass public education, public healthcare, and governmental bodies that regulate markets. By the century’s end, it was believed that the modern economic revolution was politically stable, even Pareto-improving, and spreading globally.

Underpinning the modern transformation of Europe was a new ideological system, so natural to contemporary readers that its profound effects are hidden in plain sight. Traditional hereditary elites were not just surpassed financially, but their status and legitimacy were also heavily undermined. The divine right of kings and aristocratic paternalism, or *noblesse oblige*, gave way to a radically different ideology – the meritocracy – that is, the allocation of power, prestige, and privilege based on merit, not birth. The concept animated Enlightenment thinking, the bourgeois revolution, and many of the social democratic institutions developed over the 20th century, including the expansion of public education and public sector employment. It is exemplified by the protagonist’s exclamation to the Spanish Count in *The Marriage of Figaro*, written in 1778 by Pierre Beaumarchais:

‘What have you done to deserve such advantages? Put yourself to the trouble of being born – nothing more. For the rest – a very ordinary man!! Whereas I, lost among the obscure crowd, have had to deploy more knowledge, more calculation and skill merely to survive than has sufficed to rule all the provinces of Spain for a century!’ (Beaumarchais, 1964: Act V, 199)

Based on this monolog, Louis XVI forbade the public performance of the play. But, this could hardly forestall the Enlightenment and the decline of hereditary rule in Europe. The Enlightenment thinkers were themselves inspired by the Confucian concept of meritocracy dating from the 6th century BC. Confucian philosophy, more broadly, shaped state institutions from the Han dynasty (206 BC to 220 AD) until the end of the Qing Dynasty in 1911, before reemerging in China in the 21st century (Bell, 2016; Jiang, 2018). Confucian meritocracy was embodied in the system of Imperial civil service examinations that every government official had to pass. Success in the exam indicated not only intelligence and education, but also virtue. That is, the exam required knowledge of Confucian classics, which were thought to produce gentleman who would serve as incorruptible officials (Jiang, 2018). Thus, from the beginning, the notion of meritocracy was tied to education and bureaucracy, and imbued with moral significance.

A system based on merit is self-evidently superior to one based on birth, and is an essential component of inclusive institutions (Acemoglu and Robinson, 2012). Nevertheless, two fundamental critiques of the current system have emerged and should be carefully considered. First, the very idea of meritocracy is being reexamined, as are the institutions through which it is expressed. The term was originally coined as a pejorative by Michael Young in his prescient (fictionalized) book *The Rise of the Meritocracy* in 1958 (Young, 2017). His vision of the transition from caste to class was a dystopian one in which the new cognitive elite, identified by intelligence tests and education, would demoralize and oppress the ‘undeserving’. Because the meritocracy created social divisions that could be justified, it could also exaggerate inequality (see Bénabou and Tirole, 2006). Of course, the old caste system was also justified, but by custom and religion rather than merit, and hereditary elites would scarcely have felt any less conceit or vindication, before the rise of meritocratic thinking. Despite its pejorative

origins, it was a positive interpretation of meritocracy that took hold, beginning with Bell (1972). Recently, attention has returned to the negative aspects, not of merit *per se*, but of the current system of meritocracy (Goodhart, 2020; Markovits, 2020; Sandel, 2020). I identify five main objections:

- (1) *Narrow definition*: Merit is defined in terms of intelligence and education, devaluing the skill required in manual and care work (Goodhart, 2020). The emphasis on education and written examinations may also favor certain types of intelligence and character traits, such as conformity and rule-following, over sometimes more valuable traits such as originality and creativity.
- (2) *Biased assessment*: Gatekeepers may have biases in assessing merit, however defined including racial, ethnic, religious, and gender biases (e.g. Bordalo *et al.*, 2016; Sarsons *et al.*, 2021). Individual choice might also be influenced by what is considered ‘normal’ or ‘appropriate’ based on an individual’s identity (e.g. gender and race), which could in turn be shaped by historical biases or underrepresentation (Carvalho and Pradelski, 2020).
- (3) *Uneven playing field* One’s success in schooling is determined not only by individual effort and characteristics, but also by socioeconomic factors such as poverty, school quality, and peer effects, which are beyond an individual’s control (Arrow *et al.*, 2018; Sandel, 2020). Hence what is deemed to be merit may be partly inherited socioeconomic advantages.
- (4) *Stigma*: The losers in the contest for high-status positions among the cognitive elite are stigmatized and demoralized by the perception that their low status is justified (Sandel, 2020). This is especially pernicious when outcomes are due to factors beyond an individual’s control (see points 2 and 3 above).
- (5) *Excessive competition*: Even the winners may lose from meritocratic competition through the costs of excessive competition (Markovits, 2020). This is especially so when the number of elite positions is close to being fixed and independent of investments in education. We know from contest theory that it can be more efficient to allocate positions at random, than based on effort, in such environments (see Skaperdas, 1996). See also Frank (2012) on excessive consumption and status competition.

The second development is that a number of social outcomes are no longer moving in lockstep with economic growth (Lima de Miranda and Snower, 2020). This is a marked departure from the modern growth experience. Most notably, in their recent book *Deaths of Despair*, Case and Deaton (2020) show that life expectancy in the United States started falling in 2015. This trend was due primarily to rising mortality rates for middle-aged whites, due to suicides, drug overdoses, and alcoholic liver disease, the so-called deaths of despair. On closer inspection, this trend is concentrated among non-Hispanic whites without a college degree. Unlike the college educated, this group experienced a range of negative economic and social shocks, due to automation, globalization, and changes in social norms. Between 1979 and 2017, median earnings of white men without a 4-year college degree fell by 13%. Job security and quality also declined. In addition, marriage rates fell for this sociodemographic group, and rates of cohabitation and out-of-wedlock births rose. Case and Deaton attribute this decoupling of GDP growth and social outcomes for those without a college degree to dysfunction in the US healthcare system and the shift in market and political power from labor to capital through rent-seeking. They also note the stigma attached to the uneducated in the exaggerated US system of meritocracy: ‘A four-year degree has become *the* key marker of social status, as if there were a requirement for nongraduates to wear a circular scarlet badge bearing the letters *BA* crossed through by a diagonal red line’ (Case and Deaton, 2020: 3).

In this paper, I hypothesize that both the emerging disaffection with meritocracy and the polarization in social outcomes between the college educated and uneducated in the United States are due to a common, but hitherto neglected, side-effect of the meritocratic system. Economic mobility – the hallmark of meritocracy – sorts people based on certain traits that are conducive to achieving ‘merit’, socially separates them into sorted groups, and thus alters peer effects and other social externalities within groups. This sorting–separation–externalities (SSE) mechanism can create a connection

between economic and social outcomes that has important consequences. The precise sign and magnitude of the correlation depends on the particular economic system and definition of merit in operation. In the archetypal meritocratic system, based on education and passing examinations over long periods of time, it could be that the traits conducive to success are non-cognitive ones such as patience, self-control, conscientiousness, rule-following, and conformity, as proposed by Hopkins (2019). Hopkins focuses on self-control, using the Gul and Pesendorfer (2001, 2004) representation. I am agnostic as to the precise set of traits that is selected for by the SSE mechanism. In the model, however, I focus on present bias. An individual subject to present bias discounts the future more heavily, so spends more on immediate gratification and less on investment in future rewards.

Consider a model in which individuals are born into one of two classes H and L . Individuals can move into (or remain in) the H class by making a costly investment (e.g. college education). Initially, both classes start out with the same distribution of time preferences. However, only those with low present bias make the investment, so that the final distribution of present bias in class L first-order stochastically dominates the final distribution in class H . After this, individuals choose a (continuous) social behavior which can be more or less unhealthy (e.g. drug abuse and excessive consumption of alcohol). After social decisions are made, individuals experience negative social externalities based on the degree of unhealthy behavior in their class. Several important effects on social outcomes emerge from the initial sorting based on economic investments. First, because individuals with greater present bias engage in unhealthier behavior, there is polarization in social outcomes between classes. The H class engages in the lower levels of unhealthy behavior than when there is no sorting, while the L class engages in more. This is one example of the SSE mechanism at work. That social outcomes are anticipated when making economic investments can further sharpen selection based on time preferences. Second, the aggregate level of unhealthy behavior in the population (over both classes) can rise under certain conditions, which we characterize. Third, the initial burst of economic mobility created by the move from a static caste-based society to a meritocratic society dissipates over time due to sorting based on time preferences combined with intergenerational transmission of traits. Thus, the meritocratic class system can come to resemble the old caste system. Finally, the polarization in social outcomes and even the perception of class differences in traits can have important political consequences. Recall the objection to stigmatization of those who lose out in meritocratic competition. A positive correlation between economic and social outcomes could amplify this stigma and, moreover, create less empathy for those who do not 'make it'. If less political weight is placed on their interests or their interests change dramatically, then populist movements might arise.

This concentration of negative social outcomes among the uneducated is not natural or immutable; it is the product of the particular system of meritocracy that has emerged in the United States in particular, based around education and examinations. I shall discuss an alternative narrative called the 'experimental society' in which success is still based on one's own actions, not birth, but in which failure does not hold the same moral significance. This conception of society is rooted in the work of the Austrian school of economics (Hayek, 1960, 1988; Kirzner, 2015; Menger, 1892; Mises, 1996/1949), as well as modern cultural evolutionary theory (Boyd and Richerson, 1985; Henrich, 2017) and evolutionary economics (Nelson and Winter, 1982; Simon, 1997; Young, 1998). In an experimental society, there is radical uncertainty and it might be that no individual possesses special insight or ability. Instead, knowledge accumulates at the population level through experimentation (trial-and-error individual learning) and social learning. Individual success is based on risk taking and a large measure of luck. Selection into the elite on the basis of risk preferences, which are not associated with healthy social decisions, could break the positive correlation between economic and social outcomes, and the double disadvantage of the less successful. This is likely to attenuate the stigma they face. Such a model may not be simply an alternative story we can tell about our society today, but could be closer to materializing due to developments in communications technology and artificial intelligence and the reliance on innovation for economic growth.

Before proceeding, let us mention the following related work. Although the paper's application is restricted to Case and Deaton's findings about non-Hispanic whites in the United States, the

hypothesis I propose bears some resemblance to William Julius Wilson's thesis that poverty in US inner cities rose not only through urban deindustrialization in the 1970s, but also the relocation of community leaders to the suburbs following desegregation (Wilson, 2011). Relatedly, Fryer (2007) presents a model in which economic mobility can undermine cooperation in low-income neighborhoods. To be trustworthy, an individual in a low-income neighborhood might reduce investment in education, so that potential partners in social interactions know they are not economically mobile and the probability of continuing interaction with them is high. The consequent polarization between education in low- and high-income communities increases with social mobility. Fryer applies his model to explain deteriorating outcomes in black neighborhoods in the 1970s, among other things. Other work on the social costs of economic development and mobility includes Putnam (2000), Skaperdas (2003), and Rajan (2019). Skaperdas (2003) demonstrates how rising real wages can reduce welfare by lowering voluntary contributions to public goods, including social interaction and communal rituals.

This paper is more broadly related to the literature on ideas and institutions. That institutions are not physical facts, but depend on social attitudes and expectations, has been established by Searle (2005), Aoki (2007), Boettke and Fink (2011), and Greif and Mokyr (2017). In fact, ideas are sometimes seen as creative forces selecting between alternative self-enforcing institutional equilibria. Tabellini (2016) argues that ideas have had a major impact on long-term economic development. The prime example is the rising status of the bourgeoisie in Europe following the Enlightenment (McCloskey, 2010), and its close connection to the meritocracy. Another idea that is bound up with meritocracy is individualism. In a meritocracy, one is judged based on one's own achievements and not the identity and status of one's family, social circle, or ethnic group. This individualistic orientation can contribute to the atomization of society and the erosion of institutions that limit negative social externalities. It is thus critical to the operation of the SSE mechanism analyzed in this paper. Institutions can also shape ideology (Bisin and Verdier, 2017). Accordingly, the concept of meritocracy begins with an idea that one's status should depend on merit, not birth. But, meritocracy is more than that; it is a complex of ideas and institutions for selecting elites based on a particular conception of merit. Such an ideological–institutional complex can emerge in response to inefficiencies in the institutional *status quo ante*. The ideology and institutions may be mutually reinforcing for some time. But ultimately, they could set in motion forces that undermine their existence (Greif and Laitin, 2004). Accordingly, I will present a model of meritocracy that produces economic mobility for some time, with upward mobility associated with certain 'productive' traits, but which quickly becomes ossified and hence self-undermining.

I shall now set out my model of the SSE mechanism, working through time preferences, followed by a discussion of the results and their implications. I then briefly describe an alternative model, the experimental society, before concluding.

2. The SSE mechanism

As Akerlof (2020) has rightly pointed out, economic models have tended to focus on narrow and 'hard' problems over important but 'softer' ones. Yet, parsimonious models can be used to contribute to large and complex subjects such as the meritocracy and its consequences. In this section, I develop a model of the SSE mechanism which identifies a hitherto neglected (negative) consequence of meritocratic institutions.

Consider an economy consisting of a continuum of agents with unit mass divided at birth into two classes indexed by $k \in \{L, H\}$. Mass m_H is born into the high-income class H and mass m_L into the low-income class L . Economic mobility is possible through an investment (e.g. education) in entering (or remaining in) the H class. Once an agent's final class status is determined, she makes a social decision, choosing a degree of healthy or unhealthy behavior and being subject to externalities from the behavior of others in her class.

In an example of the SSE mechanism described above, working through time preferences, I show that greater economic mobility (between classes) can generate a more complete sorting of individuals based on this trait. As each agent's economic and social decisions are connected by her time

preferences, this sorting can also lead to greater inequality of social outcomes and raise aggregate levels of unhealthy behavior. In particular, I assume time preferences take the β, δ form (Laibson, 1997; Phelps and Pollak, 1968; Pollak, 1968). Future payoffs are discounted by the per-period factor δ in the standard exponential manner, with an additional discount factor β applied to all future payoffs. That is, when making a decision at time T , payoffs received at any time $t > T$ are each multiplied by the factor β , which is called the degree of present bias. This makes preferences time-inconsistent. To limit notation, we set $\delta = 1$ and focus on the degree of present bias β , which we assume varies across agents.

The initial distribution of β in each class k is identical and given by the c.d.f. F , which is continuous and strictly increasing on $(\underline{\beta}, 1]$, where $\underline{\beta} \in (0, 1)$ and $F(\underline{\beta}) = 0$. Any difference in present bias across classes will emerge endogenously through mobility between classes.

2.1 Timing

Interactions take place as follows:

Date 1. Each agent makes an investment $x = 1$ (e.g. college education), or chooses not to do so $x = 0$. The cost of investment is $c > 0$ for those born into the H class and $c + d$ for those born into the L class, where $d > 0$ is the disadvantage faced by L class members. The cost of investment is paid immediately.

Date 2. If $x = 1$, the agent ends up in the H class and receives an economic payoff of $w > c$. If $x = 0$, she ends up in the L class and receives a payoff normalized to zero.

Date 3. Each agent chooses a social action $s \in R^+$, where higher values of s denote more unhealthy choices. Unhealthy choices yield immediate gratification, but at a deferred cost (e.g. overconsumption of alcohol or drugs). The immediate payoff from social action s is bs , where $b > 0$.

Date 4. A member of class k who chooses social action s , bears a cost of $(1/2)s^2 + s_k^\sigma$ where s_k^σ is the mean social action among k members and σ is a positive constant. Hence individuals not only bear the costs of their own choices, but also negative externalities are generated by the unhealthy choices of others in their class.

There are many ways to micro-found this externality. One example is through social influence as follows. At date 3, each agent chooses a social action s as described, at cost $(1/2)s^2$. At date 4, social action occurs again, except now under social influence. Specifically, each individual takes the average social decision chosen in her class k at date 3, at cost s_k^σ (i.e. $\sigma = 2$).

2.2 Social and economic decisions

Economic decisions are made at date 1 and social decisions at date 3. Working backward from date 3, an individual who ends up in class k chooses s as follows:

$$\max_{s \in \mathbb{R}^+} bs - \beta \left(\frac{1}{2} s^2 + s_k^\sigma \right). \tag{1}$$

The maximizer is

$$s^* = b/\beta \tag{2}$$

which is independent of k .

Now at date 1, an agent who starts out in class H receives the following payoff from investing $x = 1$:

$$\begin{aligned} u(1; H) &= \beta \left[w + bs^* - \frac{1}{2} (s^*)^2 - s_H^\sigma \right] - c \\ &= \beta \left[w + b^2 \left(\frac{\beta - 1/2}{\beta^2} \right) - s_H^\sigma \right] - c. \end{aligned} \tag{3}$$

An agent who starts out in class H receives the following payoff from not investing $x = 0$:

$$U(0; H) = \beta \left[b^2 \left(\frac{\beta - 1/2}{\beta^2} \right) - s_L^\sigma \right]. \tag{4}$$

Hence such an agent chooses to invest if

$$\beta[w + (s_L^\sigma - s_H^\sigma)] \geq c. \tag{5}$$

This defines a cutoff β_H such that an agent who starts out in class H invests if and only if

$$\beta \geq \frac{c}{w + (s_L^\sigma - s_H^\sigma)} \equiv \beta_H \tag{6}$$

Hence only agents with low present bias (high β) invest. As we shall see, this sorting has an effect on social outcomes. In particular, $s_L \geq s_H$. This means the assumption that $w > c$ is sufficient to guarantee that $\beta_H < 1$, so that a positive mass of agents ends up in the H class.

The equivalent payoffs for agents who start out in class L are:

$$u(1; L) = \beta \left[w + b^2 \left(\frac{\beta - 1/2}{\beta^2} \right) - s_H^\sigma \right] - (c + d)$$

$$u(0; L) = \beta \left[b^2 \left(\frac{\beta - 1/2}{\beta^2} \right) - s_L^\sigma \right].$$

Hence such an agent chooses to invest if

$$\beta[w + (s_L^\sigma - s_H^\sigma)] \geq c + d. \tag{7}$$

An agent who starts out in class L invests if and only if

$$\beta \geq \frac{c + d}{w + (s_L^\sigma - s_H^\sigma)}. \tag{8}$$

Again, this defines a cutoff

$$\beta_L = \min \left\{ \frac{c + d}{w + (s_L^\sigma - s_H^\sigma)}, 1 \right\}. \tag{9}$$

We assume d is such that $\beta_L > \underline{\beta}$, so that a positive mass of agents ends up in the L class. Hence, once again, there is sorting based on present bias.

2.3 Equilibrium

As evident from equations (6) and (9), agents choosing whether to invest at date 1 must form a conjecture about the composition of each class at date 2, in order to compute the social externalities s_L and s_H at date 4. In equilibrium, their investment decisions must fulfill the conjectures on which they were

based. Specifically, the equilibrium is given by the following system:

$$s_H = b \frac{m_H \int_{\beta_H}^1 (1/\beta) dF + m_L \int_{\beta_L}^1 (1/\beta) dF}{m_H(1 - F(\beta_H)) + m_L(1 - F(\beta_L))} \tag{10}$$

$$s_L = b \frac{m_H \int_{\underline{\beta}}^{\beta_H} (1/\beta) dF + m_L \int_{\underline{\beta}}^{\beta_L} (1/\beta) dF}{m_H F(\beta_H) + m_L F(\beta_L)}, \tag{11}$$

noting that β_L and β_H are functions of s_L and s_H .

Proposition 1. *There exists a subgame perfect equilibrium (SPE) in terms of investment cutoffs β_L and β_H and social decisions $s^*(\beta) = b/\beta$.*

All proofs are relegated to the Appendix.

There can be multiplier effects as follows. As high β agents invest in entering the H class, the social externalities in the L class worsen. This can induce a larger mass of agents to invest in order to avoid these externalities. Hence, anticipation of social outcomes when making economic investments can sharpen selection based on time preferences.

2.4 Social outcomes

Economic mobility produces sorting based on time preferences. To examine the social impact of this, compare equilibrium social outcomes to the counterfactual case in which there is no sorting and the distribution of β in each class is the initial distribution F . The mean social decision in the no-sorting case is

$$S = b \int_{\underline{\beta}}^1 \frac{1}{\beta} dF. \tag{12}$$

Thus sorting produces the following effect:

Proposition 2. *In every SPE:*

$$s_L > S > s_H.$$

Hence economic mobility produces more healthy social decisions among the H class and less healthy social decisions among the L class, compared to the counterfactual no-sorting case.

The question remains of whether aggregate social outcomes across the two classes are better than those in the no-sorting case. Although the average social decision in the population is the same regardless of sorting, the aggregate social externality depends on how the population is divided into classes. In equilibrium, the aggregate social externality is

$$\bar{s}(\sigma) = [m_H F(\beta_H) + m_L F(\beta_L)]s_H^\sigma + [m_H(1 - F(\beta_H)) + m_L(1 - F(\beta_L))]s_L^\sigma. \tag{13}$$

In the counterfactual no-sorting case, it is simply S^σ . This yields the following result:

Proposition 3. *In every SPE, the aggregate social externality is larger in equilibrium than that in the no-sorting case, that is*

$$\bar{s}(\sigma) > S^\sigma,$$

if and only if $\sigma > 1$.

Hence the sorting produced by economic mobility can worsen aggregate social outcomes when externalities are convex, that is, when extremely unhealthy decisions have a greater effect on others. In this case, the social gain from mobility in terms of producing healthier behavior in the H class is outweighed by the loss in terms of less healthy behavior in the L class. Recall that the condition $\sigma > 1$ is satisfied in the social influence example mentioned above, in which $\sigma = 2$. Another mechanism that can lead to convexity is the greater social visibility of extreme actions. For example, drinking behavior by an alcoholic could be more visible than that of a moderate drinker and may thus have greater social influence. The only empirical evidence I am aware of on this subject is presented by Hammond and Ornstein (2014). They show that individuals with extremely high body mass index (BMI) can raise the perception of mean BMI and thereby increase actual BMI through social influence.

Case and Deaton (2020) attribute declining longevity and rising deaths of despair among non-Hispanic white Americans without a college degree to exogenous changes that have made life harder for the less educated and less successful. In particular, they point to automation and globalization combined with dysfunction in the US healthcare system and the shift in market and political power from labor to capital through rent-seeking. My analysis suggests that rising deaths of despair may also be the unintended result of the particular meritocracy that has emerged in the United States, with the massive expansion of education and educational mobility between 1900 and 1980 (Goldin and Katz, 2008), and the role of higher education in selecting the American elite. This has reshaped society through SSE, with negative social consequences.

2.5 Policy

The adverse social outcomes in our model of meritocracy are not easily overcome through standard policy instruments. It is difficult to find a policy that allows for economic mobility without worsening social outcomes for the least well off. Consider the following policy interventions:

- (i) *Reducing the (common) cost of education c .* Educational subsidies increase the share of both classes who choose education, lowering the thresholds β_L and β_H . This can be verified by the inspection of equations (6) and (9). Hence, the L class ends up being composed of individuals with even higher present bias (lower β). Overall, it could be that social outcomes improve as fewer individuals end up in the L class, but social outcomes do get worse for the least well off (i.e. the L class) due to negative social externalities.

There is a further problem: contemporary meritocracies use education to screen individuals for entry into the elite. If c is reduced so far that everyone gets educated, then a new screening instrument would be sought which could be even less accessible and inclusive.

- (ii) *Reducing educational disadvantage d .* Similarly, an educational subsidy for individuals starting out in the L class (equivalent to a reduction in d) would lower the threshold for choosing education β_L . This can be verified by the inspection of equation (14). Again, the L class ends up being composed of individuals with even higher present bias. Hence their social outcomes worsen.
- (iii) *Taxation and redistribution.* If part of the return to education w was taxed and re-distributed to those who did not get educated, rates of education would decline. The effect would be the same as if one raised the cost of education c . The thresholds β_L and β_H would rise and social outcomes would get better for those who end up in the L class. However, the overall effect on social outcomes is ambiguous as a larger share of the population ends up in the L class.

2.6 Intergenerational mobility

Let us now examine intergenerational mobility by embedding the one-shot game described so far in an infinite-horizon model. Time is discrete and indexed by $t = 0, 1, 2, \dots$. Each period is a new

generation. For simplicity, we assume each agent cares only about her own payoffs and not about her offspring's, although the model could certainly be extended to include such concerns. Parameters are fixed for all time, only the distribution of traits in the population is endogenously determined. At time $t = 0$, the game just analyzed is played with symmetric type distributions F in each class. At the end of each period, the old generation dies and is replaced by a new generation. I assume the type distribution of the new generation $t \geq 1$ born into class k , denoted by F^t , is simply the final type distribution in class k in the previous period. This is consistent with vertical transmission of traits from parent to child (Bisin and Verdier, 2000; Cavalli-Sforza and Feldman, 1981) and/or social transmission of traits within groups (Carvalho, 2016).

Define upward mobility U^t as the share of agents born into the L class in period t that ends up in the H class. Downward mobility D^t is the share of agents born into the H class in period t that ends up in the L class. From the one-shot game, we know that $U^0 > 0$ and $D^0 > 0$, in the initial period. For all subsequent periods, we can state the following result:

Proposition 4. *There exists a subgame perfect equilibrium of the infinite-horizon game in which upward and downward mobility stall completely: $U^t = D^t = 0$ for all $t > 0$.*

Consider a caste-based society in which there is no upward or downward mobility. Because there is no selection, the distribution of traits in each caste could be more or less the same. Moving to a meritocratic society (period 0) produces a burst of economic mobility, with low present bias types entering or remaining in the H class, through investments in college education for example. But, mobility is only temporary. Sorting combined with intergenerational transmission of traits completely eliminates both upward and downward mobility in all subsequent periods. Hence the meritocratic class-based society comes to resemble the old caste-based society, only with worse social outcomes for the worst off and possibly overall.

2.7 Discussion

I shall now provide some guidance on interpreting the model and its implications.

2.7.1 Interpretation

I have illustrated the SSE mechanism through an example based on time preferences, and in particular present bias β , because it is the simplest possible case. This is consistent with Hopkin's (2019) hypothesis that college education today serves mainly to signal non-cognitive traits such as self-control. My point, however, is more general: under certain conditions the SSE mechanism creates a connection between economic and social outcomes. I am agnostic as to precisely which traits form the basis for sorting. For example, in a slightly more complicated analysis one could examine sorting based on patience, as measured by the standard exponential discount factor δ . Beyond time preferences, there is substantial evidence, surveyed by Almlund *et al.* (2011), that the Big Five personality traits predict educational attainment and achievement. In fact, they have just as much explanatory power as cognitive measures. Most notably, high conscientiousness is associated with higher levels of educational attainment and achievement. It also predicts lower levels of risky behavior, including alcohol and nicotine consumption, gambling, and unprotected sex (Dash *et al.*, 2019; Hagger-Johnson *et al.*, 2011). The SSE mechanism may also operate through traits such as rule-following and conformity, which have received less attention. Such traits may be valued by large organizations that require cooperation, teamwork, and loyalty, but are not necessarily conducive to creativity and innovation.

2.7.2 Empirical studies of economic/social mobility

An important literature examines the determinants of economic/social mobility. Chetty *et al.* (2014) find that mobility is decreasing in segregation and inequality, and increasing in social capital, family stability, and school quality. In Italy, Güell *et al.* (2018) similarly find that mobility is increasing in education and social capital, and decreasing in inequality. In addition, they find no association between mobility and other socio-political variables, such as life expectancy, crime rates, and suicide

rates. Based on the analysis in this paper, two remarks can be made. First, the SSE mechanism makes social capital, inequality, life expectancy, crime rates and suicide rates a consequence of economic mobility, as much as a cause. In fact, when social externalities are convex ($\sigma > 1$), the true causal effect of these variables on economic mobility is underestimated. For example, a rise in social mobility caused by declining crime rates can lead to a subsequent rise in crime rates via the SSE mechanism when $\sigma > 1$ (Proposition 3). Thus, cross-sectional analyses would underestimate the causal effect of low crime rates on social mobility. Given this endogeneity problem, the lack of association between mobility and life expectancy, crime, and suicide rates in Italy is not surprising. In contrast, when social externalities are concave ($0 < \sigma < 1$), economic/social mobility can improve overall social outcomes and the causal effect of social conditions on mobility will be amplified. Second, the intergenerational analysis in Section 2.6 suggests that the relationship between mobility and the variables considered by these studies depends on where the economy is in its life-cycle. Once sorting takes hold and economic mobility dissipates, the causal effect of these variables could attenuate.

2.7.3 Positive externalities from education

There are many positive externalities from education that are ignored by my model. *Inter alia*, education is associated with higher rates of economic growth (Mankiw *et al.* 1992), positive human capital spillovers (Bénabou, 1993), greater support for democracy (Glaeser *et al.*, 2007), and lower rates of smoking (De Walque, 2007) and illicit drug use (Carpenter *et al.*, 2017). When it comes to social outcomes, it is useful to distinguish between three types of effects of education. First, education can produce positive outcomes at the individual level. In the model, this would mean that an individual's education $x = 1$ lowers her own present bias (raises β). Second, education can produce positive outcomes at the class level. This would mean that an individual's education $x = 1$ lowers the present bias of all others who choose education (and end up in the *H* class). Third, education can produce positive outcomes at the population level. This would mean that an individual's education $x = 1$ lowers the present bias of all other individuals regardless of their class status. The third type of effect would offset the negative social externalities analyzed in this paper. If the effect is strong enough, it could even overwhelm the negative effects of education through the SSE mechanism. On the other hand, the first and second types of effect exacerbate the social inequality caused by the SSE mechanism. Not only would the educated be sorted according to time preferences, but there would also be a positive *treatment effect* of education, with any spillovers confined to the educated. To know more, we need further empirical work identifying the precise nature of educational spillovers.

2.7.4 Internalizing negative social externalities

Why do institutions not emerge to internalize the negative social externalities analyzed in this paper (see Coase, 1960; Demsetz, 1967)? There are institutions that have performed such a function, including the following:

- (i) *'Sin' taxes and prohibitions.* Past societies developed a wide range of prohibitions to regulate social externalities, such as usury laws (Koyama, 2010), sumptuary laws (Desierto and Koyama, 2020), inheritance laws (Kuran, 2010), restrictions on disposal of property (entails), and most relevantly bans on alcohol and drugs (Miron and Zwiebel, 1991). More recently, so-called 'sin taxes' (i.e. excise taxes) have been imposed on cigarettes and have been proposed for sugar and fast food.
- (ii) *Social norms and religious prohibitions.* Social externalities have also been regulated through informal institutions, often at the group level. For example, most societies have social and religious norms against extramarital sex (Francesconi *et al.*, 2016). Religious groups also monitor and regulate excessive drinking, drug use, and other forms of antisocial behavior by members (Iannaccone, 1992; McBride, 2007). For example, the prohibition on alcohol consumption in the United States from 1920 to 1933 was partly driven by the religious temperance movement. Perhaps as a consequence of religious prohibitions, religiosity is associated with better mental and physical health, and educational and marital outcomes (e.g. Stark, 2002).

Such formal and informal institutions could mitigate the negative social externalities that are part of the SSE mechanism. However, they clearly have not prevented the adverse social outcomes and deaths of despair documented by Case and Deaton (2020). Why not? A crucial factor is ideology. Both ‘sin taxes’ and religious prohibitions are increasingly resisted as a violation of individual freedom. This concept of individualism is part of the complex of ideas that gave birth to meritocratic institutions. The meritocracy eschews traditional collectivist notions of identity: one is not judged by the status of one’s family, social circle, or ethnic group, but rather on one’s own achievements. This emphasis on individual freedom and agency limits support for (coercive) institutions that regulate social externalities.

2.7.5 Political implications

The results of Sections 2.4 and 2.6 also hold a number of political implications. First, by stigmatizing those who do not ‘make it’, the meritocracy is said to reduce political empathy for the less educated. A positive correlation between economic and social outcomes induced by the SSE mechanism can add to this, further reducing the weight placed on their interests in policymaking. A more subtle point is that even an erroneous perception that the SSE mechanism works through traits such as self-control (as we have shown is possible) can exaggerate the decline in political empathy. This may not only be undesirable in and of itself, but could also spur populist movements. Second, independent of trends in racial and economic segregation, there has been a large increase in geographic segregation based on education in the United States since the 1940s (Domina, 2006). The sorting effect of education in terms of personality and other traits adds to the effect of geographic segregation on electoral outcomes, given the structure of the electoral college. Finally, the dynamic results suggest that meritocracy may be self-undermining. Unfortunately, it could even be that those without a college degree come to support policies that reduce educational access and economic mobility. That is, regrettably, if a new model is not found, the old caste-system may not be entirely behind us.

3. An alternative model: the experimental society

A single economic (and moral) hierarchy defined in terms of education is not the only alternative to the various social systems in which one’s status is inherited. We do not know what will replace the current meritocracy, if anything does, nor do we know what should. Nevertheless, we can speculate on alternative conceptions of society that may not suffer from the problems of meritocracy described above. In particular, I will briefly set out what I call ‘the experimental society’, which is rooted in Austrian economics, especially the work of Friedrich Hayek (see Boettke, 2018). It is no coincidence that Hayek (1960: 94–99, 387–388) prefigured some of the objections to meritocracy, stating: ‘This means an official ranking of people into a hierarchy, with the certified genius on top and the certified moron at the bottom, a hierarchy made worse by the fact that it is presumed to express “merit” and will determine access to the opportunities in which value can show itself’ (p. 387).

The experimental society has different criteria for allocating resources, and a different conception of itself. It is in the first instance a ‘narrative’, that is, a conceptual model of the economy and society in which we live that has at least a grain of truth (Akerlof *et al.*, 2020; Mullainathan *et al.*, 2008). Moreover, if widely adopted as a working model, the narrative can bring about changes in institutions that make it a better descriptor of society. We describe some potential institutional changes below. This self-fulfilling property of the ‘experimental society’ narrative is consistent with institutions as ‘social facts’ that depend on social attitudes and expectations (Aoki, 2007; Boettke and Fink, 2011; Greif and Mokyr, 2017; Searle, 2005). For reasons I shall discuss, such a society might avoid a positive correlation between economic and social outcomes, and its attendant problems. It could also be better suited to the emerging technological environment.

Continued economic growth and development in advanced countries such as the United States depend on innovation, that is, new forms and means of utilizing knowledge. The meritocratic system, originally bound to bureaucracy and the expansion of state capacity, is not ideally suited to this.

Widespread public education was introduced in 19th century Europe as part of nation-building programs. The goal was not to foster innovation and disseminate knowledge, but rather to replace traditional regional, class, and ethnic identities with a national identity spearheaded by a new bureaucratic elite (Weber, 1976). The new elites received rents that motivated them to properly perform the expanding executive functions of the state. For this new political system to be stable, rents enjoyed by members of the new bureaucratic elite had to be considered legitimate by the citizenry at large. For example, Hoffman *et al.* (1994) show that individuals are willing to tolerate greater inequality in experimental games when the decisive position is allocated based on performance in a test, rather than at random. This notion of desert runs counter to philosophical notions of luck egalitarianism (Schmidtz, 2006) and might be hard-wired components of human psychology (Carvalho and Koyama, 2010). Hence the need for legitimization of the political and bureaucratic rents that hold the modern state together might be the impetus behind the narrative of meritocracy in its current form (see also Bowles and Gintis, 1976).

Meritocratic ideology supposes there is a set of individuals in society who are capable of possessing superior knowledge and character, and who can be identified based on education and examinations. This happens to gel with standard economics with its focus on closed systems in which all unknowns are known unknowns. For example, it is usually assumed that individuals know their own strategy set, the strategy set of all other individuals, and the mapping from strategy profiles to their own payoffs. In addition, the probability distribution over states and player types is common knowledge. Innovation, however, takes place in a radically different environment of ignorance and Knightian uncertainty (Knight, 1921). It is difficult to identify who will make the largest contributions beforehand and knowledge is only partially generated through rational, deliberative procedures. More often, it is discovered through experimentation and social learning. This idea is most famously expressed by Hayek:

If we possess all the relevant information, if we can start out from a given system of preferences and if we command complete knowledge of available means, the problem which remains is purely one of logic. [...] This, however, is emphatically not the economic problem which society faces. And the economic calculus which we have developed to solve this logical problem, though an important step toward the solution of the economic problem of society, does not yet provide an answer to it. The reason for this is that the 'data' from which the economic calculus starts are never for the whole society 'given' to a single mind which could work out the implications, and can never be so given. Hayek (1945: 519–520, emphasis in original)

An alternative approach which accounts for the incomplete, fragmented, and distributed nature of knowledge is the evolutionary approach, which is consonant with early classical economics. It was more fully developed by the Austrian school, beginning with the evolutionary theory of fiat money by Menger (1892), followed by the work of Mises (1996/1949), Hayek (1960, 1988), Kirzner (2015), and others. The evolutionary approach has been formalized and extended further by modern evolutionary economics (Nelson and Winter, 1982; Simon, 1997) and evolutionary game theory (Sandholm, 2010; Young, 1998). These fields do not assume individual agents understand the system in which they interact, nor can they guide it in any meaningful way. Rather: 'Agents adapt – they are not devoid of rationality – but they are not hyper-rational. They look around them, they gather information, and they act fairly sensibly on the basis of their information most of the time' (Young, 1998: 5). The evolutionary process this generates is a form of distributed computation. Thus, rationality and knowledge are not properties of the individual, but can be thought of as emerging at the population level.

With regard to innovation, the appropriate starting point is cultural evolutionary theory (e.g. Boyd and Richerson, 1985; McElreath and Boyd, 2008). Hayek anticipates modern cultural evolutionary theory by identifying the key to progress as 'selection by imitation of successful institutions and habits', from which emerges 'ideas and skills – in short, the whole cultural inheritance which is

passed on by learning and imitation' (Hayek, 1960: 591). This is the current conception of culture as cumulative intergenerational learning, and the distinguishing characteristic of humans (Henrich, 2017). Knowledge circulates through the population through imitation and other forms of social learning, combined with the codification, storage, and intergenerational transmission of new ideas and skills. But, this would count for little without new knowledge being injected into the system through trial-and-error learning and random experimentation (Giuliano and Nunn, *forthcoming*; Rogers, 1988). This notion is extended to networked technologies by evolutionary game theory, in which idiosyncratic choice and experimentation is critical to the social adoption of new technologies (Kreindler and Young, 2014; Young, 2011). As Hayek points out repeatedly, individual learning and experimentation has often occurred not out of commercial interest, but plain curiosity (Hayek, 1960: 392). Hence for cultural evolution to proceed there must be incentives for risk taking and experimentation, as well as a tolerance of eccentricity (Harper, 2018; Witt, 2008). This approach to innovation is developed into a theory of entrepreneurial discovery and market dynamics by Israel Kirzner (2015).

The experimental society suggests some modifications to current meritocratic institutions. Under the current system, college education is the predominant screening device for entry into the elite and the university system is at the center of the organization of innovation. For some time in the United States, professional positions have required a 4-year college degree. Doctoral programs require candidates to have performed well in an undergraduate program, and increasingly in a prior Masters-level program. Even in Silicon Valley, investors place some importance on the founders' educational background when deciding whether to fund a startup. The most peculiar convention is for professional sporting teams in the United States to recruit almost exclusively from college teams, so that aspiring professional athletes must go through the university system. There are of course good reasons for the prominent role played by universities. Modern economic growth has been driven by the interplay between science and technology, beginning with the scientific and industrial revolutions (Mokyr, 2002; Mokyr, 1990). In addition, the universities perform a useful role in screening individuals *at scale* and directing opportunities toward those who have the talent and temperament to make the most of them. However, to make it through the university system and qualify for such opportunities today requires over 16 years of education, of sitting patiently and learning how to pass examinations. This strongly selects for individuals with non-cognitive traits such as patience and self-control, and may screen out highly talented and creative individuals without these traits. In the current meritocracy, the individuals who are screened out are more or less invisible. Only those who have distinguished themselves in the education system are qualified for elite positions. In the experimental society, however, where knowledge is distributed and originality key, the demotivation and loss of such individuals through the education process would be recognized. New channels for gaining access to capital and positions of influence would be opened up. This is the difference between a bureaucratic and commercial meritocracy.

To get to the experimental society requires a change in perspective and norms. In fact, we are already seeing changes along these lines. France recently closed the Ecole Nationale d'Administration which has educated the country's political and bureaucratic elite since 1945. Technology companies such as Tesla, Google, and Apple no longer require employees to have a 4-year college degree and are hiring based on performance in coding platforms and competitions. In addition, NBA teams increasingly recruit basketball players directly from high school. Indeed, developments in communications technology and artificial intelligence are pushing society in the direction of the experimental society. Through big data and machine learning, artificial intelligence has already surpassed humans in many routine cognitive tasks (more so than in some manual tasks), for which higher education used to prepare individuals. Although it is difficult to predict which areas will be immune from competition with machines even in the short term, as the advantage of artificial intelligence expands, educational institutions will have to focus on preparing students for new non-routine and creative tasks (Acemoglu and Restrepo, 2019). In addition, the internet makes possible free communication among large numbers of geographically dislocated individuals,

as well as the transmission of up-to-date knowledge at low cost to people who lack access to elite education. Although such technological developments were and are feared as enablers of authoritarian regimes (Guriev and Treisman, 2019; Hayek, 1960), they might also produce less gated, less bureaucratic, more creative and experimental forms of organization.

Such an experimental society would be less prone to the problems of the archetypal meritocracy. First and foremost, economic success would be based more on luck, risk-taking, and curiosity than traits such as patience and self-control which are associated with healthy social decisions. This would weaken the connection between economic and social outcomes. The experimental society is also less likely to ossify into a static class-based society, because it exposes individuals to fluctuations in the economic and technological environment and the inherent randomness of success based on experimentation. Second, the emphasis on experimentation in the face of radical uncertainty and knowledge emerging at the population level would remove some of the stigma faced by the less educated. They would no longer be in the shadow of an elite with superior and esoteric knowledge. Rather, it would be recognized that everyone has the capacity to produce new ideas and refine ways of doing things. Third, it could be that socioeconomic disadvantages are less relevant to innovation through experimentation than to formal education, which requires extensive knowledge of social codes. Fourth, the experimental society would be less gated, so the definition and assessment of merit are less of a concern. Of course, there is no way to solve all of these problems and the differences are a matter of degree. Increasing specialization and technological sophistication means that at least some education is required for innovation in a modern society. Many fields in the hard sciences will continue to require years of university education to master. The days of self-taught geniuses and polymaths working by themselves may be behind us. Also, innovation often depends on mastering and recombining existing knowledge, which may select for similar non-cognitive skills as patience and self-control. Thus, it may not be possible to completely neutralize the SSE mechanism, but it may be possible to weaken it by moving toward a more experimental society.

4. Conclusion

The United States today is sharply divided on the basis of college education, with the college educated experiencing significantly better economic and social outcomes. I propose that this polarization is due, in part, to the current system of meritocracy, based on education and examinations. In particular, meritocratic competition sorts individuals according to traits such as patience, self-control, conscientiousness, and rule-following, socially separates them into classes, and thus regulates social externalities. This SSE mechanism not only polarizes, but can also worsen aggregate social outcomes over all classes. The connection between bad economic and social outcomes can further reduce political empathy for those who lose out in meritocratic competition. In addition, the meritocratic system tends to be self-undermining, with economic mobility dissipating over time due to sorting and intergenerational transmission of traits. Thus, changes to the meritocratic system are required to avoid both populist movements and regressive moves toward a caste-based society. Based on Austrian economics and cultural evolutionary theory, I have proposed an alternative conception of society – the experimental society – which is less susceptible to these problems.

Acknowledgements. I am grateful for comments by the editor and referees, Daniel D'Amico, Pete Boettke, Mark Koyama, Adam Martin, Stergios Skaperdas, Virgil Storr, Cole Williams, and participants at the Sociality and Austrian Economics colloquium at Texas Tech University. All errors are mine.

References

- Acemoglu, D. and P. Restrepo (2019), 'Automation and New Tasks: How Technology Displaces and Reinstates Labor', *Journal of Economic Perspectives*, 33(2): 3–30.
- Acemoglu, D. and J. A. Robinson (2005), *Economic Origins of Dictatorship and Democracy*, Cambridge, UK: Cambridge University Press.

- Acemoglu, D. and J. A. Robinson (2012), *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*, New York, NY: Currency.
- Akerlof, G. A. (2020), 'Sins of Omission and the Practice of Economics', *Journal of Economic Literature*, **58**(2): 405–418.
- Akerlof, R., N. Matouschek and L. Rayo (2020), 'Stories at work', AEA Papers and Proceedings, vol. **110**, pp. 199–204.
- Almlund, M., A. L. Duckworth, J. Heckman, and T. Kautz (2011) 'Personality Psychology and Economics', in *Handbook of the Economics of Education* (Vol. 4), Oxford, UK: Elsevier, pp. 1–181.
- Aoki, M. (2007), 'Endogenizing Institutions and Institutional Changes', *Journal of Institutional Economics*, **3**(1): 1–31.
- Arrow, K., S. Bowles, and S. N. Durlauf eds. (2018), *Meritocracy and Economic Inequality*. Princeton, NJ: Princeton University Press.
- Beaumarchais, P. (1964), *The Marriage of Figaro* (J. Wood, Trans.), London, UK: Penguin.
- Bell, D. (1972), 'Meritocracy and Equality', *The Public Interest*, **29**: 29–68.
- Bell, D. A. (2016), *The China Model: Political Meritocracy and the Limits of Democracy*, Princeton, NJ: Princeton University Press.
- Bénabou, R. (1993), 'Workings of a City: Location, Education, and Production', *Quarterly Journal of Economics*, **108**(3): 619–652.
- Bénabou, R. and J. Tirole (2006), 'Belief in a Just World and Redistributive Politics', *Quarterly Journal of Economics*, **121**(2): 699–746.
- Bisin, A. and T. Verdier (2000), 'Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits', *Quarterly Journal of Economics*, **115**(3): 955–988.
- Bisin, A. and T. Verdier (2017) 'On the Joint Evolution of Culture and Institutions', Working paper, National Bureau of Economic Research.
- Boettke, P. J. (2018) *F.A. Hayek: Economics, Political Economy and Social Philosophy*, London, UK: Palgrave Macmillan.
- Boettke, P. J. and A. Fink (2011), 'Institutions First', *Journal of Institutional Economics*, **7**(4): 499–504.
- Bordalo, P., K. Coffman, N. Gennaioli and A. Shleifer (2016), 'Stereotypes', *Quarterly Journal of Economics*, **131**(4): 1753–1794.
- Bowles, S. and H. Gintis (1976), *Schooling in Capitalist America*, New York, NY: Basic.
- Boyd, R. and P. J. Richerson (1985), *Culture and the Evolutionary Process*, Chicago, IL: University of Chicago Press.
- Carpenter, C. S., C. B. McClellan and D. I. Rees (2017), 'Economic Conditions, Illicit Drug use, and Substance use Disorders in the United States', *Journal of Health Economics*, **52**: 63–73.
- Carvalho, J. P. (2016), 'Identity-Based Organizations', *American Economic Review Papers and Proceedings*, **106**(5): 410–414.
- Carvalho, J. P. and M. Koyama (2010), 'Instincts and Institutions: The Rise of the Market', in *The Social Science of Hayek's 'The Sensory Order'*. Bingley, UK: Emerald Group, pp. 285–309.
- Carvalho, J. P. and B. Pradelski (2020), 'Identity and Underrepresentation: Interactions between Race and Gender', Working paper, UC Irvine.
- Case, A. and A. Deaton (2020), *Deaths of Despair and the Future of Capitalism*, Princeton, NJ: Princeton University Press.
- Cavalli-Sforza, L. L. and M. W. Feldman (1981), *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton, NJ: Princeton University Press.
- Chetty, R., N. Hendren, P. Kline and E. Saez (2014), 'Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States', *Quarterly Journal of Economics*, **129**(4): 1553–1623.
- Clark, G. (2008), *A Farewell to Alms: A Brief Economic History of the World*, Princeton, NJ: Princeton University Press.
- Coase, R.H. (1960), 'The Problem of Social Cost', in C. Gopalakrishnan (ed.), *Classic Papers in Natural Resource Economics*, New York, NY: Springer, pp. 87–137.
- Dash, G. F., W. S. Slutske, N. G. Martin, D. J. Statham, A. Agrawal and M. T. Lynskey (2019), 'Big Five Personality Traits and Alcohol, Nicotine, cannabis, and Gambling Disorder Comorbidity', *Psychology of Addictive Behaviors*, **33**(4): 420.
- Demsetz, H. (1967), 'Toward a Theory of Property Rights', *American Economic Review*, **1**(2): 347–359.
- Desierto, D. and M. Koyama (2020), 'The Political Economy of Status Competition: Sumptuary Laws in Preindustrial Europe', CEPR Discussion Papers.
- De Walque, D. (2007), 'Does Education Affect Smoking Behaviors?: Evidence Using the Vietnam Draft as an Instrument for College Education', *Journal of Health Economics*, **26**(5): 877–895.
- Domina, T. (2006), 'Brain Drain and Brain Gain: Rising Educational Segregation in the United States, 1940–2000', *City & Community*, **5**(4): 387–407.
- Durkheim, E. (2002), *Suicide: A Study in Sociology*, Abingdon, UK: Routledge Classics.
- Fogel, R. W. (2004a) *The Escape From Hunger and Premature Death, 1700–2100: Europe, America, and the Third World* (Vol. 38), Cambridge, UK: Cambridge University Press.
- Fogel, R. W. (2004b), 'Technophysio Evolution and the Measurement of Economic Growth', *Journal of Evolutionary Economics*, **14**(2): 217.
- Francesconi, M., C. Ghiglino and M. Perry (2016), 'An Evolutionary Theory of Monogamy', *Journal of Economic Theory*, **166**: 605–628.

- Frank, R. H. (2012), *The Darwin Economy: Liberty, Competition, and the Common Good*, Princeton, NJ: Princeton University Press.
- Fryer, R. G. (2007), 'A Model of Social Interactions and Endogenous Poverty Traps', *Rationality and Society*, **19**(3): 335–366.
- Giuliano, P. and N. Nunn (forthcoming), 'Understanding Cultural Persistence and Change', *Review of Economic Studies*.
- Glaeser, E. L., G. A. Ponzetto and A. Shleifer (2007), 'Why Does Democracy Need Education?', *Journal of Economic Growth*, **12**(2): 77–99.
- Goldin, C. and L. F. Katz (2008), *The Race Between Education and Technology*, Boston, MA: Belknap Press.
- Goodhart, D. (2020) *Head, Hand, Heart: Why Intelligence is Over-Rewarded, Manual Workers Matter, and Caregivers Deserve More Respect*, New York, NY: Free Press.
- Greif, A. and D. D. Laitin (2004), 'A Theory of Endogenous Institutional Change', *American Political Science Review*, **98**(4): 633–652.
- Greif, A. and J. Mokyr (2017), 'Cognitive Rules, Institutions, and Economic Growth: Douglass North and Beyond', *Journal of Institutional Economics*, **13**(1): 25–52.
- Güell, M., M. Pellizzari, G. Pica and J. V. Rodríguez Mora (2018), 'Correlating Social Mobility and Economic Outcomes', *The Economic Journal*, **128**(612): F353–F403.
- Gul, F. and W. Pesendorfer (2001), 'Temptation and Self-Control', *Econometrica*, **69**(6): 1403–1435.
- Gul, F. and W. Pesendorfer (2004), 'Self-control and the Theory of Consumption', *Econometrica*, **72**(1): 119–158.
- Gurieff, S. and D. Treisman (2019), 'Informational Autocrats', *Journal of Economic Perspectives*, **33**(4): 100–127.
- Hagger-Johnson, G., B. M. Bewick, M. Conner, D. B. O'Connor and D. Shickle (2011), 'Alcohol, Conscientiousness and Event-Level Condom Use', *British Journal of Health Psychology*, **16**(4): 828–845.
- Hammond, R. A. and J. T. Ornstein (2014), 'A Model of Social Influence on Body Mass Index', *Annals of the New York Academy of Sciences*, **1331**(1): 34–42.
- Harper, D. A. (2018), 'Innovation and Institutions from the Bottom up: An Introduction', *Journal of Institutional Economics*, **14**(6): 975–1001.
- Hayek, F. A. (1945), 'The use of Knowledge in Society', *American Economic Review*, **35**(4): 519–530.
- Hayek, F. A. (1960), *The Constitution of Liberty*, Chicago, IL: University of Chicago Press.
- Hayek, F. A. (1988), *The Fatal Conceit*, Chicago, IL: University of Chicago Press.
- Henrich, J. (2017), *The Secret of our Success: How Culture is Driving Human Evolution, Domesticating our Species, and Making us Smarter*, Princeton, NJ: Princeton University Press.
- Hoffman, P. T. (2017), *Why did Europe Conquer the World?*, Princeton, NJ: Princeton University Press.
- Hoffman, E., K. McCabe, K. Shachar and V. Smith (1994), 'Preferences, Property Rights, and Anonymity in Bargaining Games', *Games and Economic Behavior*, **7**(3): 346–380.
- Hopkins, E. (2019) 'College as a Signal of Self-Control: Self-Control Preferences in a High Temptation environment', Working paper, University of Edinburgh.
- Iannaccone, L. R. (1992), 'Sacrifice and Stigma: Reducing Free-Riding in Cults, Communes, and Other Collectives', *Journal of Political Economy*, **100**(2): 271–291.
- Ibrahim, S. E. (1980), 'Anatomy of Egypt's Militant Islamic Groups: Methodological Note and Preliminary Findings', *International Journal of Middle East Studies*, **12**(4): 423–453.
- Jiang, Y. H. (2018), 'Confucian Political Theory in Contemporary China', *Annual Review of Political Science*, **21**: 155–173.
- Johnson, N. D. and M. Koyama (2019), *Persecution & Toleration: The Long Road to Religious Freedom*, Cambridge, UK: Cambridge University Press.
- Jones, E. (2003), *The European Miracle*, Cambridge, UK: Cambridge University Press.
- Kirzner, I. M. (2015), *Competition and Entrepreneurship*, Chicago, IL: University of Chicago press.
- Knight, F. H. (1921), *Risk, Uncertainty and Profit*, Boston, MA: Houghton Mifflin.
- Koyama, M. (2010), 'Evading the "Taint of Usury": The Usury Prohibition as Barrier to Entry', *Explorations in Economic History*, **47**(4): 420–442.
- Kreindler, G. E. and H. P. Young (2014), 'Rapid Innovation Diffusion in Social Networks', *Proceedings of the National Academy of Sciences*, **111**(Supplement 3): 10881–10888.
- Kuran, T. (2010), *The Long Divergence: How Islamic Law Held Back the Middle East*, Princeton, NJ: Princeton University Press.
- Laibson, D. (1997), 'Golden Eggs and Hyperbolic Discounting', *Quarterly Journal of Economics*, **112**(2): 443–478.
- Lima de Miranda, K. and D. J. Snower (2020), 'Recoupling Economic and Social Prosperity', *Global Perspectives*, **1**(1): 118–167.
- Mankiw, N. G., D. Romer and D. N. Weil (1992), 'A Contribution to the Empirics of Economic Growth', *Quarterly Journal of Economics*, **107**(2): 407–437.
- Markovits, D. (2020), *The Meritocracy Trap: How America's Foundational Myth Feeds Inequality, Dismantles the Middle Class, and Devours the Elite*, New York, NY: Penguin Books.
- Marx, K. and F. Engels (1992), *The Communist Manifesto*, Oxford, UK: Oxford University Press, Originally published in 1848.

- Marx, K. and F. Engels (2009), *The Economic and Philosophic Manuscripts of 1844 and the Communist Manifesto*, New York, NY: Prometheus Books.
- McBride, M. (2007), 'Club Mormon: Free-Riders, Monitoring, and Exclusion in the LDS Church', *Rationality and Society*, **19** (4): 395–424.
- McCloskey, D. N. (2010), *Bourgeois Dignity: Why Economics Can't Explain the Modern World*, Chicago, IL: University of Chicago Press.
- McElreath, R. and R. Boyd (2008), *Mathematical Models of Social Evolution: A Guide for the Perplexed*, University of Chicago Press.
- Menger, C. (1892), 'On the Origin of Money', *The Economic Journal*, **2**(6): 239–255.
- Miron, J. A. and J. Zwiebel (1991), 'Alcohol Consumption During Prohibition', *American Economic Review*, **81**(2): 242–247.
- Mises, L.v. (1996/1949), *Human Action*. Auburn, Alabama: Ludwig von Mises Institute.
- Mokyr, J. (2002), *The Gift of Athena: Historical Origins of the Knowledge Economy*, Princeton, NJ: Princeton University Press.
- Moykr, J. (1990), *The Lever of Riches*, Oxford, UK: Oxford University Press.
- Mullainathan, S., J. Schwartzstein and A. Shleifer (2008), 'Coarse Thinking and Persuasion', *Quarterly Journal of Economics*, **123**(2): 577–619.
- Nelson, R. R. and S. G. Winter (1982), *An Evolutionary Theory of Economic Change*, Cambridge, MA: Harvard University Press.
- North, D. C. (1981), *Structure and Change in Economic History*, New York, NY: Norton.
- North, D. C. (1990), *Institutions, Institutional Change, and Economic Performance*, Cambridge, UK: Cambridge University Press.
- Phelps, E. S. and R. A. Pollak (1968), 'On Second-Best National Saving and Game-Equilibrium Growth', *Review of Economic Studies*, **35**(2): 185–199.
- Polanyi, K. (2001), *The Great Transformation*, Boston, MA: Beacon Press, Originally published in 1944.
- Pollak, R. A. (1968), 'Consistent Planning', *Review of Economic Studies*, **35**(2): 201–208.
- Putnam, R. D. (2000), *Bowling Alone: The Collapse and Revival of American Community*, New York, NY: Simon and Schuster.
- Rajan, R. (2019), *The Third Pillar: How Markets and the State Leave the Community Behind*, New York, NY: Penguin.
- Rogers, A. R. (1988), 'Does Biology Constrain Culture?', *American Anthropologist*, **90**(4): 819–831.
- Sandel, M. J. (2020), *The Tyranny of Merit: What's Become of the Common Good?*, New York, NY: Farrar, Straus and Giroux.
- Sandholm, W. H. (2010), *Population Games and Evolutionary Dynamics*, Cambridge, MA: MIT Press.
- Sarsons, H., K. Gërkhani, E. Reuben and A. Schram (2021), 'Gender Differences in Recognition for Group Work', *Journal of Political Economy*, **129**(1): 101–147.
- Schmidtz, D. (2006), *The Elements of Justice*, Cambridge, UK: Cambridge University Press.
- Searle, J. R. (2005), 'What is an Institution?', *Journal of Institutional Economics*, **1**(1): 1–22.
- Simon, H. A. (1997), *Models of Bounded Rationality: Empirically Grounded Economic Reason*, Cambridge, MA: MIT press.
- Skaperdas, S. (1996), 'Contest Success Functions', *Economic Theory*, **7**(2): 283–290.
- Skaperdas, S. (2003), *Turning 'Citizens' into 'Consumers': Economic Growth and the Level of Public Discourse*, Cambridge, UK: Cambridge University Press, pp. 30–43.
- Stark, R. (2002), *America's Blessings: How Religion Benefits Everyone, Including Atheists*, West Conshohocken, PA: Templeton Foundation Press.
- Tabellini, G. (2016), 'Ideas or Institutions? – A Comment', *Journal of Institutional Economics*, **12**(1): 43–48.
- Tilly, C. (1992), *Coercion, Capital, and European States, AD 990–1992*, Oxford, UK: Blackwell.
- Weber, E. (1976), *Peasants Into Frenchmen: The Modernization of Rural France, 1870–1914*, Stanford, CA: Stanford University Press.
- Weil, D. N. (2014), 'Health and Economic Growth', in *Handbook of Economic Growth* (Vol. 2 vols, Oxford, UK: Elsevier, pp. 623–682).
- Wilson, W. J. (2011) *When Work Disappears: The World of the New Urban Poor*, New York, NY: Vintage.
- Witt, U. (2008), 'Observational Learning, Group Selection, and Societal Evolution', *Journal of Institutional Economics*, **4**(1): 1–24.
- Young, H. P. (1998), *Individual Strategy and Social Structure*, Princeton, NJ: Princeton University Press.
- Young, H. P. (2011), 'The Dynamics of Social Innovation', *Proceedings of the National Academy of Sciences*, **108** (Supplement 4): 21285–21291.
- Young, M. (2017), *The Rise of the Meritocracy*, London, UK: Routledge.

Appendix

Proof of Proposition 1. By observation of equations (6) and (9), β_H and β_L are continuous in (s_L, s_H) . In addition, the systems (10) and (11) define a mapping from the compact subspace $S = \{(s_L, s_H) : 0 \leq s_H, s_L \leq b/\beta\}$ to itself. Hence, by Brouwer's fixed point theorem, there exists at least one equilibrium. \square

Proof of Proposition 2. The proposition follows from the inspection of equations (10)–(12), and in particular comparing the limits of integration. \square

Proof of Proposition 3. Note that

$$\begin{aligned} \bar{s}(1) &= [m_H F(\beta_H) + m_L F(\beta_L)]s_H + [m_H(1 - F(\beta_H)) + m_L(1 - F(\beta_L))]s_L \\ &= b \left[m_H \int_{\beta_H}^1 \frac{1}{\beta} dF + m_L \int_{\beta_L}^1 \frac{1}{\beta} dF + m_H \int_{\underline{\beta}}^{\beta_H} \frac{1}{\beta} dF + m_L \int_{\underline{\beta}}^{\beta_L} \frac{1}{\beta} dF \right] \\ &= b \left[m_H \int_{\underline{\beta}}^1 \frac{1}{\beta} dF + m_L \int_{\underline{\beta}}^1 \frac{1}{\beta} dF \right] \\ &= b \int_{\underline{\beta}}^1 \frac{1}{\beta} dF = S. \end{aligned}$$

Hence $\bar{s}(1)^\sigma = S^\sigma$. Moreover, $\bar{s}(\sigma) > \bar{s}(1)^\sigma$ if and only if $\sigma > 1$ by Jensen’s inequality. This establishes the proposition. \square

Proof of Proposition 4. In period $t=0$, play takes place as in the one-shot game with investment cutoffs β_H^0 and β_L^0 given by equations (6) and (9). In period 1 then, each class is sorted and the distributions of β have the following properties: $F_L^1(\beta_L^0) = 1$ and $F_H^1(\beta_H^0) = 0$. Consider a putative equilibrium in which $\beta_H^1 = \beta_H^0$ and $\beta_L^1 = \beta_L^0$. Then $s_H^1 = s_H^0$ and $s_L^1 = s_L^0$, so the cutoffs are the same as in period 0, as supposed. Hence this is an equilibrium. In this equilibrium, $U^1 = 1 - F_L^1(\beta_L^0) = 0$ and $D^1 = F_H^1(\beta_H^0) = 0$. Iterating this argument yields the proposition. \square