We thank the reviewers for their thoughtful comments. Regarding the empirical evaluation of our policy, we plan to include simulations to showcase the efficiency of the policy; due to lack of time we leave this for the final version of the paper. Next, we address each reviewer's additional questions separately.

Review 1: — Regarding non-interchangeable occupations and jobs: Thank you for raising your concern; this is an interesting aspect. We currently assume that there is one fixed job type and the same distribution of high-skilled workers across social groups. With multiple job types and belief-based bias, we could again improve the welfare of minority. Specifically, in order to learn whether an employer is discriminating (see eq. (10)), we need to condition on the job type r and the bias level β_r of discriminating employers against minority workers doing job r, as well as take into account the fraction of D employers looking for task r. Thus, we think that the DM policy can be modified to apply here, with the difference that matching decisions will also depend on job types and capacity constraints. Our results will still hold even if workers stick to their type of jobs as long as workers do accept offers from many different employers of the same type (who offer the same job). This is common in many types of jobs in online labor platforms.

— Regarding the concern about reinforcing unjustified stereotypes: Given belief-based bias, our policy reduces the effect of unjustified stereotypes. It achieves it by helping minority workers accumulate a larger number of reviews. As the available information about the worker increases, the effect of stereotypes (belief-based bias) reduces. Theoretically, bias can not be reinforced asymptotically (Theorem 2). But even in a practical setting with a finite time horizon, our policy helps decrease faster the uncertainty about minority workers; as a result, the discrimination gap also decreases.

— Thank you for the references. Indeed, studying the exit rate of workers is an interesting direction. Intuitively, we expect higher exit rates of minority workers because "they may not even have the chance to receive enough reviews or even stay long in the platform due to the competition" (l. 241-243). Hence our DM policy may help to reduce exit.

— Regarding the applicability of DM policy: There is no available law that clearly regulates such policies (see Rosenblat et al. (2017), Levy and Barocas (2017)). However, platforms are already allowed to collect data about users to optimize the platform's actions such as matching and recommendations. Policies such as hiding sensitive information (see the Airbnb policy about user photos) have not been successful. Furthermore, policies that penalize discriminating users may be problematic because 1) under belief-based bias, it is not clear how to select a threshold for discriminatory behavior, and 2) such policies may create imbalance in the market. Thus, our policy could be an effective, easy alternative to implement and/or serve as a benchmark for future policies. Nevertheless, online platforms already implement similar directed matching (e.g. new users with few reviews are more frequently shown on top of search results).

Review 2: — In comparison to Johari et al. (2017), we include agent histories on both sides of the market and incorporate strategic behaviors (social learning and hiring/review decisions) on one side of the market (employers) to the evolution of the system (lines 83-84). These two factors make the dynamical system in our paper take a non-linear, non-standard form, and differentiate (both technically and conceptually) our model from Johari et al. (2017).

— Regarding bias in reviews: We could directly extend the model to add an additional bias level for reviews. In this case, the model exhibits taste-based discrimination (and not belief-based as we mainly consider). As we discuss in Appendix E, given taste-based bias, discrimination persists asymptotically but our DM policy can still improve the welfare of minority workers. The same would hold if we assume additional bias in reviews. In both cases of belief- and taste-based bias, DM helps minority workers accumulate a larger number of reviews faster.

— Regarding comparison of multiple workers: Modeling this will complicate the model by introducing choice models. However, if we borrow ideas from search theory, then the problem of each employer reduces to a threshold-based decision rule which - in expectation - should not affect our current results.

— On notation in eq. (3): Indeed, we mean $W_q^c(K) = \mathbb{E}(\sum_{k=1}^{K} \delta^k m_k \mid Q = q, C = c)$. Thank you for the correction.

— On use of "monotonic": We actually mean non-strictly increasing/decreasing.

— On word choice (minority/majority): Our results do not rely on assumptions about the size of each worker group. Thus, priviledged/unpriviledged is a better choice and we will adopt it in the paper.

Review 4: — Thank you for acknowledging the originality of our model. We also view this as a significant contribution.

— $\mathbb{P}(g = D \mid H_n)$ is the probability that the employer belongs to group D (discriminating) given her history H_n of past hiring decisions about the $n-1$ workers she has met so far.

— "Ex-ante/ ex-post idiosyncratic": We mean the preference shocks of employers before/ after hiring the worker, respectively (see also Besbes and Scarsini 2018). We will include the definition.

— On the optimality of DM policy: The optimality of DM policy is an open, challenging question. However, the policy is simple and we show that it successfully reduces discrimination.

— On Section 4.2 and the learning pool: Some workers are in the learning pool, because we are learning from employers' decisions made for those workers. Under the given DM policy, some minority workers enter the learning pool and a few may remain there until they leave. An alternative solution (which does not affect our technical analysis) would be to carefully randomize among workers of the same history so that minority workers in the learning pool also exit the learning pool with positive probability.