Contributions

Rebecca M. Glawtschew*

Eliminating Persistent Statistical Discrimination: An Analysis of Several Policy Options

Abstract: I consider a model of endogenous human capital formation with competitively determined wages in a dynamic setting. In the presence of two distinguishable, but ex ante identical groups of workers, discrimination will be persistent in equilibrium. Using this framework, I then consider the effectiveness of three government policies designed to eliminate this discrimination. I determine the paths that workers will take after a policy is instated as well as how long a policy needs to be in place to guarantee the successful elimination of discrimination. The policies I consider are (1) a hiring subsidy that promotes the hiring of disadvantaged workers to the better job, (2) an investment voucher that defrays the monetary cost of human capital investment, and (3) an equal treatment policy under which firms are required to treat workers equally across groups. I find that all three policies have the potential to eliminate persistent discrimination if certain conditions are met. In a general equilibrium setting, I also address the welfare effects of the three policies for a parametric example.

Keywords: statistical discrimination, general equilibrium, asymmetric information, human capital

JEL Classification: D58, D62, D82, J71, J78

DOI 10.1515/bejeap-2014-0005

1 Introduction

This paper explores the consequences of three government policies designed to eliminate existing persistent statistical discrimination. This analysis is done in

*Corresponding author: Rebecca M. Glawtschew, Department of Economics, Finance and Quantitative Analysis, Kennesaw State University, Burruss Building 341, Kennesaw, GA 30144, USA, E-mail: rglawtsc@kennesaw.edu

the framework of a dynamic model of endogenous human capital formation with competitively determined wages and two distinguishable, but ex ante identical groups of workers. By examining these policies in a dynamic framework, I am able to characterize the gradual change in each group's human capital investment decisions that occurs as a result of each policy and to determine when and if the government can safely remove each policy. I am also able to meaningfully analyze the welfare implications of the three policies, because wage is endogenously determined through competition between firms.

The dynamic model of statistical discrimination on which I rely to assess the implications of various policies has three wage steady states in equilibrium. The middle steady state is unstable and paths to the high and low steady states emanate from that point. These paths can take several forms depending on the parameter values. No matter what form the dynamic paths take, a group becomes "stuck" in a steady state as soon as it is reached. As such, if the groups end up at different steady states discrimination is present and persistent.

For the purpose of this analysis, I assume that the economy is experiencing *discrimination*, so that one group of workers, which I refer to as *disadvantaged*, is stuck in the low steady state and the other group of workers, which I refer to as *advantaged*, is stuck in the high steady state. I examine the impact of three government policies from this starting point. The goal of these policies is to move the disadvantaged group from the low steady state to the high steady state.

The first policy I examine is one in which the government offers a *hiring subsidy* to firms for correctly assigning skilled workers in the disadvantaged group to the more productive job. This policy is always effective at eliminating discrimination as long as a large enough subsidy is offered.

The second policy of interest is an *investment voucher* offered to workers in the disadvantaged group. This voucher defrays the monetary cost associated with skill acquisition. As long as the monetary costs are high, this policy can eliminate discrimination.

The last policy is one in which the government enforces *equal treatment* of workers across groups. This policy may be effective depending on what proportion of the total worker population is disadvantaged. The policy will successfully eliminate discrimination if a small proportion of the population is disadvantaged. If the groups are close to equal in size, then the policy may eliminate discrimination and the final outcome depends on the specific parameterization. If the disadvantaged group is relatively large, then this policy will move the advantaged group from the high to the low steady state, resulting in a population that is worse off on average.

Using a simplified version of the dynamic model, I examine the welfare effects of the three policies. I complete this analysis in a general equilibrium

framework. Workers in the advantaged group are taxed in order to fund the policies. I find that all three policies create a net increase in welfare while they are in place. However, there are significant differences across policies with respect to how much the advantaged group is impacted over the course of the policy. The hiring subsidy results in a very large loss in welfare to the advantaged group, but it moves the disadvantaged group to the high steady state the fastest. The investment voucher reduces the welfare of the advantaged group the least of all the policies and for this reason it may be viewed as optimal by some policy makers. The equal treatment policy has a moderate impact of the advantaged group's welfare and may be viewed as optimal by a policy maker who does not want to redistribute.

The paper is structured as follows. The related literature is discussed in the remainder of Section 1. The details of the general dynamic model are described in Section 2. Section 3 develops the equilibria of the dynamic model and addresses the conditions under which discrimination will be persistent. Section 4 describes the three policies and examines their impact on the general dynamic model. In Section 5 a simple version of the dynamic model is developed; using that model I consider the welfare effects of the three policies. Finally, Section 6 concludes.

1.1 Related literature

The model presented in this paper is closely related to Coate and Loury (1993) and Moro and Norman (2003). Coate and Loury (1993) present a model of statistical discrimination where two *ex ante* groups may end up in different, Pareto ranked, equilibria. This model formalizes many of the ideas that were originally presented loosely in Arrow (1973), but it makes the additional assumption that wages are set exogenously. They find that inequality will occur if the two groups of workers achieve different equilibria. In the context of their model, Coate and Loury discuss whether affirmative action or similar policies can be temporary policies. However, owing to the static nature of the model they are unable to make a formal determination. The model presented here allows for a formal discussion of whether or not these types of policies can be removed after sufficient time has passed.

Moro and Norman (2003) relax two of the assumptions in Coate and Loury's model. In particular, they remove the linearity of the production technology and the exogeneity of wages. Allowing for curvature in the production function eliminates the separability of groups, thus allowing for true interaction effects between groups. Introducing endogenous wages allows for more meaningful

welfare analysis. They find that under these relaxed assumptions the dominant group will actually gain from discrimination and, as a result, they will be motivated to ensure its continuation. The model presented here adopts the endogeneity of wages, which allows me to perform meaningful welfare analysis of the policies.

There is relatively little literature that addresses discrimination in a dynamic environment. Fryer (2007), Blume (2006), and Kim and Loury (2014) all develop dynamic models that are adaptations of the original Coate and Loury (1993) paper. These models all differ from the one presented here in two key areas: they do not incorporate endogenous wages nor do they consider the welfare effects of potential policies designed to eliminate or mitigate discrimination.

The paper that is most related to the analysis presented here is Kim and Loury (2014). They also add long-lived workers to a static model in order to incorporate dynamics. However, unlike the model presented here, they do not include endogenous wages. Moreover, the goals of the two papers differ significantly. They develop a dynamic model of statistical discrimination in order to analyze how group reputation determines the final outcome of a group. They also complete a brief analysis of two policies that may eliminate discrimination. One of these policies, a training subsidy, is similar to the investment voucher policy that I examine. My goal is to develop a dynamic model in order to undertake a thorough analysis of the effectiveness of three policies. Specifically, I determine in the potential paths that each group will follow once a policy is put in place, how long a policy needs to be in place to be effective and the potential welfare impacts of each policy while it is in place. All of these results go well beyond what is developed in Kim and Loury (2014).

2 The model

Consider a market where short-lived firms engage in Bertrand-style competition for workers in each period. Time is continuous and all changes occur at the beginning of each period of length Δ .

There are two identifiable groups of workers, B(lack) and W(hite), and the total population of each group is constant at λ^J for $J = B, W$. Workers are long-lived and subject to a *Poisson death process* with parameter β . Consequently, at the beginning of each period $\Delta\beta$ of the existing workers are replaced by newborn individuals. Workers also discount future pay-offs at a strictly positive rate r . Prior to entering the labor market, each worker has the option to invest in human capital. Workers who make this investment become

qualified and those that do not remain *unqualified*. These premarket investment decisions are final; a worker cannot choose to invest in human capital later in life.

A worker who invests in human capital incurs a total cost c . This cost of investment consists of two important parts: a monetary cost, c_m , which is strictly positive and fixed across all workers and a nonmonetary cost, c_e , which varies across workers. The nonmonetary cost represents both the effort required to become qualified and any value that a worker places on being qualified (outside of the potential increase in expected wage); it may be positive or negative.

Total cost of investment will vary across workers because of the differences in c_e . I assume that the total cost of investment, c , is distributed over $[\underline{c}, \bar{c}] \subseteq \mathbb{R}$ according to a continuous and strictly increasing distribution $G(c)$. Workers are risk-neutral and care only about the wage that they receive and the cost they incur. Pay-offs are additively separable in income and cost of investment.

Output in each period is generated by the completion of two types of tasks, a simple task and a complex task. All workers can perform the simple task, but only those workers who are qualified can successfully perform the complex task. For a given period starting at t , the effective input of labor into the simple task, S_t , is the total number of workers employed in that task. The effective input of labor into the complex task, C_t , is the number of qualified workers employed in that task. Firms' per period output is given by the linear production function $Y_\Delta(S_t, C_t) = \Delta[\alpha_s S_t + \alpha_c C_t]$, where $\alpha_c > \alpha_s > 0$. This condition implies that firms will always want to assign a skilled worker to the complex task.

Firms cannot observe whether a worker is qualified, nor are they able to distinguish between new and old workers. The inability to distinguish between old and new workers is a limitation of the model that is required to generate manageable dynamics. The only information a firm has about a worker in a given period, aside from his group identity, is his current signal $\theta_t \in [0, 1]$, which is distributed according to f_u if a worker is unqualified or f_q if a worker is qualified. Both densities are bounded away from zero and, without further loss of generality, $f_q(\theta)/f_u(\theta)$ is strictly increasing in θ . I denote the associated cumulative distributions by F_q and F_u . I assume a law of large numbers holds so that these are also the realized frequency distributions of signals for qualified and unqualified workers, respectively.

A firm's prior belief about what proportion of the total J population is qualified at time t is represented by Π_t^J . The posterior probability that a worker from group J with signal θ_t is qualified given beliefs Π_t^J is

$$p(\theta_t, \Pi_t^J) \equiv \frac{\Pi_t^J f_q(\theta)}{\Pi_t^J f_q(\theta) + (1 - \Pi_t^J) f_u(\theta)}. \quad [1]$$

The strict monotone likelihood ratio property implies that $p(\theta_t, \Pi_t^J)$ is strictly increasing in θ_t so that a high signal reflects positively on a worker and a low signal negatively.

The timing of each period of the dynamic game is as follows: In the first stage, $\Delta\beta$ new workers enter the workforce and decide whether or not to invest in becoming qualified. At the same time, firms post wage and task assignment rules. Firms may condition wage and task assignments on θ_t . Formally, a strategy for each firm is to select some wage schedule $\omega_t^J : [0, 1] \rightarrow \mathbb{R}_+$ and a task assignment rule $\tau_t^J : [0, 1] \rightarrow [0, 1]$ for each group J . In the second stage, nature assigns a signal θ_t to each worker according to the appropriate density. Workers observe the posted wage and task assignment rules and decide where to work. Then tasks are performed, wages are paid, and output is realized. In the last stage, $\Delta\beta$ of the workers exit the workforce.

Workers only care about potential wages when comparing job offers and their investment costs are sunk. Hence they will choose to work for the firm whose wage and task assignment rule provides them with the highest wage.

3 Equilibria

There are two main components to an equilibrium of this game. First, firms and workers must behave optimally in each period. This means that firms must employ both an optimal task assignment rule and an equilibrium wage rule in each period. While members of each new cohort of workers invest only if their expected lifetime benefit at the time of birth exceeds their cost, all workers choose firms in a sequentially rational way after any history of play. Second, firms have rational beliefs about proportion of qualified workers in each period. Formally, an equilibrium of this model is described by a sequence of rational firm beliefs regarding the proportion of qualified workers in each group, $\{\Pi_t^J\}_{t=0}^\infty$ for $J = B, W$.

Let (Π_t^B, Π_t^W) denote the group-specific investment rates in the period starting at time t . Given this arbitrary investment behavior by workers, I can address the optimal behavior of firms in each period. The probability that a worker is qualified, as defined in eq. [1], is strictly increasing in signal. This implies that threshold rules are optimal when assigning workers to tasks. Under this type of rule, workers with a signal above the threshold are assigned to the complex task and workers with lower signals are assigned to the simple task. Given a threshold $\bar{\theta}_t^J$ for group J at time t the effective input of labor into the complex task is $C_t^J = \Pi_t^J(1 - F_q(\bar{\theta}_t^J))$. This is simply the number of qualified workers who are

assigned to the complex task. Any unqualified workers with signals above the threshold will fail to complete the complex task. Consequently, these workers do not appear in the effective input of labor into the complex task. The effective input of labor into the simple task is $S_t^I = \Pi_t^I F_q(\bar{\theta}_t^I) + (1 - \Pi_t^I) F_u(\bar{\theta}_t^I)$. More specifically, this is all the workers, both qualified and unqualified, with a signal below the threshold, $\bar{\theta}_t^I$.

The optimal threshold is the one that maximizes a firm's expected output conditional on investment decisions. Formally, the task assignment problem is

$$\sum_{j=B,W} \lambda^j \max_{\theta_t^I \in [0,1]} (\Delta \alpha_c [\Pi_t^I (1 - F_q(\theta_t^I))] + \Delta \alpha_s [\Pi_t^I F_q(\theta_t^I) + (1 - \Pi_t^I) F_u(\theta_t^I)]). \quad [2]$$

This task assignment problem will always have a unique solution. See Lemma 1 of Glawtschew (2013) for a formal proof. Moreover, the linear production function allows for the task assignment problem to be solved separately for each group. Consequently, there is a unique optimal threshold for each group. It is

$$\hat{\theta}(\Pi_t^I) = \begin{cases} 1 & \text{if } \alpha_c p(1, \Pi_t^I) \leq \alpha_s, \\ 0 & \text{if } \alpha_c p(0, \Pi_t^I) \geq \alpha_s, \\ \text{the unique solution to} & \\ \alpha_c p(\theta, \Pi_t^I) = \alpha_s & \text{if } \alpha_c p(0, \Pi_t^I) < \alpha_s < \alpha_c p(1, \Pi_t^I). \end{cases} \quad [3]$$

Under this task assignment rule, all workers are assigned to the simple task (a threshold of 1) if prior beliefs are such that a worker with a signal of 1 has higher expected productivity in the simple task. All workers are assigned to the complex task (a threshold of 0) if prior beliefs are such that a worker with a signal of 0 has higher expected productivity in the complex task. Otherwise, the optimal threshold is the θ at which a worker's expected productivity in the complex task is equivalent to that of the simple task.

Following Moro and Norman (2003), I call a strategy profile a *continuation equilibrium* if firms are implementing their optimal wage and task assignment rules given some arbitrary investment behavior by workers. This first result suggests that wages are given by expected marginal products and job assignments are constrained efficient in any continuation equilibrium.

Proposition 1. *Suppose for a given period beginning at t that fractions $\Pi_t = (\Pi_t^B, \Pi_t^W)$ of the workers invest and that $(\hat{\theta}_t(\Pi_t^B), \hat{\theta}_t(\Pi_t^W))$ is as in eq. [3]. Then there exists a continuation equilibrium where both firms post wage schedules*

$$\omega_\Delta(\theta_t, \Pi_t^I) = \begin{cases} \Delta \alpha_s & \text{if } \theta_t < \hat{\theta}(\Pi_t^I) \\ \Delta \alpha_c p(\theta_t, \Pi_t^I) & \text{if } \theta_t \geq \hat{\theta}(\Pi_t^I) \end{cases}, \quad [4]$$

and assign a worker with characteristics (J, θ_t) to the complex task if and only if $\theta_t \geq \hat{\theta}_t(\Pi_t)$. Moreover, in any continuation equilibrium where fractions $\Pi_t = (\Pi_t^B, \Pi_t^W)$ of the workers invest the posted wage schedule for group J , $\omega_\Delta(\theta_t, \Pi_t^J)$ agrees with eq. [4] for almost all $\theta_t \in [0, 1]$ for each firm.

Proof. See Appendix A. □

In a full equilibrium, investments must be best responses to wages. The optimal investment decision for a worker is to invest in human capital if and only if the lifetime gain in expected earnings is higher than the cost c . It is helpful to define what I refer to as the *unit period gain from investment*. This is a worker's one-period expected gain from investment if periods have a length of one. For wages consistent with the continuation equilibrium, where a fraction Π_t^J of the J population is qualified, the unit period gain from investment is written as

$$\begin{aligned} \gamma(\Pi_t^J) &= \frac{1}{\Delta} \left(\int_0^1 \omega_\Delta(\theta, \Pi^J) f_q(\theta) d\theta - \int_0^1 \omega_\Delta(\theta, \Pi^J) f_u(\theta) d\theta \right) \\ &= \alpha_s [F_q(\hat{\theta}(\Pi_t^J)) - F_u(\hat{\theta}(\Pi_t^J))] + \alpha_c \int_{\hat{\theta}(\Pi_t^J)}^1 p(\theta, \Pi_t^J) [f_q(\theta) - f_u(\theta)] d\theta. \end{aligned} \quad [5]$$

I can also define the *per period gain from investment* as $\Delta\gamma(\Pi_t^J)$.

Using the per period gain from investment, I can define the *lifetime gain from investment* from the perspective of a worker at time t in group J . It is simply the discounted sum of all future per period gains from investment. It can be written as

$$\Gamma_t^J = \sum_{g=0}^{\infty} (1 - \Delta\beta - \Delta r)^g \Delta\gamma(\Pi_{t+g\Delta}^J) \quad [6]$$

$$= (1 - \Delta\beta - \Delta r) \Gamma_{t+\Delta}^J + \Delta\gamma(\Pi_t^J). \quad [7]$$

Rearranging eq. [6] yields

$$\frac{\Gamma_{t+\Delta}^J - \Gamma_t^J}{\Delta} = (\beta + r) \Gamma_{t+\Delta}^J - \gamma(\Pi_t^J). \quad [8]$$

If I let $\Delta \rightarrow 0$, then I can express how the lifetime gain from investment evolves over time with the following differential equation:

$$\dot{\Gamma}_t^J = (\beta + r) \Gamma_t^J - \gamma(\Pi_t^J). \quad [9]$$

I can also address how the investment rate of each group's population will develop over time. In each period, workers in the new cohort invest if their lifetime gain from investment exceeds their cost. The proportion of new J population births that invest in the period starting at time t is $G(\Gamma_t^J)$, where $G(\cdot)$ is the distribution of costs.

Given the Poisson death process and the assumption that the total population is constant, the investment rate of the whole J population at time t is

$$\Pi_t^J = (1 - \Delta\beta)\Pi_{t-\Delta}^J + \Delta\beta G(\Gamma_t^J). \quad [10]$$

Rearranging eq. [10] yields

$$\frac{\Pi_t^J - \Pi_{t-\Delta}^J}{\Delta} = \beta(G(\Gamma_t^J) - \Pi_{t-\Delta}^J). \quad [11]$$

If I let $\Delta \rightarrow 0$, then I can express how the overall investment rate of group J evolves over time with the following differential equation:

$$\dot{\Pi}_t^J = \beta(G(\Gamma_t^J) - \Pi_t^J). \quad [12]$$

I conclude that any equilibria of this model is characterized by a dynamic system consisting of eqs [9] and [12].

Proposition 2. *An equilibrium of this dynamic model is fully characterized by the following two-variable differential equations:*

$$\begin{aligned} \dot{\Gamma}_t^J &= (\beta + r)\Gamma_t^J - \gamma(\Pi_t^J), \\ \dot{\Pi}_t^J &= \beta(G(\Gamma_t^J) - \Pi_t^J). \end{aligned} \quad [13]$$

This proposition indicates that the difference between the investment rate of the newborn cohort and the overall investment rate of the group J determines the speed of the firms' beliefs about the fraction of workers in group J that are qualified. The change in lifetime gain from investment is determined by the difference between the discounted lifetime gain from investment and the unit period gain from investment.

It is important to note that Kim and Loury (2014) develop a very similar dynamic system in their Theorem 1. However, unlike their result, this system incorporates competitively determined wages via the lifetime incentive to invest function, Γ_t^J . The steady states of the dynamic system can be described as all beliefs satisfying

$$\Pi^J = G\left(\frac{\gamma(\Pi^J)}{\beta + r}\right). \quad [14]$$

The existence of a solution to eq. [14] is immediate since $\gamma(\Pi_t^J)$ is composed of continuous functions. There may be a unique solution, in which case groups are treated identically when they are both in the steady state. For the purposes of this analysis, multiple steady states are desirable since persistent discrimination is only possible if this is the case.

Figure 1 shows an example of the dynamic system with three steady states. From now on I denote these three steady states as $\Sigma_l(\Gamma_l, \Pi_l)$, $\Sigma_m(\Gamma_m, \Pi_m)$, and $\Sigma_h(\Gamma_h, \Pi_h)$, referring to the low, middle, and high steady states, respectively. The hump-shaped curve is the $\dot{\Gamma}_t^J = 0$ locus. The arrows indicate that to the right of the locus, when $\Gamma_t^J > \gamma(\Pi_t^J)/(\beta + r)$, Γ^J is increasing. To the left of the locus, Γ_t^J is decreasing. The s-shaped curve is the $\dot{\Pi}_t^J = 0$ locus. The arrows indicate that above this locus, when $G(\Gamma_t^J) < \Pi_t^J$, Π_t^J is increasing and below the locus it is decreasing. The shape is from the assumption that $G(\cdot)$ is continuous and strictly increasing.

Figure 1: Phase diagram of the dynamic system

The steady states in this model are a result of self-fulfilling human capital investment behavior. In particular, a group in the low steady state will have relatively low average investment rates and consequently they will be offered a less attractive wage schedule. These low wages result in low incentives to invest and as a result new generations will also have low average investment rates. In essence, a group in the low steady state is stuck in a cycle of self-perpetuating

underinvestment in human capital. A group in the high steady state will have relatively high average investment rates, and as a result will be offered a more favorable wage schedule. These high wages result in large incentives to invest and as a result new generations will also have high average investment rates. So a group in the high steady state is stuck in a self-perpetuating cycle of high investment in human capital. A group in the middle steady state has a similar story, but with intermediate investment behavior and wage schedules.

The existence of multiple solutions is not guaranteed and depends on the shape of Γ and $G(\cdot)$. The possibility of multiple solutions can be proven by construction: if I fix the parameters f_u , f_q , α_c and α_s , then I can find an appropriate $G(\cdot)$ function such that eq. [14] has multiple solutions. Graphically, if I fix the $\dot{\Gamma}_t^J = 0$ locus in Figure 1, I can ensure multiple equilibria simply by selecting a $G(\cdot)$ function that yields multiple intersections of the two loci. I can conclude that for any f_u and f_q that satisfy the monotone likelihood ratio and any technology parameters, I can find a nontrivial set of $G(\cdot)$ functions that ensure multiple steady states. These multiple steady states are Pareto rankable. In particular, it is simple to show that the steady state with the highest Π , $\Sigma_h(\Gamma_h, \Pi_h)$, is Pareto dominant. See Proposition 4 of Glawtschew (2013) for the formal proof.

It is necessary to establish which of these three steady states are stable and to determine the potential shapes of the paths leading to steady states. Understanding the behavior of the system outside of the steady states is important both for determining if discrimination is persistent and for analyzing the effectiveness of the policies in Section 4. The phase diagram indicates that both the high and the low steady states are saddle points, while the middle steady state is an unstable source. See Lemma 2 of Glawtschew (2013) for a formal proof.¹

The existence of two saddle points means that members of a group may rationally conjecture that firms will eventually have positive beliefs about their group's investment rate, Π_h , or have negative beliefs about their group's investment rate, Π_l . The only reasonable steady states for a group to approach are Σ_h and Σ_l . It is possible to characterize the shape that the paths to these saddle stable steady states will take.

The analysis of path shape relies heavily on the conclusions drawn in Krugman (1991). In that paper Krugman develops a simple model of trade in which there are three steady states with different levels of specialization in two

¹ Kim and Loury (2014) draw the same conclusions about stability in their Lemma 5. This result is not surprising considering the similarity of the relevant dynamic systems.

goods. The two extreme steady states are saddle stable and the third, intermediate steady state is unstable. He uses this model to show that depending on the structure of the economy, the steady state an economy will converge to may depend entirely on history, or it may be determined by expectations, which can be thought of as self-fulfilling prophecy. In particular, he shows that it is the shape of the paths to the steady states that determines whether history or expectations will direct an economy's final resting place. If the paths are S-shaped, then history alone is the determining factor. If the paths form interlocking spirals, then there is a range of initial values such that an economy can converge to either steady state depending on the economy's expectations. He refers to this range of initial conditions as the "overlap," as it is where the paths to the two saddle stable steady states overlap one another.

Clearly the structure of the dynamic model of statistical discrimination presented here mirrors that of the model addressed in Krugman (1991). Specifically, I have a model with three steady states, in which the high and low steady states are saddle stable and the middle steady state is unstable. Consequently, I can use the same analysis to determine the parameter values that generate S-shaped paths, which I refer to as *deterministic paths*, and that generate paths with interlocking spirals, which I refer to as *spiralling paths*. When paths are *deterministic*, historical firm beliefs completely determine the final steady state of a group. On the other hand, when paths are *spiralling*, there is a region of initial firm beliefs that can lead to either of the saddle stable steady states depending on the expectations of the group. See Lemma 3 of Glawtschew (2013) for a formal discussion of the conditions that lead to each type of path.

Figure 2 depicts a case where there are three steady states and paths are spiralling. As determined in Krugman (1991), spiralling paths to the steady states create a range of initial firm beliefs that can lead to either the high or the low steady state, which is indicated by the purple region. I call this range of initial beliefs the *uncertain region* rather than the "overlap," as in Krugman (1991), because it generates uncertainty about the final outcome of a group. This region is formed by the bottom bound of the path to the high steady state, $\hat{\Pi}$, and by the upper bound of the path to the low steady state, $\hat{\Gamma}$. If initial firm beliefs about the investment rate of a group are in this range, then that group can converge to either $\Sigma_h(\Pi_h, \Gamma_h)$ or $\Sigma_l(\Pi_l, \Gamma_l)$.

When initial firm beliefs are in the uncertain region, the steady state that a group converges to is determined by the expectations of the workers in that group. Again, expectations can be thought of as a self-fulfilling prophecy. So, if a group believes they will have high investment rates in the future and all the wage benefits that result from that, then they make their current investment

Figure 2: Spiralling paths to the steady states

decisions according to that belief. As a result, the average investment rate increases each period until the group reaches the high steady state. In other words, if a group develops and shares an optimistic view of the future, then they will converge to the high steady state. Conversely, if they develop a pessimistic view, they will converge to the low steady state. Consequently, I refer to the path that leads to the high steady state as the *optimistic path* and the path that leads to the low steady state as the *pessimistic path*. See Appendix C.1 for a discussion of how these paths are calculated.

If initial firm beliefs for a group fall outside the uncertain region, then the steady state the group converges to will depend entirely on history. In the case of spiralling paths, when initial beliefs about investment rate are such that $\Pi_0^J \in (\hat{\Pi}, 1]$, then group J will converge to the high qualification steady state with certainty. This is because the optimistic path is the only reasonable path to take. When initial firms' beliefs are such that $\Pi_0^J \in [0, \check{\Pi})$, then group J will converge to the low qualification steady state with certainty. Additionally, the high and low steady states will always lie outside of the uncertain region, which means that there is no way for a group to move from one steady state to another.

Figure 3 depicts a case where there are three steady states and paths are deterministic. In this situation there is no uncertain region, as a result initial firm beliefs are the only factor that determine the final steady state of a group. If initial beliefs about investment rate are such that $\Pi_0^J \in (\Pi_m, 1]$, then group J will converge to the high qualification steady state with certainty. If $\Pi_0^J \in [0, \Pi_m)$,

Figure 3: Deterministic paths to the steady states

then the group will converge to the low qualification steady state with certainty. Again, when paths are deterministic, groups are unable to move from one steady state to another.

From now on, regardless of the shape of the paths, I refer to the range of initial firm beliefs that lead only to the high steady state as the *certain region for Σ_h* , and I refer to the range of initial firm beliefs that lead only to the low steady state as the *certain region for Σ_l* . These certain regions make it possible for a group to get stuck permanently at either the high or the low steady state given appropriate initial firm beliefs. This suggests that once present, discrimination will persist indefinitely without a structural change to the model.² If a policy moves firm beliefs into the certain region for Σ_h , then it can be removed and the group will converge to the high steady state.

4 Potential policies

The analysis of Section 3 suggests that, if present, statistical discrimination is persistent. One way to eliminate this discrimination is through appropriate government intervention. In this section, I consider three government policies that have the potential to eliminate existing statistical discrimination.

² See Glawtschew (2013) for a full discussion of the persistence of statistical discrimination.

As a starting point for the analysis of the policies I assume that one group is in the high steady state and the other is in the low steady state, so that persistent discrimination is present. From now on I refer to the group in the low steady state as the *disadvantaged group* and the group in the high steady state as the *advantaged group*. I then determine under what circumstances the three policies are able to move the disadvantaged group from the low to the high steady state. Discrimination in this model will be eliminated if both groups are in either the high or the low steady state. Since the high steady state Pareto dominates the low steady state, I will only consider a policy *successful* if it is able to move both groups to the high steady state.

Specifically, in the analysis that follows, I determine whether each policy allows the disadvantaged group to get on the path to the high steady state. I also address if the policy needs to be in effect permanently or if there is a point at which it can be removed and the discrimination will not return. I assume that the government commits to employing each policy at least until it can be safely removed and still successfully eliminate the discrimination.

The policies that I investigate are (1) a subsidy paid by the government to the firms for every worker in the disadvantaged group that is correctly assigned to the complex task, (2) a government subsidization of investment costs for those workers in the disadvantaged group, and (3) a government policy that forces firms to assign workers in different groups with the same signal to the same task and pay them the same wage.

4.1 Hiring subsidy program

The first policy is a hiring subsidy that promotes the assignment of disadvantaged workers to the complex task. This subsidization is applied in the following way: For each worker in the disadvantaged group who is correctly assigned to the complex task, the government compensates firms with a subsidy of $\Delta\xi$.

The government cannot actually observe whether workers are correctly assigned, instead this policy is implemented through a lump sum payment of $\Delta\xi\Pi_t^d[1 - F_q(\hat{\theta}(\Pi_t^d))]$ to the firms. This is simply the value of the subsidy times the expected number of workers from the disadvantaged group that will be correctly assigned to the complex task. This type of subsidy requires that the government only be able to observe what the firms observe. Under the subsidy, the firms' profit from the disadvantaged group, given an arbitrary threshold θ_t^d , is

$$\lambda^d((\Delta\alpha_c + \Delta\xi)[\Pi_t^d(1 - F_q(\theta_t^d))] + \Delta\alpha_s[\Pi_t^d F_q(\theta_t^d) + (1 - \Pi_t^d)F_u(\theta_t^d)]). \quad [15]$$

Since there are two firms engaged in Bertrand competition for workers and the firms have linear production functions, all of the benefits of the subsidy program are passed on to the workers via an increase in continuation wage to workers assigned to the complex task. So the subsidy functions as a wage increase for those disadvantaged workers who are assigned to the complex task. This will clearly result in an increase in the gain from investment function for workers in the disadvantaged group. Workers in the advantaged group are not affected by this policy; they will remain at Σ_h . For simplicity of notation, I drop the group subscripts in this section.

The profit function in eq. [15] indicates that I can represent the hiring subsidy with a change to the productivity of qualified workers in the disadvantaged group. Given a subsidy of $\Delta\xi$, the expected productivity of a disadvantaged worker with signal θ_t who is assigned to the complex task is $\Delta(\xi + \alpha_c)p(\theta_t, \Pi_t)$. The higher productivity of qualified workers in the disadvantaged group means that, given the same firm beliefs, the optimal threshold is reduced by the policy. The optimal threshold for the disadvantage group given a subsidy of $\Delta\xi$ is

$$\hat{\theta}^\xi(\Pi_t) = \begin{cases} 1 & \text{if } (\alpha_c + \xi)p(1, \Pi_t) \leq \alpha_s, \\ 0 & \text{if } (\alpha_c + \xi)p(0, \Pi_t) \geq \alpha_s, \\ \text{the unique solution to} & \\ (\alpha_c + \xi)p(\theta, \Pi_t) = \alpha_s & \text{if } (\alpha_c + \xi)p(0, \Pi_t) < \alpha_s < (\alpha_c + \xi)p(1, \Pi_t). \end{cases} \quad [16]$$

Since workers are paid their expected productivity, the wages of those workers assigned to the complex task will be higher as a result of the policy. The continuation wage under the hiring subsidy is

$$\omega_\Delta^\xi(\theta; \Pi_t) = \begin{cases} \Delta\alpha_s & \text{if } \theta < \hat{\theta}^\xi(\Pi_t), \\ \Delta(\alpha_c + \xi)p(\theta, \Pi_t) & \text{if } \theta \geq \hat{\theta}^\xi(\Pi_t). \end{cases} \quad [17]$$

The gain from investment will also be impacted. The new unit period gain from investment is

$$\gamma^\xi(\Pi_t) = \alpha_s \left[F_q(\hat{\theta}^\xi(\Pi_t)) - F_u(\hat{\theta}^\xi(\Pi_t)) \right] + (\alpha_c + \xi) \left[\int_{\hat{\theta}^\xi(\Pi_t)}^1 p(\theta, \Pi_t) [f_q(\theta) - f_u(\theta)] d\theta \right]. \quad [18]$$

The change from $\gamma(\Pi_t)$ to $\gamma^\xi(\Pi_t)$ also changes the lifetime gain from investment Γ_t .

Consequently, under the hiring subsidy, the equilibrium for the disadvantaged group is described by a new dynamic system.

Proposition 3. *An equilibrium of the dynamic model under a hiring subsidy is fully characterized by the following two variable differential equations:*

$$\begin{aligned}\dot{\Gamma}_t &= (\beta + r)\Gamma_t - \gamma^\xi(\Pi_t), \\ \dot{\Pi}_t &= \beta(G(\Gamma_t) - \Pi_t).\end{aligned}\quad [19]$$

It is clear that the hiring subsidy policy only impacts the $\dot{\Gamma}_t = 0$ locus. The $\dot{\Pi}_t = 0$ locus is unchanged.

In order for this policy to eliminate discrimination it must move the disadvantaged group into the certain region of Σ_h from the base model. This result is guaranteed if the dynamic system under the hiring subsidy has only one steady state that is within the relevant region.

Given that the baseline dynamic system has three steady states, the only way that the dynamic system under the hiring subsidy will have one steady state is if $\gamma^\xi(\Pi_t)$ is either much larger or much smaller than $\gamma(\Pi_t)$ for every value of Π_t . If $\gamma^\xi(\Pi_t)$ is smaller, then there is only one steady state at a low investment level. If $\gamma^\xi(\Pi_t)$ is larger, then the single steady state occurs at a high investment level.

The following lemma indicates that $\gamma^\xi(\Pi_t)$ is nondecreasing in ξ (the unit period subsidy). This means that under the hiring subsidy policy, $\gamma^\xi(\Pi_t)$ is larger than $\gamma(\Pi_t)$ when ξ is positive.

Lemma 1. *If Π_t is such that $\hat{\theta}^\xi(\Pi_t) \in [0, 1)$, then the unit period gain from investment is strictly increasing in ξ . If Π_t is such that $\hat{\theta}^\xi(\Pi_t) = 1$, then the unit period gain from investment is equal to zero and unaffected by an increase in ξ .*

Proof. The partial derivative of the unit period gain from investment with respect to the hiring subsidy ξ is

$$\begin{aligned}\frac{\partial \gamma^\xi(\Pi_t)}{\partial \xi} &= \underbrace{\frac{\partial \hat{\theta}^\xi(\Pi_t)}{\partial \xi} \left[\alpha_s - (\xi + \alpha_c)p(\hat{\theta}^\xi(\Pi_t), \Pi_t) \right]}_{\leq 0} \left(f_q(\hat{\theta}^\xi(\Pi_t)) - f_u(\hat{\theta}^\xi(\Pi_t)) \right) \\ &\quad + \underbrace{\int_{\hat{\theta}^\xi(\Pi_t)}^1 p(\theta, \Pi_t) [f_q(\theta) - f_u(\theta)] d\theta}_{> 0 \text{ unless } \Pi_t = 0 \text{ or } \Pi_t = 1}.\end{aligned}\quad [20]$$

If $\Pi_t = 0$ or 1 , then $\gamma^\xi(\Pi_t) = 0$ and consequently $\partial \gamma^\xi(\Pi_t) / \partial \xi = 0$. There are three additional ranges of Π_t that need to be considered:

1. If Π_t is such that $\hat{\theta}^\xi(\Pi_t) \in (0, 1)$, then $\alpha_s - (\xi + \alpha_c)p(\hat{\theta}^\xi(\Pi_t), \Pi_t) = 0$ so that

$$\partial \gamma^\xi(\Pi_t) / \partial \xi = \int_{\hat{\theta}^\xi(\Pi_t)}^1 p(\theta, \Pi_t) [f_q(\theta) - f_u(\theta)] d\theta > 0.$$

2. If Π_t is such that $\hat{\theta}^\xi(\Pi_t) = 0$, then $\gamma^\xi(\Pi_t) = (\alpha_c + \xi) \int_0^1 p(\theta, \Pi_t) [f_q(\theta) - f_u(\theta)] d\theta$ and $\partial \gamma^\xi(\Pi_t) / \partial \xi = \int_0^1 p(\theta, \Pi_t) [f_q(\theta) - f_u(\theta)] d\theta > 0$.
3. If Π_t is such that $\hat{\theta}^\xi(\Pi_t) = 1$, then $\gamma^\xi(\Pi_t) = 0$ and consequently $\partial \hat{\theta}^\xi(\Pi_t) / \partial \xi = 0$. □

Corollary 1. If $\Pi^{\text{sup}} \equiv \sup \{ \Pi | \hat{\theta}^\xi(\Pi) = 1 \}$, then Π^{sup} is decreasing in ξ .

Proof. In order for $\hat{\theta}^\xi(\Pi_t) = 1$, it must be that $(\alpha_c + \xi)p(1, \Pi_t) \leq \alpha_s$ (see eq. [16]). As ξ increases, the highest Π_t for which this inequality holds decreases. □

Lemma 1 and the associated corollary indicate that the unit period gain from investment is strictly increasing in ξ unless Π_t is such that $\hat{\theta}^\xi(\Pi_t) = 1$, and that the range of Π_t values that induces $\hat{\theta}^\xi(\Pi_t) = 1$ is decreasing in ξ . I conclude that, as ξ increases, the $\dot{\Gamma} = 0$ locus is stretched to the right and if there is a region of Π_t values for which $\gamma^\xi(\Pi_t) = 0$, its upper bound decreases as ξ increases. At some value of ξ , which I refer to as $\xi^{\min}$, only one steady state will remain and it will be at a relatively high investment rate. I refer to this steady state as $\Sigma^\xi(\Gamma^\xi, \Pi^\xi)$.

Lemma 2. If $G(0) > 0$, then there exists a $\xi^{\min}$ such that for all $\xi > \xi^{\min}$ the dynamic system under the hiring subsidy has only one steady state, which is a saddle point.

Proof. Lemma 1 indicates that $\gamma^\xi(\Pi_t')$ is strictly increasing in ξ . This result, along with the assumption that $G(0) > 0$, implies that there exists a $\xi^{\min}$ such that for all $\xi > \xi^{\min}$ there is only one steady state. □

Figure 4 depicts a potential phase diagram for the dynamic system when the hiring subsidy is such that there is only one steady state. This steady state is a saddle point as indicated by the arrows. Consequently, when a hiring subsidy of sufficient size is put in place, the disadvantaged group will move from the low steady state to the path that leads to $\Sigma^\xi(\Gamma^\xi, \Pi^\xi)$, the only remaining steady state. In each period that the policy is in place, the group will move along the path that approaches the steady state.

If the hiring subsidy is removed when the investment rate of the disadvantaged group is in the certain region for Σ_l , then the group will return to their initial steady state. If, on the other hand, the policy is removed after the disadvantaged group reaches the certain region for Σ_h , then they will converge to Σ_h and discrimination will be eliminated.

Figure 4: Phase diagram of the dynamic system under the hiring subsidy

It is important that Σ^ξ is not in the uncertain region, or else there is no way to guarantee the success of the hiring subsidy. The following lemma indicates that for a positive subsidy Σ^ξ is always in the certain region of Σ_h .

Lemma 3. *If $G(0) > 0$ and $\xi > \xi^{\min}$, then Σ^ξ will be in the certain region for Σ_h .*

Proof. From Lemma 1, $\gamma^\xi(\Pi_t^I)$ is strictly increasing in ξ . By assumption, $G(\cdot)$ is strictly increasing in its argument. Together these imply that Σ^ξ will always involve a higher investment rate than Σ_h . Also, the upper bound of the uncertain region is always below Π_h . As a result, Σ^ξ must be in the certain region of Σ_h because Π^ξ is greater than Π_h , which is always in the certain region. $\square$

Taking Lemma 3 into account, it is clear that the hiring subsidy policy will be successful as long as a large enough subsidy is offered.

Proposition 4. *If $G(0) > 0$ and $\xi > \xi^{\min}$, then the hiring subsidy will eliminate discrimination in the general dynamic model.*

Corollary 2. *This hiring subsidy can be removed as soon as Π_t is in the certain region for Σ_h .*

Figure 5: Path to Σ^ε

Figure 5 depicts the path to the remaining steady state from the disadvantaged group's starting point at the low steady state. See Appendix D.2 for a discussion of how these paths were generated. For comparison the figure also includes the original $\gamma(\Pi_t^I)$ function. I do not show the path to Σ_h once the policy is removed, but as long as the policy is removed after $\Pi_t > \hat{\Pi}$ then the disadvantaged group will simply move to the path that leads to Σ_h as it is the only plausible path given the current investment rate of the group.

I conclude that the hiring subsidy will always effectively eliminate discrimination as long as a large enough subsidy is offered. This policy eliminates discrimination by changing the optimal behavior of the firms. Specifically, the subsidy makes hiring disadvantaged workers more attractive and this is passed on to workers with an increase in the gain from investment. The end result is that the workers in the disadvantaged group are able to move to the high steady state.

Now that it has been confirmed that the hiring subsidy policy will eliminate discrimination it is worth analyzing the costs associated with implementation of the policy. As mentioned previously, the unit period cost of the hiring subsidy policy is $\kappa_t^\xi = \xi \lambda^d \Pi_t^d [1 - F_q(\hat{\theta}(\Pi_t^d))]$. This is simply the value of the subsidy times the expected number of workers from the disadvantaged group that will be correctly assigned to the complex task. Analysis of this per period cost reveals that

$$\frac{\partial \kappa_t^\xi}{\partial \xi} = \lambda^d \Pi_t^d [1 - F_q(\hat{\theta}(\Pi_t^d))] \geq 0, \quad [21]$$

$$\frac{\partial \kappa_t^\xi}{\partial \lambda^d} = \xi \Pi_t^d [1 - F_q(\hat{\theta}(\Pi_t^d))] \geq 0, \text{ and} \quad [22]$$

$$\frac{\partial \kappa_t^\xi}{\partial \Pi_t^d} = \xi \lambda^d \left[(1 - F_q(\hat{\theta}(\Pi_t^d))) - \Pi_t^d f_q(\hat{\theta}(\Pi_t^d)) \hat{\theta}'(\Pi_t^d) \right] \geq 0. \quad [23]$$

These results indicate that the larger the subsidy, the more costly this policy becomes. Thus the government has incentive to offer the lowest effective subsidy. Also, the larger the disadvantaged group is, the more expensive the policy becomes. Finally, the cost of the policy is also increasing in Π_t^d , as the more qualified workers there are, the more frequently the subsidy will have to be paid out. Note that the longer the policy is in place, the higher Π_t^d gets, as the disadvantaged group converges to the only remaining steady state. Consequently, the longer this policy is in place the more expensive it becomes per period.

4.2 Investment voucher program

The second policy of interest approaches the elimination of discrimination more directly. Rather than providing hiring incentives to the firms, an investment voucher program is designed to affect the decisions of the workers. The voucher is applied in the following way: The government provides all workers in the disadvantaged group with an *investment voucher* of v dollars. This voucher is only good toward the purchase of human capital, so it defrays the monetary cost of investment, c_m . Consequently, a voucher larger than c_m has no more impact than one equal to c_m . By implementing the voucher in this way, the government does not need to observe whether an individual worker invested or not. This policy is applied only to the disadvantaged group, so for simplicity of notation I drop the group subscripts in this section.

I can represent this investment voucher with a change in the effective lifetime gain from investment since the voucher is only used if a worker invests. The impact of the policy on the dynamic system is a strict increase in the investment rate of new workers at time t . In particular, the proportion of investors is $G(\Gamma_t + \min\{v, c_m\})$ under the voucher, rather than $G(\Gamma_t)$, as in the baseline model.

Consequently, under the investment voucher the equilibrium for the disadvantaged group will be described by a new dynamic system.

Proposition 5. *An equilibrium of the dynamic model under an investment voucher is fully characterized by the following two variable differential equations:*

$$\begin{aligned}\dot{\Gamma}_t &= (\beta + r)\Gamma_t - \gamma(\Pi_t) \\ \dot{\Pi}_t &= \beta(G(\Gamma_t + \min\{v, c_m\}) - \Pi_t).\end{aligned}\quad [24]$$

The only resulting change to the phase diagram will be in the $\dot{\Pi}_t = 0$ locus, which is updated to $\Pi_t = G(\Gamma_t + \min\{v, c_m\})$. Given that $G(\cdot)$ is strictly increasing, a negative voucher shifts the $\dot{\Pi}_t = 0$ locus to the right, so that the investment rate is lower at every Γ_t . A positive voucher shifts the locus to the left, so that investment rate is higher at every Γ_t .

Just as in the analysis of the hiring subsidy, in order for this policy to eliminate discrimination it must move the disadvantaged group to the certain region of Σ_h . The original dynamic model assumes a $G(\cdot)$ function that generates three steady states. The investment voucher policy does not change γ or $G(\cdot)$ but instead affects how investment decisions are made by increasing the lifetime gain from investment by $\min\{v, c_m\}$. The dynamic system under the investment voucher will have only one steady state if $\min\{v, c_m\}$ is sufficiently positive or negative.

Given the goal of moving the disadvantaged group to Σ_h , an effective investment voucher must be positive. Moreover, assuming that the monetary cost of investment, c_m , is high enough, at some value of v , which I refer to as $v^{\min}$, only one steady state will remain. I refer to this steady state as $\Sigma^v(\Gamma^v, \Pi^v)$. If the monetary cost of investment is low, then it is possible that even the largest effective voucher, $v = c_m$, is not able to eliminate the two lower steady states. In this case, the voucher policy is not effective at eliminating discrimination and no matter what v is employed there will be three steady states.

Lemma 4. *If the dynamic system under the largest effective investment voucher, $v = c_m$, has a single steady state, then there exists a $v^{\min}$ such that for all $v > v^{\min}$ the dynamic system under the investment voucher has only one steady state, which is a saddle point.*

Proof. $G(\Gamma_t + \min\{v, c_m\})$ is strictly increasing in v when $v < c_m$. As v increases, the $\dot{\Pi}_t = 0$ locus shifts to the left. Given a large enough c_m , there must exist a $v^{\min}$ such that for all $v > v^{\min}$ there is one steady state. $\square$

Figure 6 depicts a potential phase diagram for the dynamic system when the investment voucher is large enough to induce a single steady state. The arrows indicate that this steady state is a saddle point. Just as in the case of the hiring

Figure 6: Phase diagram of the dynamic system under the investment voucher

subsidy, the existence of a single steady state means that when the policy is put in place the disadvantaged group moves from the low steady state to the path that leads to $\Sigma^v(\Gamma^v, \Pi^v)$. In each period the policy is in effect, the group moves along the path approaching that steady state.

In order to successfully eliminate the discrimination, the policy can only be removed when the disadvantaged group is in the certain region for Σ_h . Again it is important that Σ^v is not in the uncertain region, or else there is no way to guarantee the success of the investment voucher. The following lemma indicates that for any voucher greater than $v^{\min}$, Σ^v will be in the certain region of Σ_h .

Lemma 5. *If $v > v^{\min}$, then Σ^v will be in the certain region for Σ_h .*

Proof. $G(\Gamma_t + \min\{v, c_m\})$ is strictly increasing in v when $v < c_m$. So, the $\dot{\Pi}_t = 0$ locus is always higher under the policy. This implies that Σ^v will involve a higher investment rate than Σ_h . Also, the upper bound of the uncertain region is always below Π_h . Σ^v must be in the certain region of Σ_h because Π^v is greater than Π_h , which is always in the certain region. $\square$

Taking Lemma 5 into account, it is clear that the investment voucher policy will be successful as long as monetary costs make up a large enough proportion of total costs and a sufficient voucher is offered.

Proposition 6. *If there exists a $v^{\min}$, then for all $v > v^{\min}$, the investment voucher will eliminate discrimination in the general dynamic model.*

Corollary 3. *This investment voucher policy can be removed as soon as Π_t is in the certain region for Σ_h .*

Figure 7 depicts the path to the remaining steady state from the disadvantaged group's starting point at Σ_l . See Appendix D.3 for a discussion of how these paths were generated. For comparison, the figure also includes the original $\dot{\Pi} = 0$ locus. As long as the policy is removed after $\Pi_t > \hat{\Pi}$, the disadvantaged group will simply move to the path that leads to Σ_h as it is the only plausible path given the firms' current beliefs about the group's overall investment rate.

Figure 7: Path to Σ^p

The investment voucher approaches the discrimination issue in a different manner than the hiring subsidy. It works by changing the incentives of the workers and making it easier for a worker's gain from investment to exceed his total cost. By adjusting the optimal behavior of workers, this policy allows the firms and workers to move to the high steady state.

Now that it has been confirmed that the investment voucher policy will eliminate discrimination it is worth analyzing the costs associated with implementation of the policy. The per period cost of the investment voucher policy is

$\kappa_t^v = v\lambda^d\beta G(\Gamma(\Pi_t^d) + v)$. This is simply the voucher times the fraction of new births that choose to invest under the policy. Analysis of this per period cost reveals that

$$\frac{\partial \kappa_t^v}{\partial v} = \lambda^d\beta G(\Gamma(\Pi_t^d) + v) + v\lambda^d\beta G'(\Gamma(\Pi_t^d) + v) \geq 0, \quad [25]$$

$$\frac{\partial \kappa_t^v}{\partial \lambda^d} = v\beta G(\Gamma(\Pi_t^d) + v) \geq 0, \quad [26]$$

$$\frac{\partial \kappa_t^v}{\partial \Pi_t^d} = v\beta G'(\Gamma(\Pi_t^d) + v)\Gamma'(\Pi_t^d) \geq 0, \quad \text{and} \quad [27]$$

$$\frac{\partial \kappa_t^v}{\partial \beta} = v\lambda^d G(\Gamma(\Pi_t^d) + v) \geq 0. \quad [28]$$

The cost of the voucher policy in each period is clearly increasing in the size of the voucher offered as well as in the size and average investment rate of the disadvantaged group. Also note, that under this policy the faster the population turns over, the higher the per period costs become. Just as in the case of the hiring subsidy policy, the longer the investment voucher policy is in place the more expensive it becomes per period, since the voucher must be paid out to an increasing number of new investors as time progresses.

4.3 Government-enforced equal treatment

The final policy of interest involves intervention by the governing body with respect to the optimal behavior of the firms. Specifically, under what I call the *equal treatment policy*, the government requires that both firms ignore a workers' group identity when assigning tasks and wages. This policy affects both disadvantaged and advantaged workers.

The equal treatment policy can be formally stated as: Firms are forbidden to condition task and wage assignments on group identity. I assume that the government is able to observe the firms' task and wage assignments and if they observe a deviation from the policy, they will charge the firm a substantial fine. As long as this fine is large enough, the firms will never have incentive to deviate.

Under the equal treatment policy, the firms' optimal behavior can no longer depend on group identity. So there will be a single $\hat{\theta}(\Pi_t)$ and a single $\omega_\Delta(\theta_t; \Pi_t)$ in each period, rather than one for each group.

The updated optimal threshold at time t is the one which maximizes a firm's output conditional on investment decisions. Formally, the task assignment problem is

$$\max_{\theta_t \in [0,1]} \Delta \alpha_c [(\lambda^d \Pi_t^d + \lambda^a \Pi_t^a)(1 - F_q(\theta_t))] + \Delta \alpha_s [(\lambda^d \Pi_t^d + \lambda^a \Pi_t^a) F_q(\theta_t) + (1 - (\lambda^d \Pi_t^d + \lambda^a \Pi_t^a)) F_u(\theta_t)]. \quad [29]$$

Just as in the original dynamic model there is a unique solution to the task assignment problem. But, in this case, it is not solved separately for each group. If I define a firm's beliefs about the entire population of workers as $\Pi^{pop} = \lambda^d \Pi_t^d + \lambda^a \Pi_t^a$, then the unique optimal threshold is

$$\hat{\theta}(\Pi_t^{pop}) = \begin{cases} 1 & \text{if } \alpha_c p(1, \Pi_t^{pop}) \leq \alpha_s, \\ 0 & \text{if } \alpha_c p(0, \Pi_t^{pop}) \geq \alpha_s, \\ \text{the unique solution to} \\ \alpha_c p(\theta, \Pi_t^{pop}) = \alpha_s & \text{if } \alpha_c p(0, \Pi_t^{pop}) < \alpha_s < \alpha_c p(1, \Pi_t^{pop}). \end{cases} \quad [30]$$

The continuation wage under the equal treatment policy will also depend on Π^{pop} rather than the group investment rates. The new optimal wage schedule for both firms is

$$\omega_\Delta(\theta_t; \Pi_t^{pop}) = \begin{cases} \Delta \alpha_s & \text{if } \theta_t < \hat{\theta}(\Pi_t^{pop}), \\ \Delta \alpha_c p(\theta_t, \Pi_t^{pop}) & \text{if } \theta_t \geq \hat{\theta}(\Pi_t^{pop}). \end{cases} \quad [31]$$

This change in the task assignment rule and the associated alteration of the continuation wage results in a new unit period gain from investment (and per period gain from investment). The updated unit period gain from investment is

$$\gamma(\Pi_t^{pop}) = \alpha_s [F_q(\hat{\theta}(\Pi_t^{pop})) - F_u(\hat{\theta}(\Pi_t^{pop}))] + \alpha_c \int_{\hat{\theta}(\Pi_t^{pop})}^1 p(\theta, \Pi_t^{pop}) [f_q(\theta) - f_u(\theta)] d\theta. \quad [32]$$

Under the policy, the lifetime gain from investment for each group will be determined by the population investment rate rather than the group investment rate. This causes a subsequent change in the dynamic system.

Proposition 7. *An equilibrium of the dynamic model under the equal treatment policy is fully characterized by the following differential equations:*

$$\dot{\Gamma}_t^{pop} = (\beta + r) \Gamma_t^{pop} - \gamma_t^{pop}(\Pi_t^{pop}).$$

$$\dot{\Pi}_t^{pop} = \beta(G(\Gamma_t^{pop}) - \Pi_t^{pop}). \quad [33]$$

The dynamic system in eq. [33] is identical to that of the original dynamic model. The only difference is that it describes the entire population rather than each individual group. In order to talk about when the equal treatment policy can be removed, it is also necessary to determine how a firm's beliefs about the investment rate of each group will change over time. Under the policy these beliefs are described by $\dot{\Pi}_t^I = \beta(G(\Gamma_t^{pop}) - \Pi_t^I)$.

Since I have already assumed that the original dynamic model has three steady states, with the high and the low steady states being saddle points, I know that the dynamic system under the equal treatment policy will as well. I call these three steady states $\Sigma_l^{pop}(\Gamma_l^{pop}, \Pi_l^{pop})$, $\Sigma_m^{pop}(\Gamma_m^{pop}, \Pi_m^{pop})$, and $\Sigma_h^{pop}(\Gamma_h^{pop}, \Pi_h^{pop})$ referring to the low, middle, and high steady states, respectively.

Lemma 6. *Under the equal treatment policy the dynamic system of the general model has three steady states. Σ_l^{pop} and Σ_h^{pop} are saddle points and Σ_m^{pop} is a source.*

The uncertain region, if present, is identical to that of the original dynamic model. With that in mind, it is straightforward to determine that this policy will always be able to eliminate discrimination. However, in some cases discrimination will be eliminated by moving both groups to the high steady state, in other cases both groups will end up in the low steady state. Clearly this is not a preferable outcome as the high steady state Pareto dominates the low steady state.

Given that the starting point of this analysis is one in which discrimination is present, when the policy is put in place the investment rate of the population is $\Pi_0^{pop} = \lambda^a \Pi_h + \lambda^d \Pi_l$. If this Π_0^{pop} is in the certain region for Σ_h , then the policy will be successful as the entire population will move to the only feasible path, which happens to lead to the high steady state. However if Π_0^{pop} is in the uncertain region of firm beliefs, then the policy will move the population to the high steady state if workers are optimistic and to the low steady state if workers are pessimistic. The policy is guaranteed to result in both groups at the low steady state if Π_0^{pop} is in the certain region for Σ_l .

Proposition 8. *If Π^{pop} is in the certain region for Σ_h or if Π^{pop} is in the uncertain region and the population is optimistic, then the equal treatment policy will eliminate discrimination by moving both groups to the Pareto-dominant steady state. Otherwise, the policy will move the two groups to the Pareto-dominated steady state.*

Corollary 4. *This equal treatment policy can be removed as soon as Π_t^a and Π_t^d are both in the certain region for Σ_h or for Σ_l .*

I can restate the results of Proposition 8 in terms of the size of the disadvantaged population. Specifically, if $\lambda^d < (\Pi_h - \hat{\Pi}/\Pi_h - \Pi_l)$, then the equal treatment policy is guaranteed to move both groups to the high steady state. I conclude that this policy will have a preferable outcome if the disadvantaged population is small or if the difference between Π_l and Π_h is small.³

This policy is different from the previous two because it impacts workers in both the advantaged and the disadvantaged groups. It is attractive to consider this policy because there is no direct cost associated with it. However, the policy has the potential to have the opposite of the intended effect if the initial population investment rate is in the certain region for Σ_l , or in the uncertain region and the population is pessimistic.

Now that it has been confirmed that the equal treatment policy can eliminate discrimination it is worth analyzing the costs associated with implementation of the policy. As stated earlier, the policy involves no direct monetary costs. However, it will result in a reduction in welfare for those in the advantaged group.

Welfare for those in the advantaged group when there is no equal treatment policy is

$$\begin{aligned}
 W(\Pi_t^a) = & \underbrace{\Pi_t^a}_{\text{Proportion of qualified workers}} \left[\underbrace{\int_0^1 \omega_\Delta(\theta, \Pi_t^a) f_q(\theta) d\theta}_{\text{Expected wage of qualified workers}} - \underbrace{\beta}_{\text{Fraction of new births}} \underbrace{\left(\frac{\int_c^{\Gamma(\Pi_t^a)} c g(c) dc}{G(\Gamma(\Pi_t^a))} \right)}_{\text{Expected cost given investment}} \right], \\
 & + \underbrace{(1 - \Pi_t^a)}_{\text{Proportion of unqualified worker}} \underbrace{\int_0^1 \omega_\Delta(\theta, \Pi_t^a) f_u(\theta) d\theta}_{\text{Expected wage of unqualified workers}},
 \end{aligned}
 \tag{34}$$

where given the assumptions about the starting point of the economy when the policy is put in place $\Pi_t^a = \Pi_h$. Under the equal treatment policy, all wages and task assignments are determined by the average investment rate of the

³ In an earlier draft of Kim and Loury (2014) they analyze this type of policy and draw the same conclusion.

population, Π^{pop} , rather than the individual group investment rate. So the welfare of the advantaged group under the equal treatment policy is

$$W^a(\Pi_t^a, \Pi_t^{pop}) = \Pi_t^a \left[\int_0^1 \omega_\Delta(\theta, \Pi_t^{pop}) f_q(\theta) d\theta - \beta \left(\frac{\int_c^{\Gamma(\Pi_t^{pop})} c g(c) dc}{G(\Gamma(\Pi_t^{pop}))} \right) \right] \\ + (1 - \Pi_t^a) \int_0^1 \omega_\Delta(\theta, \Pi_t^{pop}) f_u(\theta) d\theta. \quad [35]$$

The true cost of the equal treatment policy is the difference between welfare in the high steady state and the welfare of the advantaged group under the policy: $\kappa_t^{et} = W(\Pi_h) - W^a(\Pi_t^a, \Pi_t^{pop})$. In this general dynamic environment, there are no conclusions that I can draw regarding the impacts that $\lambda^d, \lambda^a, \Pi_t^d, \Pi_t^a$, or β have on the cost of the equal treatment policy. In Section 5 I examine a simple version of the dynamic model, within that structure I am able to be more specific regarding the costs of the equal treatment policy.

5 Analysis in the simple dynamic model

In this section I examine the welfare effects of the three policies in a simplified version of the dynamic model. First, I briefly describe the simple dynamic model. Next, I determine that all results from Section 3 and Section 4 hold for the simple dynamic model. Lastly, I use the simple model to analyze the welfare effects of each policy.

5.1 The simple dynamic model

Firms face the same linear per period production function as in the original model. For simplicity workers have only two possible signals: θ_H and θ_L where $\theta_H > \theta_L$. If a worker is qualified, they receive a signal of θ_H with probability $P_q > 1/2$ and a signal of θ_L with probability $1 - P_q$. If a worker is unqualified, they receive a signal of θ_L with probability $P_u > 1/2$ and a signal of θ_H with probability $1 - P_u$. The cost of investment is distributed uniformly over $[\underline{c}, \bar{c}]$. Time is discrete, so that $\Delta = 1$.

The equilibria of this example are characterized by the same dynamic system as in Proposition 2. Given the parameterization, I can find more specific forms of the relevant functions, $\hat{\theta}(\Pi_t^I)$ and $\gamma(\Pi_t^I)$. See Appendix B.1 for the updated functions for the threshold rule and unit period gain from investment.

In this simplified version of the dynamic model it is possible for there to be one or three steady states depending on the parameter values. For the analysis in the following sections I assume that there are three steady states.⁴ Just as in the general dynamic model, the high and low steady states are saddle points and the middle steady state is an unstable source. See Appendix B.1 for a sample phase diagram for the simple dynamic model as well as a figure demonstrating the paths to the high and low steady states.

5.2 Policies in the simple dynamic model

If I analyze the three policies using the simple dynamic model described above I find that all major results from Section 4 hold. In particular, Propositions 3, 4, 5, 6, 7, and 8 are all confirmed. Using the simple dynamic model I can determine more specific functions for the dynamic systems in Propositions 3, 5, and 7 as well as for the minimum required hiring subsidy and the minimum required investment voucher mentioned in Propositions 4 and 6.

The minimum required subsidy is $\xi^{\min} = \alpha_s \left[1 - \frac{1-P_u}{P_q} \left(\frac{\bar{c}}{\underline{c}} \right) \right] - \alpha_c$. Note that $\xi^{\min}$ is decreasing in the precision of signals (P_u and P_q) as well as in α_c ; it is increasing in α_s . The minimum required voucher is $v^{\min} = (\bar{c} - \underline{c}) \left[\frac{\alpha_s(1-P_u)}{\alpha_s(1-P_u) + (\alpha_c - \alpha_s)P_q} \right] + \underline{c}$. This minimum voucher is decreasing in the precision of signals (P_u and P_q) as well as in α_c , and it is increasing in α_s . For a full description of the effects of the three policies in the simple dynamic model see Appendix B.

5.3 Welfare analysis of the three policies in the simple dynamic model

I am now able to look at the welfare effects of the three policies. The numerical and graphical results detailed in this section are for one specific parameterization of the dynamic model: $P_q = 2/3$, $P_u = 2/3$, $\alpha_s = 1$, $\alpha_c = 2$, $\beta = 0.2$, $r = 0.05$, $\underline{c} = -0.1$, and $\bar{c} = 0.9$. However, the qualitative results hold for all parameterizations that generate three steady states with the high and low being saddle points.

⁴ There are many sets of parameter values that generate three steady states in this simple parameterization. Furthermore, some of these parameterizations will generate spiralling paths and others will generate deterministic paths. See Glawtschew (2013) for further discussion.

I complete the welfare analysis in a general equilibrium framework. The hiring subsidy and investment voucher policies require that the government has funds to distribute to the firms and to the workers in the disadvantaged group, respectively. Given the general equilibrium framework, these funds are collected by the government via a tax levied on workers in the advantaged group or to state it more generally, a tax levied on workers who are offered the better continuation wage schedule.⁵

For the purpose of more meaningful comparison between the three policies, I apply the minimum necessary voucher and minimum necessary subsidies. For the parametric example $v^{\min} = 7/30$ and $\xi^{\min} = 7/2$. I also assume that the population is optimistic so as to guarantee that the equal treatment policy is successful. Furthermore I assume that the population will follow the shortest possible path to the high steady state. The government also commits to employing each policy until the impacted groups reach their new steady state.

Table 1 describes the conditions of the advantaged and disadvantaged groups before a policy is applied. See Appendix D.1 for a discussion of how these values were calculated.

Table 1: Details of the discriminatory equilibrium

$P_u = P_q = 2/3, \lambda^a = 3/4, c: U[-0.1, 0.9]$ $\alpha_s = 1, \alpha_c = 2, \beta = 0.2, r = 0.05$	Discriminatory equilibrium	
	Disadvantaged	Advantaged
Steady-state investment	$\Pi^d = 0.1000$	$\Pi^a = 0.7631$
Lifetime gain from investment	$\Gamma^d(\Pi^d) = 0$	$\Gamma^a(\Pi^a) = 0.663$
Wages	$\omega(\theta_H, \Pi^d) = 1$ $\omega(\theta_L, \Pi^d) = 1$	$\omega(\theta_H, \Pi^a) = 1.7313$ $\omega(\theta_L, \Pi^a) = 1.2340$
Average per period expected welfare	1.0010	1.4833

Figures 8–10 depict the average expected welfare of each group and of the population as a whole under the subsidy, voucher, and equal treatment policies, respectively. Each figure shows the welfare before the policy is in place, how welfare changes when the policy is applied, and then when the policy is removed. All three policies move those in the disadvantaged group to the high steady state, so the end result of each policy is that the welfare across groups is

5 This assumption is not crucial to the analysis that follows. I can achieve the same qualitative results about welfare if the policy is instead funded by a tax levied on all workers. I cannot however assume that the tax is levied on “high-wage” workers as that will have an impact on the incentives of workers.

Figure 8: Welfare effects of the hiring subsidy

Figure 9: Welfare effects of the investment voucher

Figure 10: Welfare effects of the equal treatment policy

equal and the discrimination is eliminated. See Appendices D.2, D.3, and D.4 for a discussion of how these values were calculated.

Comparing across figures indicates that the hiring subsidy is the most costly of the three policies. The reason for this is twofold. First, the transfer that is necessary to fund the policy is very large. Second, the hiring subsidy is provided for each worker correctly employed in the complex task in every period the policy is in place. The result is that the advantaged group loses a significant amount of welfare over the course of the policy. However, this policy will get the disadvantaged group to the steady state in the shortest amount of time.

The voucher policy is the least costly of the three policies. This is because the transfer necessary to support the policy is very low and in each period it is only paid to new workers that choose to invest. This policy needs to be in place for slightly longer than the hiring subsidy policy. The equal treatment policy also has a relatively small negative impact on the advantaged group, but it has the additional benefit of not requiring the government to redistribute income across groups. However, the equal treatment policy must be in place for the longest amount of time.

Table 2 shows the total change in welfare for each group and the population while each policy is in place. These numbers are simply the sum of the welfare gains of the disadvantaged group and the sum of the welfare losses of the advantaged group for each period that the policy is in place. The net welfare effect is the weighted sum of these two values, taking into account the relative sizes of the two groups. This table indicates that, while the subsidy does increase the disadvantaged group’s welfare greatly, the net result over the course of the policy is the least favorable. The most favorable, in terms of net change in welfare, is the equal treatment policy and the voucher policy falls somewhere in the middle.

Table 2: Total change in welfare while the policies are in effect

	Group <i>d</i> welfare gain	Group <i>a</i> welfare loss	Net change in pop. welfare
Hiring subsidy	61.908	17.855	2.086
Investment voucher	12.251	0.329	2.816
Equal treatment	18.495	0.887	3.959

While the numerical values of the results above are specific to the chosen parameters, the qualitative results are not. For all parameterizations of the simple dynamic model that generate three steady states and allow for all three

policies to be successful, the hiring subsidy is the most costly to implement but leads to the fastest convergence. Similarly, the investment voucher is always the least costly policy while the equal treatment policy needs to be in place the longest, but results in the most positive net change in population welfare. See Appendix D.5 for welfare results for a sampling of parameterizations.

I can conclude that depending on policy makers' preferences, it may be that the investment voucher, equal treatment policy, or hiring subsidy is preferred. If the policy maker wants to inconvenience the advantaged group the least, then the investment voucher is best. If his goal is to ensure that net welfare is improved the most, then the equal treatment policy is best. If he wants the policy in place for the shortest amount of time then the hiring subsidy is preferred.

It is important to note that both the investment voucher and equal treatment policies can fail to be successful for some parameterizations. In particular, if the monetary cost of investment is low, then it may be that there is no voucher that can eliminate the low steady state. Similarly, if the population is pessimistic, then the equal treatment policy will have the suboptimal outcome of moving both groups to the low steady state. The subsidy policy, on the other hand, will always have the intended effect given that a large enough subsidy is applied.

I can also draw a few conclusions about how the different parameters impact welfare gains, losses, and policy length. The larger the disadvantaged group is, the longer the equal treatment policy will need to be in effect. This is because the larger the disadvantaged group is, the lower the Π^{pop} becomes and the further the population is from the steady state once the policy is put in place. Across all three policies, the welfare loss to the advantaged group is increasing in the size of the disadvantaged group. Additionally, for all policies, the lower β is, the longer a policy needs to be in place and the smaller the welfare losses are.

6 Conclusion

The main contribution of this paper is to analyze the effectiveness of several government policies designed to eliminate discrimination in a dynamic environment. Unlike policy analysis in a static environment, I am able to determine the paths that each group will take once a policy is put in place and to address how long a policy needs to be in effect. The base model allows for competitive wages, and as a result I am also able to perform meaningful welfare analysis of the

three policies in a general equilibrium setting. These results are a significant and important addition to the current literature.

The dynamic environment I develop and analyze in Sections 2 and 3 is what makes these results possible. If I were to analyze the same policies in a static setting, then I would not be able to draw any conclusions about whether each policy would be successful, as it would be impossible to predict which equilibrium the two groups would re-coordinate on after the policy was put in place. Furthermore, if a policy happened to cause both groups to re-coordinate on the Pareto-dominant equilibrium, there would be no way to predict how the groups would react to the removal of the policy. A static environment offers no information about the paths each group would take when a policy is put in place or when it is removed. Consequently, without the addition of dynamics to the model, I would not be able to draw any conclusions about the welfare effects of the policies.

Using the dynamic model developed in Sections 2 and 3, I find that the hiring subsidy policy will always effectively eliminate discrimination but it results in a significant loss in welfare to the advantaged group while in place. However, this policy does have the advantage of needing to be in effect for the shortest amount of time. I also find that the investment voucher is a successful policy as long as a large enough proportion of the investment costs are monetary in nature. This policy may be attractive to policy makers because it involves very little loss in welfare to the advantaged group.

The equal treatment policy may or may not be successful depending on several conditions. The policy is safe to implement in industries where a majority of the workers are in the advantaged group or if the population as a whole is optimistic. If this is the case, then it may be the preferred policy, as it requires no transfers from advantaged to disadvantaged workers and welfare losses are relatively small. However, the equal treatment policy should not be applied if a large proportion of workers are disadvantaged or if the population as a whole is pessimistic.

I also determine that all three policies need only be in effect until the investment rates of both groups are in the certain region of the high steady state. Once this occurs the policies can be safely removed and the groups will converge to the high steady state on their own.

Acknowledgments: I thank my dissertation advisor, Peter Norman, for his direction and support during this project. I am also grateful to the remainder of my committee Gary Biglaiser, R. Vijay Krishna, Sergio Parraries, and Helen Tauchen for their time and effort.

Appendix

A Proof of Proposition 1

The following proof is a slightly modified version of one found in Moro and Norman (2002, 2003). I make only slight adjustments to suit my purposes, as such all credit should be given to them. Due to the linear nature of the firms' production function, I will address each group separately. For simplicity in notation I drop all group indicators J .

Proof (Sufficiency). Given $\pi \in (0, 1]$, let $\theta(\pi)$ be the unique solution to the task assignment problem, $\tau : [0, 1] \rightarrow [0, 1]$ be the threshold rule with cut-off $\theta(\pi)$, and $(C(\pi), S(\pi)) = (\pi F_q(\theta), \pi F_q(\theta) + (1 - \pi)F_u(\theta))$ be the associated aggregate factor inputs. Suppose that each firm posts the wage schedule $\omega : [0, 1] \rightarrow \Re$ given by $\omega(\theta; \pi)$ in eq. [4]. Moreover, suppose that workers facing indifference between working for firm 1 and firm 2 based on wage and task assignments (in equilibrium this will be all workers) simply flip a fair coin. This implies that

$$\begin{aligned} C_1(\pi) &= C_2(\pi) = \frac{C(\pi)}{2}, \\ S_1(\pi) &= S_2(\pi) = \frac{S(\pi)}{2}. \end{aligned} \quad [36]$$

The profit for firm 1 is (firm 2 is equivalent)

$$\begin{aligned} \Pi_1 &= \alpha_s S_1(\pi) + \alpha_c C_1(\pi) - \frac{1}{2} \int_0^1 \omega(\theta; \pi) [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta \\ &= \alpha_s S_1(\pi) + \alpha_c C_1(\pi) - \frac{1}{2} \int_0^{\theta(\pi)} \alpha_s [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta \\ &\quad - \frac{1}{2} \int_{\theta(\pi)}^1 \alpha_c \underbrace{\frac{\pi f_q(\theta)}{\pi f_q(\theta) + (1 - \pi)f_u(\theta)}}_{p(\theta, \pi')} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta \end{aligned} \quad [37]$$

$$\begin{aligned}
\Pi_1 &= \alpha_s S_1(\pi) + \alpha_c C_1(\pi) - \frac{1}{2} \left[\alpha_s \int_0^{\theta(\pi)} [\pi f_q(\theta) + (1-\pi)f_u(\theta)] d\theta + \int_{\theta(\pi)}^1 \pi f_q(\theta) d\theta \right] \\
&= \alpha_s S_1(\pi) + \alpha_c C_1(\pi) - \frac{1}{2} [\alpha_s S(\pi) + \alpha_c C(\pi)] = 0.
\end{aligned}
\tag{38}$$

Suppose one firm deviates to $(\omega', \tau') \neq (\omega, \tau)$. Let C' and S' denote the implied factor inputs and let $a(\theta) \in [0, 1]$ denote the fraction of workers with signal θ that accept a job at the deviating firm. Since $\omega'(\theta; \pi) \geq \omega(\theta; \pi)$ for all θ such that $a(\theta) > 0$ the profit for the deviating firm, Π'_i , satisfies

$$\Pi'_i \leq \alpha_s S' + \alpha_c C' - \int_0^1 \omega(\theta; \pi) a(\theta) [\pi f_q(\theta) + (1-\pi)f_u(\theta)] d\theta,
\tag{39}$$

where (here the assumption that ties are broken the same way by qualified and unqualified workers is used)

$$C' = \int \tau'(\theta) \pi f_q(\theta) a(\theta) d\theta \quad \text{and} \quad S' = \int (1 - \tau'(\theta)) a(\theta) [\pi f_q(\theta) + (1-\pi)f_u(\theta)] d\theta.
\tag{40}$$

Moreover $\omega(\theta; \pi) = \max\{\alpha_c p(\theta, \pi), \alpha_s\}$ so,

$$\begin{aligned}
&\int \omega(\theta; \pi) a(\theta) [\pi f_q(\theta) + (1-\pi)f_u(\theta)] d\theta \\
&= \int \tau'(\theta) \omega(\theta; \pi) a(\theta) [\pi f_q(\theta) + (1-\pi)f_u(\theta)] d\theta \\
&\quad + \int (1 - \tau'(\theta)) \omega(\theta; \pi) a(\theta) [\pi f_q(\theta) + (1-\pi)f_u(\theta)] d\theta \\
&\geq \alpha_s \int (1 - \tau'(\theta)) a(\theta) [\pi f_q(\theta) + (1-\pi)f_u(\theta)] d\theta + \alpha_c \int \tau'(\theta) \pi f_q(\theta) a(\theta) d\theta \\
&= \alpha_s S' + \alpha_c C'.
\end{aligned}
\tag{41}$$

This implies that $\Pi'_i \leq \alpha_s S' + \alpha_c C' - [\alpha_s S' + \alpha_c C'] = 0$. □

To prove necessity of the conditions in Proposition 1 I proceed by proving a sequence of intermediate results:

Lemma 7. *Each firm earns zero profit in equilibrium.*

Proof. Let Π_1 and Π_2 denote the profits for firm 1 and firm 2 and assume for contradiction that $\Pi_1 > 0$. If $\Pi_2 < 0$ there would be a profitable deviation, so I assume without loss that $0 \leq \Pi_2 \leq \Pi_1$. Total industry profits are

$$\begin{aligned} \Pi_1 + \Pi_2 &= \alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 \\ &\quad - \int_{\theta} \max\{\omega_1(\theta), \omega_2(\theta)\} [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta. \end{aligned} \quad [42]$$

I observe that $\alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 \leq \alpha_s S(\pi) + \alpha_c C(\pi)$. This simply means that aggregate output cannot exceed what a planner could achieve. Suppose firm 2 deviates by offering ω'_2 given by $\omega'_2(\theta) = \max\{\omega_1(\theta), \omega_2(\theta)\} + \varepsilon$ for some $\varepsilon > 0$, implying that firm 2 attracts all workers. In addition suppose firm 2 assigns as in the solution to eq. [2]. The corresponding profit is

$$\begin{aligned} \Pi'_2 &= \alpha_s S(\pi) + \alpha_c C(\pi) - \int_{\theta} \max\{\omega_1(\theta), \omega_2(\theta)\} [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \varepsilon \\ &= \alpha_s S(\pi) + \alpha_c C(\pi) - [\alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2] + \Pi_1 + \Pi_2 - \varepsilon \\ &\geq \Pi_1 + \Pi_2 - \varepsilon. \end{aligned} \quad [43]$$

Hence for ε sufficiently small $\Pi'_2 > \Pi_2$ is a profitable deviation, thus proving the statement by contradiction.

Lemma 8. $\omega_1(\theta) = \omega_2(\theta)$ for almost all $\theta \in [0, 1]$ in any equilibrium.

Proof. For contradiction, suppose that $\omega_1(\theta) > \omega_2(\theta)$ for all $\theta \in \Theta' \subset [0, 1]$ and let

$$\beta = \int_{\theta \in \Theta'} (\omega_1(\theta) - \omega_2(\theta)) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta, \quad [44]$$

where Θ' has a positive measure, which implies that $\beta > 0$. Let C_1, C_2, S_1, S_2 be the effective factor inputs in the hypothetical equilibrium and suppose that firm 1 deviates and offers ω'_1 given by $\omega'_1(\theta) = \omega_2(\theta) + \varepsilon$ for all θ

and assigns all workers (the deviation attracts all workers) in accordance with eq. [2], the solution to the task assignment problem. The implied profits are

$$\begin{aligned}
 \Pi'_1(\varepsilon) &= \alpha_s S(\pi) + \alpha_c C(\pi) - \int_{\theta} \omega_2(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \varepsilon \\
 &= \alpha_s S(\pi) + \alpha_c C(\pi) - \int_{\theta \in \Theta'} \omega_1(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta + \beta \\
 &\quad - \int_{\theta \in [0,1] \setminus \Theta'} \omega_2(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \varepsilon \\
 &\geq \alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 - \int_{\theta} \max\{\omega_1(\theta), \omega_2(\theta)\} [\pi f_q(\theta) \\
 &\quad + (1 - \pi) f_u(\theta)] d\theta + \beta - \varepsilon \\
 &= \beta - \varepsilon,
 \end{aligned} \tag{45}$$

where the last inequality follows from Lemma 7. The deviation is thus profitable if ε is small enough, thus the equilibrium where $\omega_1(\theta) > \omega_2(\theta)$ cannot be optimal.

Lemma 9. $\alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 = \alpha_s S(\pi) + \alpha_c C(\pi)$.

Proof. By feasibility $\alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 \leq \alpha_s S(\pi) + \alpha_c C(\pi)$, so assume for contradiction that $\alpha_s S(\pi) + \alpha_c C(\pi) - \alpha_s S_1 - \alpha_c C_1 - \alpha_s S_2 - \alpha_c C_2 = \beta > 0$. Suppose firm 1 offers $\omega'_1(\theta) = \omega_2(\theta) + \varepsilon$ for all θ and assigns all workers (the deviation attracts all workers) in accordance with eq. [2], the solution to the task assignment problem. The implied profits are

$$\begin{aligned}
 \Pi'_1(\varepsilon) &= \alpha_s S(\pi) + \alpha_c C(\pi) - \int_{\theta} \omega_2(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \varepsilon \\
 &> \alpha_s S_1 + \alpha_c C_1 + \alpha_s S_2 + \alpha_c C_2 - \int_{\theta} \omega_2(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \varepsilon.
 \end{aligned} \tag{46}$$

But (Lemma 8) $\omega_1(\theta) = \omega_2(\theta)$ almost everywhere so $\int_{\theta} \omega_2(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta$ is the sum of wages paid out by firms 1 and 2 before the deviation. By zero

profits (Lemma 7) this implies that $\Pi'(\varepsilon) = \beta - \varepsilon$, so for ε small enough the deviation is profitable. $\square$

Lemma 10. Suppose (ω_1, ω_2) is a pair of equilibrium wage schedules and let $\theta(\pi)$ be as in eq. [3]. Then there is a pair (k_s, k_c) such that (1) $\omega_i(\theta) = k_s$ for $i = 1, 2$ and for almost all $\theta < \theta(\pi)$, (2) $\omega_i(\theta) = p(\theta, \pi)k_c$ for $i = 1, 2$ and for almost all $\theta \geq \theta(\pi)$.

Proof. The two parts have almost identical proofs, I will prove only part (2), which may appear as less obvious. Let $\omega(\theta) = \max\{\omega_1(\theta), \omega_2(\theta)\}$ and $(C(\pi), S(\pi))$ be the factor inputs corresponding to eq. [2] the unique solution to the task assignment problem. For contradiction suppose there is a set $A \subset [\theta(\pi), 1]$, where $m = \int_A \pi f_q(\theta) d\theta > 0$, and for some $\beta > 0$ such that for all $\theta \in A$,

$$\begin{aligned} \frac{\omega(\theta)}{p(\theta, \pi)} &\leq \frac{1}{1 - F_q(\theta(\pi))} \int_{\theta(\pi)}^1 \frac{\omega(\theta)}{p(\theta, \pi)} f_q(\theta) d\theta - \beta \\ &= \frac{1}{\pi(1 - F_q(\theta(\pi)))} \int_{\theta(\pi)}^1 \omega(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \beta \\ &\quad + \frac{1}{C(\pi)} \int_{\theta(\pi)}^1 \omega(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \beta. \end{aligned} \quad [47]$$

By continuity there exists a set $B \in [0, \theta(\pi))$ such that $\int_B [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta = \frac{S(\pi)}{C(\pi)} m$ and

$$\begin{aligned} \omega(\theta) &\leq \frac{1}{\pi F_q(\theta) + (1 - \pi) F_u(\theta)} \int_{\theta(\pi)}^1 \omega(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \\ &= \frac{1}{S(\pi)} \int_0^{\theta(\pi)} \omega(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \end{aligned} \quad [48]$$

for every $\theta \in B$. Consider a deviation by firm i where it offers $\omega'_i(\theta) = \omega(\theta) + \varepsilon$ to workers with $\theta \in A \cup B$ and $\omega'_i(\theta) = 0$ for all other θ , and assigns workers from A to the complex task and workers from B to the simple task. The profit from this deviation is

$$\begin{aligned}
\Pi' &= \alpha_c \int_A \pi f_q(\theta) d\theta + \alpha_s \int_B [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \\
&\quad - \int_{\theta \in A \cup B} (\omega(\theta) + \varepsilon) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \\
&\geq [\alpha_c C(\pi) + \alpha_s S(\pi)] \frac{m}{C(\pi)} \\
&\quad - \left(\frac{1}{C(\pi)} \int_{\theta(\pi)}^1 \omega(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \beta \right) \\
&\quad \int_{\theta \in A} \underbrace{p(\theta, \pi) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)]}_{=\pi f_q(\theta)} d\theta \\
&\quad - \int_{\theta \in B} [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \left[\frac{1}{S(\pi)} \int_0^{\theta(\pi)} \omega(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] \right] \\
&\quad - \varepsilon \int_{\theta \in A \cup B} [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta
\end{aligned} \tag{49}$$

$$\begin{aligned}
&= [\alpha_c C(\pi) + \alpha_s S(\pi)] \frac{m}{C(\pi)} - \left(\int_{\theta(\pi)}^1 \omega(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta - \beta \right) \frac{m}{C(\pi)} \\
&\quad - \frac{S(\pi)}{C(\pi)} m \left[\frac{1}{S(\pi)} \int_0^{\theta(\pi)} \omega(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \right] \\
&\quad - \varepsilon \int_{\theta \in A \cup B} [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta \\
&= \frac{m}{C(\pi)} \left(\underbrace{\alpha_c C(\pi) + \alpha_s S(\pi) - \int_{\theta \in [0,1]} \omega(\theta) [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta}_{= 0 \text{ by Lemmas 7 and 8}} \right) \\
&\quad + \beta \frac{m}{C(\pi)} - \varepsilon \int_{\theta \in A \cup B} [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta.
\end{aligned} \tag{50}$$

Hence, $\Pi' \geq \beta \frac{m}{C(\pi)} - \varepsilon \int_{\theta \in A \cup B} [\pi f_q(\theta) + (1 - \pi) f_u(\theta)] d\theta > 0$ for ε small enough, which together with Lemma 8 establishes part (2) of the claim. The proof for

the other half is symmetric. It's simply a matter of removing β from eq. [47] and inserting a β in the inequality in eq. [48] and again constructing an A and B such that the factor ratio is as in the unique threshold in eq. [2]. The rest of the argument is unaltered.

Proof (Necessity). It remains to be shown that $k_s = \alpha_s$ and $k_c = P(\theta, \pi)\alpha_c$. Firms would make positive profits if $k_s < \alpha_s$ and $k_c < P(\theta, \pi)\alpha_c$ and negative profits if the inequalities go the other way. Thus I must only consider cases where the inequalities are of opposing directions. The arguments are symmetric so I only consider the case with $k_s > \alpha_s$ and $k_c < P(\theta, \pi)\alpha_c$. If $\theta(\pi) = 0$, (1) each firm makes positive profits (loss), so the only case to consider is when $\theta(\pi)$ is interior. A necessary condition for optimality for problem (2) is that $\alpha_c P(\theta(\pi), \pi) = \alpha_s$. Hence there must be an interval $(\theta(\pi), \theta^*)$ where $\omega_i(\theta) = p(\theta, \pi)k_c < k_s$ for all $\theta \in (\theta(\pi), \theta^*)$. Consider the deviation

$$\omega'_i(\theta) = \begin{cases} \omega(\theta) & \text{for } \theta \in (\theta(\pi), \theta^*) \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \tau'_i(\theta) = \begin{cases} 0 & \text{for } \theta \in (\theta(\pi), \theta') \\ 1 & \text{for } \theta \in (\theta', \theta^*), \end{cases} \quad [51]$$

where θ' is set so that the factor ratio is as in the solutions to eq. [2],

$$\left(\frac{\int_{\theta'}^{\theta^*} \pi f_q(\theta) d\theta}{\int_{\theta(\pi)}^{\theta'} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta} = \frac{C(\pi)}{S(\pi)} \right). \quad [52]$$

The profit is

$$\begin{aligned} \Pi' &= \alpha_c \int_{\theta'}^{\theta^*} \pi f_q(\theta) d\theta + \alpha_s \int_{\theta(\pi)}^{\theta'} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta \\ &\quad - \int_{\theta(\pi)}^{\theta'} \omega(\theta) [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta - \int_{\theta'}^{\theta^*} \omega(\theta) [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta \\ &> \frac{(\alpha_c C(\pi) + \alpha_s S(\pi)) \int_{\theta'}^{\theta^*} \pi f_q(\theta) d\theta}{C(\theta)} - \alpha_c \int_{\theta(\pi)}^{\theta'} [\pi f_q(\theta) + (1 - \pi)f_u(\theta)] d\theta - \alpha_s \int_{\theta'}^{\theta^*} \pi f_q(\theta) d\theta \\ &= \frac{1}{C(\pi)} \int_{\theta'}^{\theta^*} \pi f_q(\theta) d\theta [\alpha_c C(\pi) + \alpha_s S(\pi) - \alpha_s S(\pi) - \alpha_c C(\pi)] = 0. \end{aligned} \quad [53]$$

which completes the proof of Proposition 1. $\square$

B Simple dynamic model

B.1 Phase diagram of the simple dynamic model

The equilibria of the simple dynamic described in 5.1 are characterized by the same dynamic system as in Proposition 2. Given the parameterization, I can find more specific forms of the relevant functions. The updated optimal threshold rule is

$$\hat{\theta}(\Pi_t^J) = \begin{cases} 1 & \text{if } \Pi_t^J \leq \underline{\Pi} \\ \theta_H & \text{if } \underline{\Pi} < \Pi_t^J < \bar{\Pi} \\ 0 & \text{if } \Pi_t^J \geq \bar{\Pi}, \end{cases} \quad [54]$$

where $\underline{\Pi}$ and $\bar{\Pi}$ are calculated by updating $\hat{\theta}(\Pi_t^J)$ in eq. [3]; they are

$$\underline{\Pi} = \frac{\alpha_s(1 - P_u)}{\alpha_s(1 - P_u) + (\alpha_c - \alpha_s)P_q} \quad \text{and} \quad \bar{\Pi} = \frac{\alpha_s P_u}{\alpha_s P_u + (\alpha_c - \alpha_s)(1 - P_q)}. \quad [55]$$

The unit period gain from investment is updated to

$$\gamma(\Pi_t^J) = \begin{cases} 0 & \text{if } \Pi_t^J \leq \underline{\Pi} \\ (P_q + P_u - 1)(\alpha_c P(\theta_H, \Pi_t^J) - \alpha_s) & \text{if } \underline{\Pi} < \Pi_t^J < \bar{\Pi} \\ \alpha_c(P_q + P_u - 1)(P(\theta_H, \Pi_t^J) - P(\theta_L, \Pi_t^J)) & \text{if } \Pi_t^J \geq \bar{\Pi}. \end{cases} \quad [56]$$

The posterior probability that a worker with signal θ_H is qualified is

$$P(\theta_H, \Pi_t^J) = \frac{\Pi_t^J P_q}{\Pi_t^J P_q + (1 - \Pi_t^J)(1 - P_u)}, \quad [57]$$

and the posterior probability that a worker with signal θ_L is qualified is

$$P(\theta_L, \Pi_t^J) = \frac{\Pi_t^J(1 - P_q)}{\Pi_t^J(1 - P_q) + (1 - \Pi_t^J)P_u}. \quad [58]$$

The two loci of the dynamic system are now $\Gamma_t^J = \frac{\gamma(\Pi_t^J)}{\beta + r}$, where $\gamma(\Pi_t^J)$ is as in eq. [56], and $\Pi_t^J = G(\Gamma_t^J)$, where G is the uniform distribution over $[\underline{c}, \bar{c}]$. Using parameters $P_q = 2/3$, $P_u = 2/3$, $\alpha_s = 1$, $\alpha_c = 2$, $\beta = 0.2$, $r = 0.05$, $\underline{c} = -0.1$, and $\bar{c} = 0.9$, these generate the phase diagram in Figure 11.

Figure 12 demonstrates the spiralling paths to the two saddle points. The steady-state beliefs for this particular parametric example are $\Pi_t = 0.1$,

Figure 11: Phase diagram for the parametric example

Figure 12: Spiralling paths to the steady states in the parametric example

$\Pi_m = 0.47$, and $\Pi_h = 0.76$. The paths are generated by inducting backward from the high and low steady states using eqs [6] and [10]. See Appendix C.2 for further discussion of how these paths were generated. The figure indicates that there is an uncertain region (designated by the purple region) that results from the spiralling paths. There are also certain regions for both Σ_h (the pink area) and Σ_l (the blue area). These certain regions will help determine how long a potential policy needs to be in effect to guarantee the elimination of discrimination.

B.2 Effects of the hiring subsidy in the simple dynamic model

The equilibria of this simple model are described by the same dynamic system as in Proposition 3. The updated forms of the relevant functions are

$$\gamma^\xi(\Pi_t) = \begin{cases} 0 & \text{if } \Pi_t \leq \underline{\Pi}^\xi \\ (P_q + P_u - 1)[(\alpha_c + \xi)P(\theta_H, \Pi_t) - \alpha_s] & \text{if } \underline{\Pi}^\xi < \Pi_t < \overline{\Pi}^\xi \\ (P_q + P_u - 1)(\alpha_c + \xi)[P(\theta_H, \Pi_t) - P(\theta_L, \Pi_t)] & \text{if } \Pi_t \geq \overline{\Pi}^\xi, \end{cases} \quad [59]$$

where $\underline{\Pi}^\xi$ and $\overline{\Pi}^\xi$ are calculated by updating $\hat{\theta}(\Pi_t^\xi)$ in eq. [16]. They are

$$\underline{\Pi}^\xi = \frac{\alpha_s(1 - P_u)}{\alpha_s(1 - P_u) + (\alpha_c + \xi - \alpha_s)P_q} \quad \text{and} \quad \overline{\Pi}^\xi = \frac{\alpha_s P_u}{\alpha_s P_u + (\alpha_c + \xi - \alpha_s)(1 - P_q)} \quad [60]$$

Figure 13 depicts the phase diagram of the new dynamic system given that $\xi = 6$. All other parameter values are as in Section 5.1. The result of this hiring subsidy is a dynamic system with only one steady state, which I refer to as $\Sigma^\xi(\Gamma^\xi, \Pi^\xi)$. The arrows indicate that Σ^ξ is a saddle point. Consequently, all results from Section 4.1 hold for this parametric example.

The minimum subsidy necessary for the dynamic system to have only one steady state is $\xi^{\min} = \alpha_s \left[\frac{1 - P_u}{P_q} \frac{1 - G(0)}{G(0)} + 1 \right] - \alpha_c$. Note that $\xi^{\min}$ is decreasing in the precision of signals (P_u and P_q) as well as in α_c ; it is increasing in α_s .

Figure 14 demonstrates the path the disadvantaged group will take once the policy is put in place as well as the group's starting point. Figure 15 demonstrates the path the group will take to Σ_h assuming that the policy is removed after the group reaches Σ^ξ , see Appendix D.2 for a discussion of how these paths were generated.

Figure 13: Phase diagram for the parametric example under the hiring subsidy

Figure 14: Path to Σ^c in the parametric example

Figure 15: Path to Σ_h once the hiring subsidy policy is lifted

B.3 Effects of the investment voucher program in the simple dynamic model

The equilibria of this model are described by the same dynamic system as in Proposition 6. The relevant functions are just as in Appendix B.1. The only change resulting from the policy is that the $\dot{\Pi} = 0$ locus is updated to $\Pi_t = G(\Gamma_t^I + \min\{v, c_m\})$. In this simple model, $G(\cdot)$ is uniform so the policy shifts the $\dot{\Pi} = 0$ locus to the left by the value of the voucher.

Figure 16 depicts the updated phase diagram given that $v = 0.3$ and $c_m = 0.5$. All other parameter values are as in Section 5.1. The result is a dynamic system with only one steady state that is a saddle point. Consequently, all results from Section 4.1 hold for this parameterization.

The minimum voucher necessary to produce only one steady state, assuming that c_m is large enough, is $v^{\min} = (\bar{c} - \underline{c})\bar{\Pi} + \underline{c}$. This minimum voucher is decreasing in the precision of signals (P_u and P_q) as well as in α_c , and it is increasing in α_s . Note that if $v^{\min} > c_m$ then the policy will not be effective.

Figure 17 demonstrates the path that the disadvantaged group will take to the new steady state from their starting point at the low steady state. Figure 18 depicts the path that will be taken if the policy is removed after the group reaches Σ^v . See Appendix D.2 for a discussion of how these paths were generated.

Figure 16: Phase diagram for the parametric example under the investment voucher

Figure 17: Path to Σ^v for the parametric example

Figure 18: Path to Σ_h once the investment voucher policy is lifted

B.4 Effects of the equal treatment policy in the simple dynamic model

In the simple dynamic model under the equal treatment policy the optimal threshold rule for both groups will be

$$\hat{\theta}_t(\Pi_t^{pop}) = \begin{cases} 1 & \text{if } \Pi_t^{pop} \leq \underline{\Pi} \\ \theta_H & \text{if } \underline{\Pi} < \Pi_t^{pop} < \bar{\Pi} \\ 0 & \text{if } \Pi_t^{pop} \geq \bar{\Pi}, \end{cases} \quad [61]$$

where $\underline{\Pi}$ and $\bar{\Pi}$ are just as in the original parametric example. They are

$$\underline{\Pi} = \frac{\alpha_s(1 - P_u)}{\alpha_s(1 - P_u) + (\alpha_c - \alpha_s)P_q} \quad \text{and} \quad \bar{\Pi} = \frac{\alpha_s P_u}{\alpha_s P_u + (\alpha_c - \alpha_s)(1 - P_q)}. \quad [62]$$

Given the optimal task assignment rule in eq. [61], the unit period gain from investment is

$$\gamma(\Pi_t^{pop}) = \begin{cases} 0 & \text{if } \Pi_t^{pop} \leq \underline{\Pi} \\ (P_q + P_u - 1)(\alpha_c P(\theta_H, \Pi_t^{pop}) - \alpha_s) & \text{if } \underline{\Pi} < \Pi_t^{pop} < \bar{\Pi} \\ \alpha_c(P_q + P_u - 1)(P(\theta_H, \Pi_t^{pop}) - P(\theta_L, \Pi_t^{pop})) & \text{if } \Pi_t^{pop} \geq \bar{\Pi}. \end{cases} \quad [63]$$

The two loci of the dynamic system under the equal treatment policy are $\Gamma_t^{pop} = \frac{\gamma(\Pi_t^{pop})}{\beta+r}$, where $\gamma(\Pi_t^{pop})$ is as in eq. [63], and $\Pi_t^{pop} = G(\Gamma_t^{pop})$, where $G(\cdot)$ is the uniform distribution over $[\underline{c}, \bar{c}]$.

I can draw the same conclusions for the parametric example under the equal treatment policy as I did for the general dynamic model. If the original parameterization has three steady states, so will the parameterization under the equal treatment policy. Similarly, if Π_0^{pop} is in the certain region for Σ_h , then the policy will definitely eliminate discrimination. On the other hand, if Π_0^{pop} is in the uncertain region, then discrimination will be eliminated only if the population is optimistic.

Figure 19 shows the paths to the steady states for the same parameter values as in Section 5.1. I make the additional assumption that $\lambda^a = 0.75$ and $\lambda^d = 0.25$. From the analysis in Section 5.1 I know that the steady-state beliefs in this example are $\Pi_l = 0.1$, $\Pi_m = 0.47$, and $\Pi_h = 0.76$. So the resulting initial beliefs about the proportion of the population that is qualified are $\Pi_0^{pop} = 0.60$; this starting point is indicated in the figure by the purple dotted line. For this parameterization, the uncertain region of initial firm beliefs is $\Pi_t \in [0.32, 0.65]$.

Figure 19: Path to Σ_h given an optimistic population

It is clear that Π_0^{pop} is in this region. As such, the policy may or may not be successful depending on whether the population is optimistic or pessimistic. The purple dots in Figure 19 indicate the three points that the population may move to after the policy is put in place. Note that there are two points on the pessimistic path and one on the optimistic path.

Figure 20 shows the path that the population will take if it is optimistic. I also include the paths that the two groups will take from their individual starting points. See Appendix D.4 for a discussion of how these paths were generated. The disadvantaged group starts with an investment rate of Π_l because that is the steady state that they were previously operating in. The advantaged group starts with an investment rate of Π_h for the same reason. Figure 21 shows one of the paths that the population may take if it is pessimistic.

Figure 20: Path to Σ_h given an optimistic population

Just as in the general dynamic model under the equal treatment policy, if the proportion of the population that is in the disadvantaged group is very small, then the policy will be successful at eliminating the discrimination. For this particular set of parameter values if $\lambda^d < 0.17$, then the policy is guaranteed to eliminate the discrimination.

Figure 21: Path to Σ_l given a pessimistic population

C Calculating paths to the steady states

C.1 Paths to the steady states in the general dynamic model

Calculating the paths to the steady states is relatively straightforward. First, it is necessary to determine all the steady-state values by finding the solutions to the following system of equations:

$$\Pi^* = G(\Gamma^*), \quad [64]$$

$$\Gamma^* = \frac{\gamma(\Pi^*)}{\beta + r}. \quad [65]$$

Once $\Sigma_h(\Pi_h, \Gamma_h)$, $\Sigma_m(\Pi_m, \Gamma_m)$, and $\Sigma_l(\Pi_l, \Gamma_l)$ are known, I can determine the path to the high steady state by evaluating

$$\Pi_{t-1}^J = \frac{\Pi_t^J - \beta G(\Gamma_t^J)}{1 - \beta} \quad [66]$$

and

$$\Gamma_{t-1}^J = (1 - \beta - r)\Gamma_t^J + \gamma(\Pi_{t-1}^J) \quad [67]$$

at $\Pi_T = \Pi_h - \varepsilon$ and $\Gamma_T = \Gamma_h - \varepsilon$. This process yields Π_{T-1} and Γ_{T-1} , which can then be used to find Π_{T-2} and Γ_{T-2} . I repeat this process until the middle steady state is reached. The result is a series, $\{\Gamma_t, \Pi_t\}_{t=n}^T$, that maps out the path from the middle steady state to the high steady state.

I can determine the path to the low steady state by evaluating eqs [66] and [67] at $\Pi_T = \Pi_l + \varepsilon$ and $\Gamma_T = \Gamma_l + \varepsilon$. This process yields Π_{T-1} and Γ_{T-1} , which can then be used to find Π_{T-2} and Γ_{T-2} . I repeat this process until the middle steady state is reached. The result is a series, $\{\Gamma_t, \Pi_t\}_{t=n}^T$, that maps out the path from the middle steady state to the low steady state.

C.2 Paths to the steady states in the simple dynamic model

For the parametric example used in Section 5.3 the steady states are the solutions to the following system of equations:

$$\Pi^* = \Gamma^* + 0.10, \quad [68]$$

$$\Gamma^* = 4\gamma(\Pi^*), \quad [69]$$

where

$$\gamma(\Pi_t^J) = \begin{cases} 0 & \text{if } \Pi_t^J \leq 1/3 \\ \frac{1}{3} \left(\frac{3\Pi_t^J - 1}{\Pi_t^J + 1} \right) & \text{if } 1/3 < \Pi_t^J < 2/3 \\ \frac{2}{3} \left(\frac{2\Pi_t^J}{\Pi_t^J + 1} - \frac{\Pi_t^J}{2 - \Pi_t^J} \right) & \text{if } \Pi_t^J \geq 2/3. \end{cases} \quad [70]$$

Solving this yields the following three steady states: $\Sigma_h(\Gamma_h = 0.6631, \Pi_h = 0.7631)$, $\Sigma_m(\Gamma_m = 0.3687, \Pi_m = 0.4687)$, and $\Sigma_l(\Gamma_l = 0.0, \Pi_l = 0.1000)$.

In order to generate the optimistic path pictured in Figure 12, I induct backward from $\Pi_T = \Pi_h - 0.0001$ and $\Gamma_T = \Gamma_h - 0.0001$ using the following equations:

$$\Pi_{t-1}^J = 1.25\Pi_t^J - 0.25\Gamma_t^J - 0.025, \quad [71]$$

$$\Gamma_{t-1}^J = .75\Gamma_t^J + \gamma(\Pi_{t-1}^J). \quad [72]$$

The resulting series $\{\Gamma_t, \Pi_t\}_{t=n}^T$ describes the optimistic path. Similarly, I generate the pessimistic path by inducting backward from $\Pi_T = \Pi_l + 0.0001$ and $\Gamma_T = \Gamma_l + 0.0001$. The resulting series, $\{\Gamma_t, \Pi_t\}_{t=n}^T$, describes the pessimistic path.

D Welfare calculations

D.1 Welfare in the base model

In order to understand the welfare effects of the three policies, it is necessary to determine the average expected welfare in both the high and low steady states of the base model for the parameterization described in Section 5.3. The general equation for the average expected welfare in one period is

$$\begin{aligned}
 W(\Pi_t^I) = & \underbrace{\Pi_t^I}_{\text{Proportion of qualified workers}} \left[\underbrace{\int_0^1 \omega_\Delta(\theta, \Pi_t^I) f_q(\theta) d\theta}_{\text{Expected wage of qualified workers}} - \underbrace{\beta}_{\text{Fraction of new births}} \underbrace{\left(\frac{\int_c^{\Gamma(\Pi_t^I)} c g(c) dc}{G(\Gamma(\Pi_t^I))} \right)}_{\text{Expected cost given investment}} \right] \\
 & + \underbrace{(1 - \Pi_t^I)}_{\text{Proportion of unqualified worker}} \underbrace{\int_0^1 \omega_\Delta(\theta, \Pi_t^I) f_u(\theta) d\theta}_{\text{Expected wage of unqualified workers}}.
 \end{aligned} \quad [73]$$

For the parameterization used in Section 5.3 the average expected welfare can be simplified to

$$W(\Pi_t^I) = \Pi_t^I \left[E_t[\omega_\Delta|q] - \beta \left(\frac{\Gamma(\Pi_t^I) - 0.10}{2} \right) \right] + (1 - \Pi_t^I) E_t[\omega_\Delta|u], \quad [74]$$

where

$$E_t[\omega_\Delta|u] = \begin{cases} 1 & \text{if } \Pi_t^I \leq 1/3 \\ \frac{2}{3} \left(\frac{3\Pi_t^I + 1}{\Pi_t^I + 1} \right) & \text{if } 1/3 < \Pi_t^I < 2/3 \\ \frac{2}{3} \left(\frac{2\Pi_t^I}{2 - \Pi_t^I} + \frac{2\Pi_t^I}{1 + \Pi_t^I} \right) & \text{if } \Pi_t^I \geq 2/3 \end{cases} \quad [75]$$

and

$$E_t[\omega_\Delta|q] = \begin{cases} 1 & \text{if } \Pi_t^I \leq 1/3 \\ \frac{1}{3} \left(\frac{9\Pi_t^I + 1}{\Pi_t^I + 1} \right) & \text{if } 1/3 < \Pi_t^I < 2/3 \\ \frac{2}{3} \left(\frac{\Pi_t^I}{2 - \Pi_t^I} + \frac{4\Pi_t^I}{1 + \Pi_t^I} \right) & \text{if } \Pi_t^I \geq 2/3. \end{cases} \quad [76]$$

Evaluating eq. [74] at the high and low steady states provides the baseline welfare level for workers in the advantaged and disadvantaged groups, respectively. In this case, $W(\Pi_h) = 1.4833$ and $W(\Pi_l) = 1.0010$.

D.2 Welfare under a hiring subsidy

The first step to determining the welfare effects of the hiring subsidy policy is to calculate the path that the disadvantaged group will take when the policy is put in place as well as the path that they will take to the high steady state when the policy is lifted. Once these paths are established, I can calculate the welfare of each group for each period of interest.

I begin by determining the single steady state of the dynamic system under the policy. This steady state is the solution to the following system of equations:

$$\Pi^\xi = G(\Gamma^\xi), \quad [77]$$

$$\Gamma^\xi = \frac{\gamma^\xi(\Pi^\xi)}{\beta + r}. \quad [78]$$

I refer to the single steady state as $\Sigma^\xi(\Gamma^\xi, \Pi^\xi)$. Once this is determined I find the disadvantaged group's path to the steady state by inducting backward from $\Gamma_T = \Gamma^\xi - \varepsilon$ and $\Pi_T = \Pi^\xi - \varepsilon$ using the following equations:

$$\Pi_{t-1}^d = \frac{\Pi_t^d - \beta G(\Gamma_t^d)}{1 - \beta}, \quad [79]$$

$$\Gamma_{t-1}^d = (1 - \beta - r)\Gamma_t^d + \gamma^\xi(\Pi_{t-1}^d). \quad [80]$$

To determine the path taken after the policy is removed I use eqs [66] and [67] and induct backward from $\Pi_T = \Pi_h + \varepsilon$ and $\Gamma_T = \Gamma_h + \varepsilon$.

For the parameterization of interest the steady state is $\Sigma^\xi(\Gamma^\xi = 0.8159, \Pi^\xi = 0.9159)$. This is determined by finding the solution to the following system of equations:

$$\Pi^\xi = \Gamma^\xi + 0.10, \quad [81]$$

$$\Gamma^\xi = 4\gamma^\xi(\Pi^\xi), \quad [82]$$

where

$$\gamma^\xi(\Pi_t^I) = \begin{cases} 0 & \text{if } \Pi_{t-1}^I \leq 0.10 \\ \frac{1}{3} \left(\frac{10\Pi_t^I - 1}{\Pi_t^I + 1} \right) & \text{if } 0.10 < \Pi_t^I < 0.31 \\ \frac{11}{6} \left(\frac{2\Pi_t^I}{\Pi_t^I + 1} - \frac{\Pi_t^I}{2 - \Pi_t^I} \right) & \text{if } \Pi_t^I \geq 0.31. \end{cases} \quad [83]$$

In order to find the path that the disadvantaged group will take to this steady state, I induct backward from $\Gamma_T = \Gamma^\xi - \varepsilon$ and $\Pi_T = \Pi^\xi - \varepsilon$ using the following equations:

$$\Pi_{t-1}^d = 1.25\Pi_t^d - 0.25\Gamma_t^d - 0.025, \quad [84]$$

$$\Gamma_{t-1}^d = .75\Gamma_t^d + \gamma^\xi(\Pi_{t-1}^d), \quad [85]$$

where $\varepsilon = 0.00000001$.

I use eqs [71] and [72] and induct backward from $\Pi_T = \Pi_h + \varepsilon$ and $\Gamma_T = \Gamma_h + \varepsilon$ to find the path taken after the policy is removed.

To determine the welfare effects while the subsidy is in place, I look at the Γ_t, Π_t pairs on the path to Σ^ξ starting from Π_t , the disadvantaged group starting point, through the steady-state values under the policy. To find how welfare changes after the policy is lifted I look at the Γ_t, Π_t pairs starting from Π^ξ until the high steady state is reached. With these values I am able to calculate each group's welfare during each period that the policy is in place as well as during the periods after it is removed.

The policy is applied in a general equilibrium framework so that the workers in the advantaged group are taxed in order to fund the policy. As such, the welfare of both groups will be impacted by the policy, even though the behavior of the advantaged group is not affected. The welfare of the advantaged group while the policy is in place is the following

$$W^a(\Pi_t^a, \Pi_t^d) = \underbrace{W(\Pi_h)}_{\text{Welfare in the high steady state}} - \underbrace{\xi}_{\text{Value of the subsidy}} \underbrace{\frac{\lambda^d}{\lambda^a}}_{\text{Fraction of the subsidy that each worker provides}} \underbrace{\Pi_t^d}_{\text{Proportion of qualified workers}} \underbrace{\int_{\hat{\theta}(\Pi_t^d)}^1 f_q(\theta) d\theta}_{\text{Probability a qualified worker is assigned to C}} \quad [86]$$

Per worker transfer in each period

This is simply the average welfare in the high steady state minus the average expected transfer. Note that the transfer each worker in the advantaged group

provides depends on the relative sizes of the two groups. Clearly, if more of the population is disadvantaged, then the loss in welfare to those in the advantaged group is higher.

The average expected welfare of the disadvantaged group while the policy is in place is the following

$$W^d(\Pi_t^d) = \Pi_t^d \left[\underbrace{\int_0^1 \omega_\Delta^\xi(\theta, \Pi_t^d) f_q(\theta) d\theta}_{\text{Expected wage of qualified workers under the subsidy policy}} - \beta \left(\frac{\int_{\mathcal{C}}^{\Gamma(\Pi_t^d)} c g(c) dc}{G(\Gamma(\Pi_t^d))} \right) \right] + (1 - \Pi_t^d) \underbrace{\int_0^1 \omega_\Delta^\xi(\theta, \Pi_t^d) f_u(\theta) d\theta}_{\text{Expected wage of unqualified workers under the subsidy policy}}. \quad [87]$$

The only way that this differs from eq. [73] is that it depends on the expected wage under the subsidy policy, $\omega_\Delta^\xi(\theta, \Pi_t^d)$; all other components are the same.

For the parameterization, the advantaged group's average expected welfare while the policy is in place is

$$W^a(\Pi_t^a, \Pi_t^d) = 1.4833 - \frac{7\Pi_t^d}{6} P[\theta_t > \hat{\theta}(\Pi_t^d) | q], \quad [88]$$

where

$$P[\theta_t > \hat{\theta}(\Pi_t^d) | q] = \begin{cases} 0 & \text{if } \Pi_t^d \leq 0.10 \\ \frac{2}{3} & \text{if } 0.10 < \Pi_t^d < 0.31 \\ 1 & \text{if } \Pi_t^d \geq 0.31. \end{cases} \quad [89]$$

The average expected welfare for the disadvantaged group is

$$W_t^d(\Pi_t^d) = \Pi_t^d \left(E_t[\omega_\Delta^\xi | q] - \beta \left(\frac{\Gamma(\Pi_t^d) - 0.10}{2} \right) \right) + (1 - \Pi_t^d) E_t[\omega_\Delta^\xi | u], \quad [90]$$

where

$$E_t[\omega_\Delta^\xi | u] = \begin{cases} 1 & \text{if } \Pi_t^d \leq 0.10 \\ \frac{1}{3} \left(\frac{13\Pi_t^d + 2}{\Pi_t^d + 1} \right) & \text{if } 0.10 < \Pi_t^d < 0.31 \\ \frac{11}{3} \left(\frac{\Pi_t^d}{2 - \Pi_t^d} + \frac{\Pi_t^d}{1 + \Pi_t^d} \right) & \text{if } \Pi_t^d \geq 0.31 \end{cases} \quad [91]$$

and

$$E_t[\omega_\Delta^\xi | q] = \begin{cases} 1 & \text{if } \Pi_t^d \leq 0.10 \\ \frac{1}{3} \left(\frac{23\Pi_t^d + 1}{\Pi_t^d + 1} \right) & \text{if } 0.10 < \Pi_t^d < 0.31 \\ \frac{11}{6} \left(\frac{\Pi_t^d}{1 - \Pi_t^d} + \frac{4\Pi_t^d}{1 + \Pi_t^d} \right) & \text{if } \Pi_t^d \geq 0.31. \end{cases} \quad [92]$$

To generate Figure 8 I evaluate eqs [88] and [90] at all Γ_t, Π_t pairs along the path starting from $\Pi_l = 0.10$ till the group is within 0.0001 of the single steady state of the dynamic system under the hiring subsidy. To calculate the welfare after the policy is removed I evaluate eq. [74] at all Γ_t, Π_t pairs along the path from Π^ξ till the group is within 0.0001 of the high steady state.

D.3 Welfare under an investment voucher

In order to find the welfare effects of the investment voucher policy, I must determine the path that the disadvantaged group will take once the policy is put in place as well as the path that they will take to the high steady state once the policy is lifted. Once these paths are established, I can calculate the welfare of each group for each period that the policy is in place.

The first step to calculating the path that the disadvantaged group will take once the investment voucher policy is implemented is to determine the single steady state of the dynamic system under the policy. This steady state is the solution to the following system of equations:

$$\Pi^v = G(\Gamma^v + v), \quad [93]$$

$$\Gamma^v = \frac{\gamma(\Pi^v)}{\beta + r}. \quad [94]$$

I refer to it as $\Sigma^v(\Gamma^v, \Pi^v)$. Once this is determined, I can find the disadvantaged group's path to the steady state by inducting backward from $\Gamma_T = \Gamma^v - \varepsilon$ and $\Pi_T = \Pi^v - \varepsilon$ using the following equations:

$$\Pi_{t-1}^d = \frac{\Pi_t^d - \beta G(\Gamma_t^d + v)}{1 - \beta}, \quad [95]$$

$$\Gamma_{t-1}^d = (1 - \beta - r)\Gamma_t^d + \gamma(\Pi_{t-1}^d). \quad [96]$$

I then use eqs [66] and [67] and induct backward from $\Pi_T = \Pi_h + \varepsilon$ and $\Gamma_T = \Gamma_h + \varepsilon$ to find the path taken after the policy is removed.

For the parameterization, the steady state is Σ^v ($\Gamma^v = 0.5058, \Pi^v = 0.8391$). This is determined by finding the solution to the following system of equations:

$$\Pi^v = \Gamma^v + \frac{1}{3}, \quad [97]$$

$$\Gamma^v = 4\gamma(\Pi^v), \quad [98]$$

where $\gamma(\cdot)$ is as in eq. [70].

In order to find the path that the disadvantaged group will take to this steady state, I induct backward from $\Gamma_T = \Gamma^v - \varepsilon$ and $\Pi_T = \Pi^v - \varepsilon$ using the following equations:

$$\Pi_{t-1}^d = 1.25\Pi_t^d - 0.25\Gamma_t^d - 0.083\bar{3}, \quad [99]$$

$$\Gamma_{t-1}^d = .75\Gamma_t^d + \gamma(\Pi_{t-1}^d), \quad [100]$$

where $\varepsilon = 0.00000001$.

To determine the path taken after the policy is removed I use eqs [71] and [72] and induct backward from $\Pi_T = \Pi_h + \varepsilon$ and $\Gamma_T = \Gamma_h + \varepsilon$.

I can calculate the welfare effects while the voucher is in place by looking at all the Γ_t, Π_t pairs along the path to Σ^v starting from Π_l , the disadvantaged group starting point. To find how welfare changes after the policy is lifted, I look at all the Γ_t, Π_t pairs along the path to Σ_h starting from the Π^v . With these values, I can calculate each group's welfare during each period that the policy is in place as well as during the periods after it is removed.

The policy is applied in a general equilibrium framework so that the workers in the advantaged group are taxed in order to fund the policy, as such the welfare of both groups will be impacted by the policy, even though the behavior of the advantaged group is not affected.

The welfare of the advantaged group while the policy is in place is

$$W^a(\Pi_t^a, \Pi_t^d) = \underbrace{W(\Pi_h)}_{\text{Welfare in the high steady state}} - \underbrace{\frac{v}{\lambda^{\frac{d}{a}}}}_{\substack{\text{Value of the} \\ \text{voucher}}} \underbrace{\frac{\beta}{\lambda^{\frac{d}{a}}}}_{\substack{\text{Fraction of the} \\ \text{voucher that each} \\ \text{worker provides}}} \underbrace{\frac{G(\Gamma(\Pi_t^d) + v)}{\lambda^{\frac{d}{a}}}}_{\substack{\text{Fraction of} \\ \text{new births} \\ \text{Proportion of} \\ \text{workers that invest} \\ \text{under the voucher}}} \\ \text{Per worker transfer in each period.} \quad [101]$$

Note that the transfer that each worker in the advantaged group provides depends on the relative sizes of the two groups. Clearly if more of the population is disadvantaged, then the loss in welfare to those in the advantaged group is higher.

The welfare of the disadvantaged group while the policy is in place is

$$W^d(\Pi_t^d) = \Pi_t^d \left[\int_0^1 \omega_\Delta(\theta, \Pi_t^d) f_q(\theta) d\theta - \beta \left(\underbrace{\frac{\int_c^{\Gamma(\Pi_t^d)+v} c g(c) dc}{G(\Gamma(\Pi_t^d) + v)}}_{\text{Expected cost given investment under the voucher policy}} - \underbrace{v}_{\text{Value of the voucher}} \right) \right] + (1 - \Pi_t^d) \int_0^1 \omega_\Delta(\theta, \Pi_t^d) f_u(\theta) d\theta. \quad [102]$$

The only way that this differs from eq. [73] is that workers are more likely to invest (as a result of the voucher) and the costs associated with this investment are reduced by v .

For the parameterization, the average expected welfare for the advantaged group while the policy is in place is

$$W^a(\Pi_t^a, \Pi_t^d) = 1.4833 - 0.0156(\Gamma(\Pi_t^d) + 0.1333). \quad [103]$$

The average expected welfare for the disadvantaged group while the policy is in place is

$$W^d(\Pi_t^d) = \Pi_t^d \left(E_t[\omega_\Delta|q] - \beta \left(\frac{\Gamma(\Pi_t^d) - 0.33}{2} \right) \right) + (1 - \Pi_t^d) E_t[\omega_\Delta|u], \quad [104]$$

where $E_t[\omega_\Delta|u]$ and $E_t[\omega_\Delta|q]$ are as in eqs [75] and [76], respectively.

To generate Figure 9, I evaluate eqs [103] and [104] at all Γ_t, Π_t pairs along the path to Σ^v starting from Π_l , the disadvantaged groups starting point, till they are within 0.0001 of Σ^v . To calculate the welfare after the policy is removed I evaluate eq. [74] at all Γ_t, Π_t pairs along the path to Σ_h starting from Π^v till they are within 0.0001 of the high steady state.

D.4 Welfare under the equal treatment policy

Finding the welfare effects of the equal treatment policy is slightly more involved than the process for the previous policies. Two steps are necessary.

First, I must find the path that the entire population takes to the high steady state once the policy is put in place. Using this path I can then find the individual paths of the two groups. It is important to consider the fact that the population as a whole will reach the steady state before the two groups do, as such I must account for several periods where the population investment rate is not changing but the group investment rates still are. Once the two groups reach the steady state, then the policy can be safely removed.

The dynamic system is the same under the equal treatment policy as it is without the policy, but it depends on the firms' beliefs about the average investment rate of the entire population. As such the first step is to calculate the path that the population as a whole will take once the policy is put in place. I assume that the population is optimistic so that they will take the path to the high steady state. I determine the population's path to the high steady state by inducing backward from $\Gamma_T^{pop} = \Gamma_h - \varepsilon$ and $\Pi_T^{pop} = \Pi_h - \varepsilon$ using the following equations:

$$\Pi_{t-1}^{pop} = \frac{\Pi_t^{pop} - \beta G(\Gamma_t^{pop})}{1 - \beta}, \quad [105]$$

$$\Gamma_{t-1}^{pop} = (1 - \beta - r)\Gamma_t^{pop} + \gamma(\Pi_{t-1}^{pop}). \quad [106]$$

After this process I have all the Γ_t^{pop} , Π_t^{pop} pairs that form the path to the high steady state. I am only interested in the path starting at $\Pi^{pop} = \lambda^a \Pi_h + \lambda^d \Pi_l$. Once I calculate the population path, I can find the paths that the individual groups will take by forward inducing from $\Pi_0^d = \Pi_l$ and $\Pi_0^a = \Pi_h$ using the following equations:

$$\Pi_t^d = (1 - \beta)\Pi_{t-1}^d + \beta G(\Gamma_t^{pop}), \quad [107]$$

$$\Pi_t^a = (1 - \beta)\Pi_{t-1}^a + \beta G(\Gamma_t^{pop}). \quad [108]$$

Note that I know the relevant Γ_t^{pop} values from the previous calculation of the population's path of the high steady state.

For the parameterization of interest the high steady state is $\Sigma_h(\Gamma_h = 0.6631, \Pi_h = 0.7631)$. In order to find the path that the population will take to this steady state I induct backward from $\Gamma_T^{pop} = \Gamma_h - \varepsilon$ and $\Pi_T^{pop} = \Pi_h - \varepsilon$ where $\varepsilon = 0.00000001$ using the following equations:

$$\Pi_{t-1}^{pop} = 1.25\Pi_t^{pop} - 0.25\Gamma_t^{pop} - 0.025, \quad [109]$$

$$\Gamma_{t-1}^{pop} = .75\Gamma_t^{pop} + \gamma^{pop}(\lambda^a\Pi_{t-1}^a + \lambda^d\Pi_{t-1}^d). \quad [110]$$

After this process I know all the Γ_t^{pop} , Π_t^{pop} pairs that form the path to the high steady state. I am only interested in the path starting at $\Pi^{pop} = 0.5973$. With this path I am able to determine the paths that the two groups will take by forward inducting from $\Pi_0^d = 0.1000$ and $\Pi_0^a = 0.7631$ using the following equations:

$$\Pi_t^d = 0.80\Pi_{t-1}^d + 0.20\Gamma_t^{pop} + 0.02, \quad [111]$$

$$\Pi_t^a = 0.80\Pi_{t-1}^a + 0.20\Gamma_t^{pop} + 0.02. \quad [112]$$

In this parameterization it is also necessary to include ten periods where Π^{pop} is at the steady-state level but the individual groups are still approaching it.

In order to determine the average expected welfare of the disadvantaged group while the equal treatment policy is in place, I look at both the Γ_t^{pop} , Π_t^{pop} pairs that form the population path and the Γ_t^{pop} , Π_t^d pairs that form the disadvantaged group's path starting from $\Pi^d = \Pi_l$ and $\Pi^{pop} = \lambda^a\Pi_h + \lambda^d\Pi_l$ until Σ^{pop} is reached. For the advantaged group I consider both the Γ_t^{pop} , Π_t^{pop} pairs that form the population path and the Γ_t^{pop} , Π_t^a pairs that form the advantaged group's path starting from $\Pi^a = \Pi_h$ and $\Pi^{pop} = \lambda^a\Pi_h + \lambda^d\Pi_l$ until Σ^{pop} is reached. With these values I can then calculate each group's welfare during each period that the policy is in place.

The welfare of each group while the policy is in place is

$$W^J(\Pi_t^J, \Pi_t^{pop}) = \underbrace{\Pi_t^J}_{\text{Proportion of qualified workers}} \left[\underbrace{\int_0^1 \omega_\Delta(\theta, \Pi_t^{pop}) f_q(\theta) d\theta}_{\text{Expected wage of qualified workers}} - \underbrace{\beta}_{\text{Fraction of new births}} \underbrace{\left(\frac{\int_c^{\Gamma(\Pi_t^{pop})} cg(c) dc}{G(\Gamma(\Pi_t^{pop}))} \right)}_{\text{Expected cost given investment}} \right] \\ + \underbrace{(1 - \Pi_t^J)}_{\text{Proportion of unqualified worker}} \underbrace{\int_0^1 \omega_\Delta(\theta, \Pi_t^{pop}) f_u(\theta) d\theta}_{\text{Expected wage of unqualified workers}} \text{ for } J = a, d. \quad [113]$$

For the parameterization the average expected welfare of each group while the policy is in place is

$$W^J(\Pi_t^J, \Pi_t^{pop}) = \Pi_t^J \left(E_t[\omega_\Delta^{pop}|q] - \beta \left(\frac{\Gamma(\Pi_t^{pop}) - 0.10}{2} \right) \right) + (1 - \Pi_t^J) E_t[\omega_\Delta^{pop}|u],$$

[114]

where $E_t[\omega_\Delta^{pop}|u]$ and $E_t[\omega_\Delta^{pop}|q]$ are as in eqs [75] and [76], respectively. They are simply evaluated at Π^{pop} rather than Π^J .

To generate Figure 10, I evaluate eq. 114 at all relevant pairs along the paths from the starting point $\Pi^{pop} = 0.5973$, $\Pi^d = 0.1000$, $\Pi^a = 0.7631$ till both groups are within 0.0001 of the high steady state.

D.5 Welfare results for a sampling of parameterizations

All these results assume starting parameter values of $P_q = 2/3$, $P_u = 2/3$, $\alpha_s = 1$, $\alpha_c = 2$, $\beta = 0.2$, $r = 0.05$, $\underline{c} = -0.1$, $\bar{c} = 0.9$, $\lambda^d = .25$, and $\lambda^a = 0.75$, only those parameters mentioned are varied from this point.

Table 3: Welfare effects of varying λ^d

	Parameter value	Group <i>d</i> welfare gain	Group <i>a</i> welfare loss	Net change in pop. welfare	Policy length
Hiring subsidy	$\lambda^d = 0.05$	61.908	3.124	0.128	20
	$\lambda^d = 0.25$	61.908	17.855	2.086	20
	$\lambda^d = 0.67$	61.908	106.550	6.317	20
Investment voucher	$\lambda^d = 0.05$	12.251	0.053	0.563	25
	$\lambda^d = 0.25$	12.251	0.329	2.816	25
	$\lambda^d = 0.67$	12.251	2.028	7.539	25
Equal treatment	$\lambda^d = 0.05$	19.460	0.189	0.794	41
	$\lambda^d = 0.25$	18.495	0.887	3.959	43
	$\lambda^d = 0.67$	21.817	2.622	13.752	52

Table 4: Welfare effects of varying P_u and P_q

	Parameter value	Group <i>d</i> welfare gain	Group <i>a</i> welfare loss	Net change in pop. welfare	Policy length
Hiring subsidy	$P_u = 0.8$	49.516	13.299	2.404	21
	$P_q = 0.5$				
	$P_u = 2/3$	61.908	17.855	2.086	20
	$P_q = 2/3$				
	$P_u = 0.6$	65.543	18.777	2.303	19
	$P_q = 0.75$				
	$P_u = 0.8$	11.508	0.266	2.678	27
Investment voucher	$P_q = 0.5$				
	$P_u = 2/3$	12.251	0.329	2.816	25
	$P_q = 2/3$				
	$P_u = 0.6$	14.331	0.370	3.306	25
	$P_q = 0.75$				
Equal treatment	$P_u = 0.8$	17.238	0.859	3.665	48
	$P_q = 0.5$				
	$P_u = 2/3$	18.495	0.887	3.959	43
	$P_q = 2/3$				
	$P_u = 0.6$	21.129	1.558	4.114	41
	$P_q = 0.75$				

Table 5: Welfare effects of varying α_s and α_c

	Parameter	Group <i>d</i> welfare n	Group <i>a</i> welfare loss	Net change in pop. welfare	Policy length
Hiring subsidy	$\alpha_s = 0.95$	63.427	17.634	2.63	21
	$\alpha_c = 2$				
	$\alpha_s = 1$	61.908	17.855	2.086	20
	$\alpha_c = 2$				
	$\alpha_s = 1$	59.647	16.434	2.587	19
	$\alpha_c = 2.1$				
	$\alpha_s = 0.95$	11.883	0.289	2.754	24
	$\alpha_c = 2$				
Investment voucher	$\alpha_s = 1$	12.251	0.329	2.816	25
	$\alpha_c = 2$				
	$\alpha_s = 1$	13.937	0.294	3.264	24
	$\alpha_c = 2.1$				
	$\alpha_s = 0.95$	21.237	1.540	4.154	48
	$\alpha_c = 2$				
Equal treatment	$\alpha_s = 1$	18.495	0.887	3.959	43
	$\alpha_c = 2$				
	$\alpha_s = 1$	25.274	1.026	5.549	46
	$\alpha_c = 2.1$				

Table 6: Welfare effects of varying β and r

	Parameter	Group d welfare gain	Group a welfare loss	Net change in pop. welfare	Policy length
Hiring subsidy	$\beta = 0.1$	87.994	29.918	-0.440	35
	$r = 0.1$				
	$\beta = 0.2$	61.908	17.855	2.086	20
	$r = 0.05$				
	$\beta = 0.05$	195.706	54.867	7.776	63
	$r = 0.2$				
Investment voucher	$\beta = 0.1$	26.318	0.322	6.331	48
	$r = 0.1$				
	$\beta = 0.2$	12.251	0.329	2.816	25
	$r = 0.05$				
	$\beta = 0.05$	39.153	0.321	9.547	126
	$r = 0.2$				
Equal treatment	$\beta = 0.1$	51.824	-0.724	13.499	91
	$r = 0.1$				
	$\beta = 0.2$	19.495	0.887	3.959	43
	$r = 0.05$				
	$\beta = 0.05$	90.728	-3.183	25.069	186
	$r = 0.2$				

Table 7: Welfare effects of varying $\underline{c}$ and $\bar{c}$

	Parameter	Group d welfare gain	Group a welfare loss	Net change in pop. welfare	Policy length
Hiring subsidy	$\underline{c} = -0.01$	706.875	233.596	1.522	23
	$\bar{c} = 0.81$				
	$\underline{c} = -0.1$	61.908	17.855	2.086	20
	$\bar{c} = 0.9$				
	$\underline{c} = -0.2$	31.980	7.600	2.295	21
	$\bar{c} = 1$				
Investment voucher	$\underline{c} = -0.01$	13.896	0.383	3.187	25
	$\bar{c} = 0.81$				
	$\underline{c} = -0.1$	12.251	0.329	2.816	25
	$\bar{c} = 0.9$				
	$\underline{c} = -0.2$	11.469	0.278	2.658	26
	$\bar{c} = 1$				
Equal treatment	$\underline{c} = -0.01$	22.847	0.902	5.035	36
	$\bar{c} = 0.81$				
	$\underline{c} = -0.1$	18.495	0.887	3.959	43
	$\bar{c} = 0.9$				
	$\underline{c} = -0.2$	17.733	0.784	3.845	42
	$\bar{c} = 1$				

References

- Arrow, K. 1973. "The Theory of Discrimination." In *Discrimination in Labor Markets*, edited by O. Ashenfelter and A. Rees, 3–33. Princeton, NJ: Princeton University Press.
- Blume, L. 2006. "The Dynamics of Statistical Discrimination." *Economic Journal* 116:F480–F498.
- Coate, S., and G. Loury. 1993. "Will Affirmative-Action Policies Eliminate Negative Stereotype?" *American Economic Review* 83:1220–40.
- Fryer, R. 2007. "Belief Flipping in a Dynamic Model of Statistical Discrimination." *Journal of Public Economics* 91:1151–66.
- Glawtschew, R. 2013. *Persistent Statistical Discrimination in a Competitive Economy*. Chapel Hill, NC: The University of North Carolina.
- Kim, Y., and G. Loury. 2014. "Collective Reputation and the Dynamics of Statistical Discrimination." Unpublished Manuscript.
- Krugman, P. 1991. "History Versus Expectations." *The Quarterly Journal of Economics* 106:651–67.
- Moro, A., and P. Norman. 2002. "A General Equilibrium Model of Statistical Discrimination: Omitted Calculations and Proofs." Unpublished Manuscript.
- Moro, A., and P. Norman. 2003. "A General Equilibrium Model of Statistical Discrimination." *Journal of Economic Theory* 114:1–30.