Attentional Role of Quota Implementation*

Andrei Matveenko[1,2] and Sergei Mikhalishchev[†2]

[1]University of Copenhagen[‡]
[2]CERGE-EI[§]

Abstract

In this paper, we introduce a new role of quotas, e.g. labor market quotas: the attentional role. We study the effect of quota implementation on the attention allocation strategy of a rationally inattentive (RI) manager. We find that quotas induce attention: a RI manager who is forced to fulfill a quota, unlike an unrestricted RI manager, never rejects minority candidates without acquiring information about them. We also demonstrate that in our model quotas are behaviorally equivalent to subsidies. In addition, we analyze different goals that the social planner can achieve by implementing quota. First, quotas can eliminate statistical discrimination, i.e. make the chances of being hired independent from the group identity. Second, when the hiring manager has inaccurate belief about distribution of candidate's producitvities, the social planner can make the manager to behave as if she had correct beliefs. Finally, we show how our results can be used to set a quota level that increases the expected value of the chosen candidates and induces truthful type revelation by candidates.

Keywords: discrete choice, rational inattention, multinomial logit, quotas
JEL classification codes: D63, D81, D83, H23, J08

*This project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreements No. 678081 and No. 740369) and was supported by Charles University, GAUK project 36119. We are grateful to Hassan Afrouzi, Mark Dean, Randall Filer, Mogens Fosgerau, Dino Gerardi, Daniel Martin, Nikolaj Nielsen, Debraj Ray, Sandro Shelegia, Peter Norman Sørensen, Jakub Steiner and Jan Zápal for helpful comments and suggestions. We thank Filip Matějka for excellent supervision and patient guidance. All remaining errors are our own.

†Corresponding author. E-mail address: sergei.mikhalishchev@cerge-ei.cz.

‡Department of Economics, University of Copenhagen, Øster Farimagsgade 5, 1307 Copenhagen, Denmark.

§CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politických vězňů 7, 111 21 Prague, Czech Republic.

1 Introduction

Labor market quotas have become a heavily-used governmental policy instrument in recent years. For example, in 2006 all publicly listed companies in Norway were required to increase female representation on their boards of directors to 40 percent. Following Norway's lead, the European Union and several countries worldwide have passed similar reforms (Bertrand et al., 2019). While there is a large body of literature that studies the effect of quota implementation on market outcomes, there is a lack of research that focuses on individual decision-making when an agent is forced to fulfill a quota. This paper fits into the latter category and introduces a new role of quotas: the attentional role.

We consider the following setup. A human resources (HR) manager in a large firm is a rationally inattentive (RI) decision-maker. Each day she faces a group of candidates. First, she sees candidate's ethnicity and gender (or other observable characteristics), which forms her prior beliefs of the candidate's qualities and potential future productivity levels. After that, the manager can acquire additional information about candidates – she can read resumes, ask questions, conduct tests, make comparisons of candidates and use other learning strategies. The key feature is that the information acquisition is flexible and endogenous – the manager does not have a fixed guide on how to learn about potential worker's future productivity. At the same time, she has cognitive (and/or time) limitations. We model these limitations as costly information acquisition. Therefore, the manager faces a trade-off between acquiring more precise information about candidates and the cost of this information. After acquiring optimal information, the manager hires the candidate with the highest expected value for the firm.

We follow the setup introduced by Matějka and McKay (2015), in which the agent is uncertain about the values of available options. The values of the options are modeled as an unknown draw from the known distribution. The agent has an opportunity to receive additional information about the realization of the draw in the manner that is optimal given the costs, which we model using the rational inattention framework introduced by Sims (1998, 2003). We think of the labor market candidates as available options and candidates' productivity levels as the values of options. Matějka and McKay (2015) show that the choice of a RI agent is typically stochastic and is characterized by the vectors of conditional and unconditional choice probabilities. In this paper we explore the effect of quota implementation on the behavior of a RI manager. We model a quota as a constraint on the unconditional choice probability of choosing a candidate from a particular group. Due to the law of large numbers, such a limitation on unconditional choice probability is essentially a limitation on the share of workers from a particular group in the overall composition of workers in the firm.

We analyze the behavior of the RI manager when quotas restrict her choice, and compare it with an unrestricted case and the situation in which a social planner subsidizes the manager's choice of certain alternatives. We find that choice probabilities of the manager in the constrained problem have the form of a generalized multinomial logit as in Matějka and McKay (2015) with an additional state independent component. In a choice among N candidates with the realized values $v(i|\omega)$ for $i \in \{1, ..., N\}$, our modified logit formula implies that the probability of choosing a candidate from group i is:

$$\mathcal{P}(i|\omega) = \frac{q_i e^{(v(i|\omega)-\varphi_i)/\lambda}}{\sum_{j=1}^{N} q_j e^{(v(j|\omega)-\varphi_j)/\lambda}},$$

where λ is the marginal cost of information, the q_i terms are quotas, and φ_i are state independent components. The form of choice probabilities shows that the manager behaves as if the value of the candidate is lower by φ_i. That is, the φ_i component induces an additive utility shifter in the decision-maker's preferences. Therefore, if a choice of a particular alternative is subsidized by $-\varphi_i$ then such a subsidy has exactly the same effect as the quota, which is the result we show in Section 3.3.

These adjustments to the logit model lead to the following change in the manager's behavior. If the choice problem is nontrivial, the RI manager who is forced to fulfill a quota always acquires information about candidates (Proposition 2). This feature is absent in the unconstrained problem, in which there are prior beliefs of the manager for which she decides not to acquire any additional information.

Further, we move to the analysis of implications of choosing different vectors of quotas by the social planner. In Section 4 we prove that the social planner using quotas can induce choice probabilities which coincide with choice probabilities of the unrestricted RI manager with any prior. The result has two immediate implications. Firstly, the social planner can always find such a quota that makes the probability of certain candidate being hired independent from the group identity, i.e. induce the fair (meritocratic) choice. Secondly, when the manager has inaccurate prior beliefs about distribution of candidates' productivities, the social planner using quotas can induce the manager to behave as if she has correct beliefs.

Finally, using the example with candidates from the two distinct social groups, we show two scenarios when a quota is socially optimal. In the first situation, the social planner maximizes the expected value of the chosen candidates without taking into account the information costs. In general, the social planner benefits by forcing the manager to fulfill quotas, but for some priors the social planner prefers not to impose a quota. In the second situation, we show that a quota can be useful instrument when applicants can falsify their types. In this situation, introduction of quotas incentivizes the applicant to truthfully reveal his type and leads to higher expected payoff for the manager compared to the situation without affirmative action.

Although we primarily focus on the effect of a quota in the labor market, the results of our analysis can be applied to study the individual behavior in other areas, e.g. a quota on the proportion of safe assets that must be in the portfolio of a financial manager or a quota on the number of orders a taxi driver can reject when searching for a client using peer-to-peer ride sharing applications (such as Uber, Lyft, or Yandex). We briefly discuss these applications in Section 7.

In the next section we review the related literature. Section 3 states the formal model of the manager's behavior with quotas and subsidies. Section 4 shows what conditional probabilities the social planner can achieve using the quota and studies how the social planner can induce fair choice and correct wrong beliefs. Section 5 demonstrates the implications of the model using the binary example. Section 6 discusses the socially optimal quota.

2 Literature

Our work contributes to the research on affirmative action and labor market discrimination. Affirmative action is "...any measure, beyond simple termination of a discriminatory practice, adopted to correct or compensate for past or present discrimination or to prevent discrimination from recurring in the future" (U. S. Commission on Civil Rights, 1977 p. 2). One of the most hotly debated types of affirmative action is the implementation of quotas. Coate and Loury (1993), in their famous paper, analyze a model of job assignment and show that quotas may lead to equilibria with persistent discrimination, due to feedback effects between expected job assignments and incentives to invest in human capital. Moro and Norman (2003) study the same problem in the general equilibrium setting and confirm the possibility that quotas can hurt the intended beneficiaries. These articles examine how affirmative action influences the behavior of the target group, and then its interaction with the behavior of the firm. In contrast, our study aims to investigate the individual decision-making process under quotas and consequences for policy design. A review of early studies on affirmative action can be found in Fang and Moro (2010).

To the date there is mixed empirical evidence of the effect of quotas on the quality of workers and firm's revenue. For example, Ahern and Dittmar (2012) show that firm value declined with a law mandating 40% representation of each gender on the board of public limited liability companies in Norway. In addition, the authors show that the average age and experience of the new female directors were significantly lower than that of the existing male directors and argue that this change led to a deterioration in operating performance. At the same time, Eckbo et al. (2019), using econometric adjustments and a larger data set, argue that the effect of implementation of quotas on both the value of firms and on the quality of directors was insignificant. Bertrand et al. (2019), by exploiting the same intervention, document that a quota resulted in significant improvement of the average observable qualifications of the women appointed to the boards and a decrease in the gender gap in earnings within boards, Besley et al. (2017), using Swedish data on the performance of politicians, show that a gender quota on the ballot increased the competence of male politicians. Ibanez and Riener (2018) use data from three field experiments in Colombia and show that the gains from attracting female applicants far outweigh the losses from deterring male applicants. Our paper proposes a mechanism that can possibly explain why the evidence on consequences of quotas is mixed. Namely, our model demonstrates that quotas may lead to increase in labor productivity under certain conditions, while it will have negative consequences under others.

Our study fits into the rational inattention literature, which originated in studies by Sims (1998, 2003). As a benchmark, we use the modified multinomial logit model of Matějka and McKay (2015), in which agents choose among alternatives without precise information about their values, but with an opportunity to study the options for some cost.[1] We analyze this model with an additional constraint on the unconditional probabilities of the choice of a certain alternative. Lindbeck and Weibull (2017) analyze investment decisions with delegation to a RI agent. They find that optimal contracts for an agent include a high reward for good investments and punishment for bad investments. Lipnowski et al. (2020) study a model, where

[1] See also Caplin and Dean (2015) for an alternative method for characterizing solutions in a similar environment.

a principal provides information to a RI agent, but the principal does not internalize this cost. They show that if there are more than two alternatives, a principal can improve material benefits from the choice by manipulating information. We analyze a similar principal-agent problem in Section 6.1, but with a different mechanism. We show that, for a set of parameters, a principal can force the manager to acquire more information by defining the level of quotas on unconditional choice probabilities and, thereby, increase the expected value of the manager's choice.

Fosgerau et al. (2020b) characterize equilibrium with a RI employer and candidates who choose how much effort to invest before being screened and use it to examine categorical inequality, including statistical discrimination, prejudice, and social capital. Acharya and Wee (2020) consider search and matching model with RI firms and find that during recessions firms become more selective. The results of our paper are complementary to these studies and can be used to investigate how affirmative action influences equilibrium outcomes.

Bartoš et al. (2016), in a field experiment, show that HR managers and landlords allocate their attention to job and rental applicants in line with rational inattention theory. For example, a non-European name or recent unemployment induces the HR manager to read a job application and CV in less detail, which negatively affects the probability of the applicant being invited for a job interview. The results of our study predict the attention allocation of decision-makers, such as HR managers, in the presence of quotas, i.e. whether they would blindly choose the quoted option or whether a quota would lead to greater information acquisition about the target group. Thus, the results of this study can provide a starting point for the empirical investigation of the effect of a quota on attention allocation. A detailed review of the RI literature can be found in Maćkowiak et al. (2020).[2]

Our study also relates to discussion on whether directly administering an activity is better than fixing transfer prices and relying on utility maximization to achieve

[2]Recent papers on strategic interactions with RI agents include Ravid (2019) and Yang (2019).

the same results in a decentralized fashion (Weitzman, 1974). We contribute to the discussion on this issue by comparing the behaviour of a manager operating under quotas and a manager whose choice is subsidized by a social planner.

3 The model

This section begins with a benchmark model – we describe the standard RI problem as in Matějka and McKay (2015) and Caplin et al. (2019) and its implications. Then we state our problem involving quotas and discuss the properties of the solution. Finally, we analyze the RI problem with subsidies.

The HR manager faces N candidates and wants to select the candidate with the highest value for the employer. There are finitely many states of the world Ω, with $\omega \in \Omega$ denoting a generic state. The values of the candidates differ from state to state, $v(i|\omega) \in \mathbb{R}$ is the value of the candidate from category $i \in \{1, ..., N\}$ in the state $\omega \in \Omega$. We refer to a candidate i as to a candidate from a category i. The decision-maker is uncertain about the realization of the state of the world. However, she knows a distribution of possible states of the world – this prior knowledge is described by a distribution $\mu \in \Delta(\Omega)$, where $\Delta(\Omega)$ denotes the set of all probability distributions over Ω, we assume that $\mu(\omega) > 0 \; \forall \omega \in \Omega$. She can refine her knowledge by processing costly information about the realization. Information processing results in a stochastic (possibly not purely stochastic) choice and, at the optimum, the decision problem can be treated as a problem of choosing conditional choice probabilities rather than the choice of information structure (Corollary 1 in Matějka and McKay 2015). We denote the conditional probability of a candidate i being selected when the realized state is ω as $\mathcal{P}(i|\omega)$.

3.1 Standard RI problem

The standard RI manager's problem is formalized as follows.

Standard (unconstrained) RI problem. *The manager's problem is to find a vector function of conditional choice probabilities* $\mathcal{P}^U = \{\mathcal{P}^U(i|\omega)\}_{i,\omega},\ \ i \in \{1, ..., N\},\ \ \omega \in \Omega$, *(the superscript "U" stands for "unrestricted") that maximizes expected payoff less the information cost:*

$$\max_{\{\mathcal{P}^U(i|\omega)\}_{i=1}^N} \left\{ \sum_{i=1}^N \sum_{\omega \in \Omega} v(i|\omega)\mathcal{P}^U(i|\omega)\mu(\omega) - \lambda\kappa(\mathcal{P}^U) \right\}$$

subject to

$$\forall i \in \{1, ..., N\},\quad \forall \omega \in \Omega: \quad \mathcal{P}^U(i|\omega) \geq 0, \tag{1}$$

$$\forall \omega \in \Omega: \quad \sum_{i=1}^N \mathcal{P}^U(i|\omega) = 1, \tag{2}$$

where unconditional choice probabilities are

$$\mathcal{P}^U(i) = \sum_{\omega \in \Omega} \mathcal{P}^U(i|\omega)\mu(\omega), \quad i \in \{1, ..., N\}.$$

The cost of information is $\lambda\kappa(\mathcal{P}^U)$, *where* $\lambda \in (0, +\infty)$ *is a given unit cost of information and* κ *is the amount of information that the manager processes, which is measured by the expected reduction in the entropy (Shannon, 1948; Cover and Thomas, 2012):*

$$\kappa(\mathcal{P}^U) = -\sum_{i=1}^N \mathcal{P}^U(i) \log \mathcal{P}^U(i) + \sum_{i=1}^N \sum_{\omega \in \Omega} \mathcal{P}^U(i|\omega) \log \mathcal{P}^U(i|\omega)\mu(\omega). \tag{3}$$

The entropy shape of information cost is common in the literature on rational inattention. Its use has been justified both axiomatically and through links to optimal coding in information theory (see Sims 2003; Matějka and McKay 2015; Denti et al. 2019 for discussions).

Matějka and McKay (2015) show that, at the optimum, the conditional probabilities of choosing a candidate $i \in \{1, ..., N\}$ follow the generalized logit form.

Theorem (Matějka and McKay, 2015)**.** *Conditional on the realized state of the* *world $\omega \in \Omega$ choice probabilities satisfy:*

$$\mathcal{P}^U(i|\omega) = \frac{\mathcal{P}^U(i)e^{v(i|\omega)/\lambda}}{\sum_{j=1}^N \mathcal{P}^U(j)e^{v(j|\omega)/\lambda}}.$$

The shape of conditional choice probabilities is similar to a multinomial logit, but is weighted with the coefficients $\mathcal{P}^U(i)$ which are endogenous to the decision problem and represent the probability of selecting a candidate i before the manager starts processing any information. These adjustments to the logit model reflect the fact that some candidates may look a priori better than others.

An important property of the solution is that parameters of the model may exist for which the manager decides not to acquire any information, and instead makes her decision based solely on her prior knowledge. In this situation, she simply chooses the candidate with the highest a priori expected value.[3] In terms of the labor market, this means that some categories of workers may not be given any attention and are consequently not hired. As we show in the next section, the social planner can force the manager to receive at least some information about the candidates – this can be achieved via the use of quotas.

3.2 Quotas

We consider a departure from the standard RI problem when the manager is not completely free in her choice.[4] Instead, some authority limits her choice in that, for all categories, the share of the candidates hired from a category i should be equal to $q_i \in (0, 1)$. [5]

[3]See Caplin and Martin (2017), who analyze a discrete choice problem of a RI agent with costly information acquisition, and show that if there is a high quality default option the manager chooses zero attentional effort.

[4]In the context of information theory the problem of computing the capacity of constrained discrete channel is considered, among others, by Blahut (1972).

[5]We focus on the case with binding quotas for all alternatives, since (a) if the quotas have a form of weak inequality, then, if it is not binding, the solution is the same as the solution to the unconstrained problem; and (b) the case with quotas only for some categories is considered in Appendix D where we show that results are similar. In addition, we restrict the quota vector to be interior because if some of its components are 0, the problem boils down to the situation in which

RI problem with quotas. *The manager's problem is to find a vector function of conditional choice probabilities* $\mathcal{P} = \{\mathcal{P}(i|\omega)\}_{i,\omega}, \quad i \in \{1, ..., N\}, \quad \omega \in \Omega,$ *that maximizes expected payoff less the information cost:*

$$\max_{\{\mathcal{P}(i|\omega)\}_{i=1}^{N}} \left\{ \sum_{i=1}^{N} \sum_{\omega \in \Omega} v(i|\omega)\mathcal{P}(i|\omega)\mu(\omega) - \lambda\kappa(\mathcal{P}) \right\}$$

subject to (1)-(3) and

$$\forall i \in \{1, ..., N\}: \quad \mathcal{P}(i) = \sum_{\omega \in \Omega} \mathcal{P}(i|\omega)\mu(\omega) = q_i, \quad q_i \in (0, 1), \tag{4}$$

where $\mathbf{q} = (q_1, ..., q_N)^T$ *is the vector of quotas and*

$$\sum_{i=1}^{N} q_i = 1.$$

Choice probabilities at the optimum follow:

$$\mathcal{P}(i|\omega) = \frac{q_i e^{(v(i|\omega) - \varphi_i)/\lambda}}{\sum_{j=1}^{N} q_j e^{(v(j|\omega) - \varphi_j)/\lambda}}. \tag{5}$$

This result is formalized in the following proposition:

Proposition 1. *Choice probabilities that are the solution of the RI manager problem with quotas are of a generalized logit form: logit choice probabilities with an additive state-independent component. Conditional choice probabilities* $\{\mathcal{P}(i|\omega)\}_{i,\omega}$, $i \in \{1, ..., N\}, \quad \omega \in \Omega$, *are the solution of the RI problem with quotas if they satisfy (4) and (5). Moreover, solution of the RI problem with quotas is unique.*

Proof. See Appendix A.1. □

The terms φ_i are the Lagrange multipliers on the constraints on unconditional choice probabilities. In choice probabilities they play a role of utility shifters, that

the choice is limited and the solution coincides with the solution to the standard RI problem with restricted menu. When the quota equals 1 for a particular category, the solution is trivial – the manager does not acquire any information.

is, the manager's behavior follows logit rule, but with utilities which are changed by some value that depends on the marginal cost of information, prior beliefs and the value of a quota. In Section 3.3 we relate φ to subsidies which are required to be paid to the manager when she hires workers from a certain category.

Proposition 1 states that the solutions to the standard RI problem and the RI problem with quotas have a similar form. However, there is a crucial difference in the information acquisition strategies of a RI manager with and without quotas. We express it in the following proposition:

Proposition 2. *If $v(i|\omega) - v(j|\omega)$ is not constant across states $\forall i, j \in \{1, ..., N\}$ and $i \neq j$, then the following holds: the RI manager with quotas always acquires information.*

Proof. See Appendix A.2. □

Proposition 2 means that the employer will never blindly choose a candidate from a certain group based only on her prior beliefs, but will acquire some information.

The assumption that the cost of information is expressed as the expected reduction in entropy is not crucial for Proposition 2 to hold. For example, it also holds for the cost functions which are based on the generalized entropy (Fosgerau et al., 2020a). More generally, Proposition 2 holds for any non-negative continuously differentiable function $\kappa(\mathcal{P})$, for which the cost of marginal change in choice probabilities is zero at the point of initial uncertainty. The intuition for this result comes from the fact that a quota forces the manager to choose not between candidates from different categories, but also between candidates from the same category, which brings incentives for information acquisition. According to conditions in Proposition 2, there is a beneficial deviation from the state-independent probabilities. Therefore, the manager can benefit from acquiring at least some information in order to improve her choice.

3.3 Subsidies

We are interested in understanding how a manager's attention strategy depends on the particular form of affirmative action chosen by the government. One of the forms of affirmative action, alternative to quotas, is an employment subsidy policy, when a firm receives subsidies if it employs certain categories of workers (for surveys see, e.g. Card et al. 2010, 2017). In this situation, the government does not have access to the values of specific candidates for a firm and introduces the same subsidy for all possible realizations of the values of the hired candidates from a particular category.

Such policy in our setting changes the values of candidates from $v(i|\omega)$ to $v(i|\omega)+S_i \quad \forall\omega \in \Omega$, where S_i is a subsidy for choosing a candidate from category i. Let us start the analysis with the definition of the RI problem with subsidies. If the government introduces an employment subsidy policy, then the manager solves the following problem:

RI problem with subsidies. *The manager's problem is to find a vector function of conditional choice probabilities* $\mathcal{P}^S = \{\mathcal{P}^S(i|\omega)\}_{i,\omega}, \quad i \in \{1, ..., N\}, \quad \omega \in \Omega$, *(the superscript "S" stands for "subsidy") that maximizes expected payoff less the information cost:*

$$\max_{\{\mathcal{P}^S(i|\omega)\}_{i=1}^N} \left\{ \sum_{i=1}^{N} \sum_{\omega\in\Omega} (v(i|\omega) + S_i)\mathcal{P}^S(i|\omega)\mu(\omega) - \lambda\kappa(\mathcal{P}^S) \right\},$$

subject to (1)-(3) and where S_i *is a subsidy for choosing a candidate* i.

In this case, the solution to the manager's problem follows the standard modified generalized multinomial logit formula, but with the changed value of the candidate i by S_i:

$$\mathcal{P}^S(i|\omega) = \frac{\mathcal{P}^S(i)e^{(v(i|\omega)+S_i)/\lambda}}{\sum_{j=1}^{N} \mathcal{P}^S(j)e^{(v(j|\omega)+S_j)/\lambda}}. \tag{6}$$

And unconditional choice probabilities are

$$\mathcal{P}^S(i) = \sum_{\omega \in \Omega} \mathcal{P}^S(i|\omega)\mu(\omega).$$

Equation (6) provides intuition about the nature of the additive component φ_i, $i \in \{1, ..., N\}$ from the solution to the RI problem with quotas. This component can be interpreted as the government subsidies that are needed to be added to the values of candidates in order to induce the RI manager to choose them with required unconditional probabilities.

We refer to a vector of subsidies $\mathbf{S}$ as *the subsidy which supports quotas* $\mathbf{q}$ when unconditional probabilities of the manager's choice are equal to quotas, $\mathcal{P}^S(i) = q_i, i \in \{1, ..., N\}$.[6] In the following proposition we provide sufficient conditions for existence of a subsidy which supports quotas $\mathbf{q}$. Namely, it requires that a solution of the RI problem with a subsidy vector $\mathbf{S} = -\boldsymbol{\varphi}$, where $\boldsymbol{\varphi}$ is a vector of Lagrange multipliers from solution to the RI problem with quotas $\mathbf{q}$, has a unique solution.

Proposition 3. *If for a vector of quotas* $\mathbf{q}$ *the vector of resulting Lagrange multipliers* $\boldsymbol{\varphi}$ *has the following property: there does not exist a set* $\{a_j\}_j$ *such that* $\sum_{j \neq i} a_j = 1$ *and for any* $\omega \in \Omega$

$$e^{(v(i|\omega) - \varphi_i)/\lambda} = \sum_{j \neq i} a_j e^{(v(j|\omega) - \varphi_j)/\lambda},$$

then there exists a vector of subsidies which supports quotas $\mathbf{q}$.

Proof. Let us consider a problem with a subsidy vector $\mathbf{S} = -\boldsymbol{\varphi}$. According to Lemma S2 from the Online Supplement to Matějka and McKay (2015), the condition of the proposition guarantees that the solution to this problem results in a unique vector of unconditional choice probabilities $\mathbf{q}$. □

[6] It is important to note that the vector of supporting subsidies is not unique. That is so because if $\mathbf{S}^*$ is a vector of subsidies which supports quotas $\mathbf{q}$, then any vector which is obtained by adding any number to all components of $\mathbf{S}^*$ is also a vector of supporting subsidies.

Proposition 3 establishes that for some vectors of quotas $\mathbf{q}$ we can find a vector of subsidies $\mathbf{S}$, such that unconditional choice probabilities in the problem with quotas $\mathbf{q}$ and in the problem with subsidy $\mathbf{S}$ coincide. Moreover, since the solution to the RI problem with quotas is unique, if one compares a quota $\mathbf{q}$ with the subsidy which supports this quota, not only unconditional but also conditional choice probabilities coincide. This means that subsidies and quotas are behaviorally equivalent in our setting.[7] It is formalized in the following corollary.

Corollary 1. *If the conditions of Proposition 3 hold, then the information acquisition strategy and conditional choice probabilities of the RI manager when her choice is restricted by a quota* $\mathbf{q}$ *are identical to those which correspond to the problem with the subsidy which supports this quota.*

Let us briefly discuss how the results in Proposition 3 depend on the functional form of the cost of information. The two elements are needed for the proposition to hold: (i) the quota leads to a solution which has the same shape as the solution to the unrestricted problem but with an additive utility shifter; and (ii) the solution to the problem with values of the options changed by this shifter should be unique. It is hard to expect these conditions to hold for an arbitrary cost function. However, the results of Proposition 3 hold for the class of Generalized Entropies (Fosgerau et al., 2020a) if there is a unique solution for the quoted problem[8]. It is so, since for the case of Generalized Entropy similar steps as in the proof of Proposition 1 can be made in order to show that the Lagrange multiplier on a quota constraint plays a role of a utility shifter in the solution.

[7] In Appendix C we provide a solution for a binary example with subsidies. We show that while the behavior of the manager under quotas and subsidies is the same, the utility of the manager is different.

[8] The uniqueness of a solution in such models is not guaranteed.

4 Effect of quotas on conditional choice probabilities

So far we have analyzed the manager's problem for a given level of quota. In this section, we analyze the features of conditional probabilities which the social planner can induce by quotas. We start by stating the following proposition:

Proposition 4. *For any vector of weighting coefficients* $\boldsymbol{\beta} \in [0,1]^N$*, such that* $\sum_{i=1}^N \beta_i = 1$*, there exists a vector of quotas* $\mathbf{q} \in [0,1]^N$*, such that* $\sum_{i=1}^N q_i = 1$*, which induces the following choice probabilities as a solution to the RI problem with quotas* $\mathbf{q}$

$$\forall i \in \{1,...,N\}, \quad \omega \in \Omega : \quad \mathcal{P}(i|\omega) = \frac{\beta_i e^{v(i|\omega)/\lambda}}{\sum_{j=1}^N \beta_j e^{v(j|\omega)/\lambda}}. \tag{7}$$

Proof. See Appendix A.3. □

Proposition 4 establishes an important feature of quotas: for any distribution of states of the world the social planner can implement such vector of quotas that the manager's behavior replicates the behavior of the unconstrained RI manager with any prior beliefs. Now we consider two situations in which implementing such a quota is optimal.

4.1 Fair quota

Original goal of the affirmative action was and for many policies is to ensure that candidates are "treated [fairly] during employment, without regard to their race, creed, color, or national origin" (John F. Kennedy, 1961). In terms of our model this goal is to ensure that the probability of certain worker being hired does not depend on the group identity and depends only on the job-relevant characteristics of candidates. In this section, we show that there exists a quota that achieves this goal.

Formally, the solution to the standard RI maximization problem is

$$\forall i \in \{1, ..., N\}, \quad \omega \in \Omega: \quad \mathcal{P}^U(i|\omega) = \frac{\mathcal{P}^U(i) e^{v(i|\omega)/\lambda}}{\sum_{j=1}^{N} \mathcal{P}^U(j) e^{v(j|\omega)/\lambda}}, \tag{8}$$

where $\mathcal{P}^U(i)$ corresponds to the bias towards hiring a candidate from group i. Such choice probabilities may result in a situation in which $v(i|\omega) = v(j|\omega)$ for some $i \neq j$, $\omega \in \Omega$, but $\mathcal{P}^U(i|\omega) > \mathcal{P}^U(j|\omega)$ just because a candidate i comes from a group that seems a priori better.

The goal of the social planner is to make choice probabilities depend on the realized values only, i.e. eliminate $\mathcal{P}^U(i)$ on the right-hand side in the equation (8). Affirmative action policy of such social planner results in equal chances of being hired for candidates with the same productivity.

One of the policies that aspires to achieve this goal is blind or de-identified approach to reviewing candidates (Goldin and Rouse, 2000). In practice, however, it is not always possible to fully hide the group identifier, since it can be approximated by other information which can lead to lower quality of the choice (see, for example, Ray and Sethi (2010) and Antonovics and Backes (2014) on the adverse effects of color-blind affirmative action). Moreover, such policies can have unexpected results and lower the representation of the discriminated group if there exists positive discrimination (Hiscox et al., 2017).

Corollary 2 states that the social planner can choose such quota, which we call *fair,* that makes conditional choice probabilities dependent on candidates' job-relevant characteristics only. Therefore, the resulting choice outcomes with fair quota is as if there is a blind or de-identified policy. At the same time, by using fair quota social planner can eliminate the bias, while keeping the information about the group identity.

Corollary 2. *There exists such* $\mathbf{q}$ *which induces*

$$\forall i \in \{1, ..., N\}, \quad \forall \omega \in \Omega : \quad P(i|\omega) = \frac{e^{v(i|\omega)/\lambda}}{\sum_{j=1}^{N} e^{v(j|\omega)/\lambda}}.$$

4.2 Inaccurate priors

Usually the literature highlights two types of discrimination: taste-based (Becker, 1957) and statistical discrimination (Phelps, 1972). Without any additional assumptions, the theory of rational inattention allow us to model the latter type, which is typically assumed to be driven by limited information about the particular candidate's characteristics and correct beliefs about the group distributions of the relevant outcome. At the same time, people often have inaccurate beliefs about the performance distributions of particular groups (Bohren et al., 2020). Such misperception can not only increase discrimination, but also is harmful for the manager, since she would acquire information in a sub-optimal way and choose candidates with lower expected productivity. In this section, we show that the social planner, who knows the correct distribution of productivities within and between different groups, can implement the quota that induces the manager to behave as if she also has correct beliefs.

In terms of our model, it would mean that the manager has a wrong prior: she believes that distribution of possible states of the world is $\mu^I(\omega) \in (0, 1), \quad \forall \omega \in \Omega$, while in reality this distribution is $\mu(\omega) \in (0, 1), \quad \forall \omega \in \Omega$. Such incorrect belief results in conditional choice probabilities $\mathcal{P}^I(i|\omega)$.

Let us consider the following possible goal of the social planner. It is to make conditional choice probabilities appear as if the manager has correct beliefs. That is, to make $\mathcal{P}^U(i|\omega) = \mathcal{P}^I(i|\omega), \forall i \in \{1, ..., N\}, \forall \omega \in \Omega$. Corollary 3 states that the social planner can choose such a quota that achieves this goal. It is important to notice that a quota is not simply $q_i = \mathcal{P}^U(i), \quad \forall i \in \{1, ..., N\}$, but is adjusted to the prior of the manager. Therefore, she will be unable to meet the quota and unconditional probabilities will be equal to the desired one.

Corollary 3. *For any misperceived distribution of possible states of the world* $\mu^w \in (0,1)$ *there exists such* $\mathbf{q}$ *which induces*

$$\forall i \in \{1, ..., N\}, \quad \forall \omega \in \Omega : \quad \mathcal{P}^I(i|\omega) = \frac{\mathcal{P}^U(i) e^{v(i|\omega)/\lambda}}{\sum_{j=1}^{N} \mathcal{P}^U(j) e^{v(j|\omega)/\lambda}},$$

where $\mathcal{P}^U(i|\omega)$ *is a conditional choice probability of a manager with the correct prior belief.*

5 Example with two groups of candidates

In order to illustrate the logic of the model, we consider a simple example in which the manager chooses between two candidates from two different social groups, $i \in \{1, 2\}$. There are two states of the world, $\omega \in \{1, 2\}$. For simplicity, the type 1 candidate is the safe choice that always has the constant value $v(1) = v(1|1) = v(1|2) = C$. The type 2 candidate is the risky choice that can take values $v(2|1) = 0$ with the probability b and $v(2|2) = 1$ with the probability $1 - b$. In a labor market context, we can assume that the share b of workers from a category 2 has low productivity, while the share $1 - b$ has high productivity. These probabilities (or shares) are priors of the manager and she does not know what the realization of the state of the world is. The manager has an opportunity to acquire some costly information about the realization.

The manager's choice is restricted in that, on average, the share q of hired candidates should be type 2 (risky) and share $1 - q$ of chosen candidates should be type 1 (safe). In terms of rational inattention the manager has restrictions on unconditional choice probabilities.

To solve the problem we must find conditional probabilities $\mathcal{P}(i|\omega)$. We show in

Appendix B that the solution is

$$\mathcal{P}(2|0) = \frac{-b - q + (b+q-1)e^{\frac{1}{\lambda}} + \sqrt{(b+q-(b+q-1)e^{\frac{1}{\lambda}})^2 + 4q(be^{\frac{1}{\lambda}} - b)}}{2(be^{\frac{1}{\lambda}} - b)},$$

$$\mathcal{P}(2|1) = \frac{q - b\mathcal{P}(2|0)}{1-b}.$$

It is worth noticing that in the formulas above conditional choice probabilities $\mathcal{P}(2|0)$ and $\mathcal{P}(2|1)$ do not depend on the value of the safe candidate C unlike in the unconstrained RI problem. There are two intuitive explanations for this. First, it follows from Proposition 3 that the required unconditional choice probabilities can be achieved by subsidies. In this example it is enough to subsidize (or put a fine on) candidates from the safe group, making its value for the firm equal to some $\tilde{C}(q)$, which, of course, does not depend on the initial C. The second intuitive reason is as follows. Whereas an unrestricted manager is interested in the relative payoffs of hiring candidates from different groups in the same state, a quota-bounded manager compares the relative payoffs of hiring candidates from the same group in different states.

For a given set of parameters, Figure 1 shows the expected reduction in entropy as a function of b. In the standard RI problem, when b is close to 0 or 1 the manager decides not to process information and selects one of the candidates with certainty. However, when the manager is forced to fulfill the quota, she always acquires information and, hence, there are no non-learning areas. For example, when b is close to 1, she is forced to choose a risky candidate with positive probability, and it is profitable to acquire information in order to choose the risky candidate with high value rather than to randomly pick him.

At the same time, the quota-bounded manager can acquire less information than in the standard RI problem (Figure 1). Accordingly, the effect of the quota on the amount of acquired information is ambiguous.

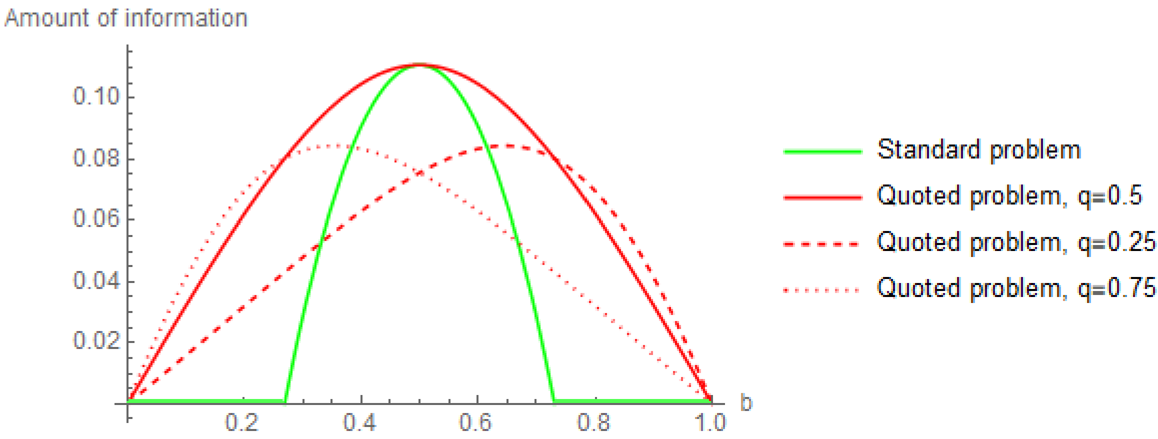

Figure 1: Amount of information as a function of b and $\lambda = 0.5$, $C = 0.5$. The green curve is for the standard RI problem and the red curves are for the quoted RI problem: the solid curve is for $q = 0.5$, the dotted curve for $q = 0.75$ and the dashed curve for $q = 0.25$.

We now explore how the quota affects the expected value of the chosen candidates. In terms of the labor market this question can be restated in the following way: does a quota necessarily mean that the expected value (or productivity) of the hired workers will fall? The definition of the expected value of the chosen risky candidate can be found below.

Definition. *The expected value of the chosen risky candidate is* $\frac{(1-b)\mathcal{P}(2|1)}{\mathcal{P}(2)}$. *This is the ratio of the probability of the chosen risky candidate being of high value to the probability of choosing any risky candidate.*

Figure 2 illustrates that the expected value of the chosen risky candidate is higher (lower) when the quota on it is smaller (larger) than the unconditional probability of choosing it in the standard RI problem.

6 Socially optimal quota

In this section we use the example from Section 5 and discuss two scenarios when a quota can be socially optimal.

Firstly, we analyze the situation in which the social planner wants to maximize

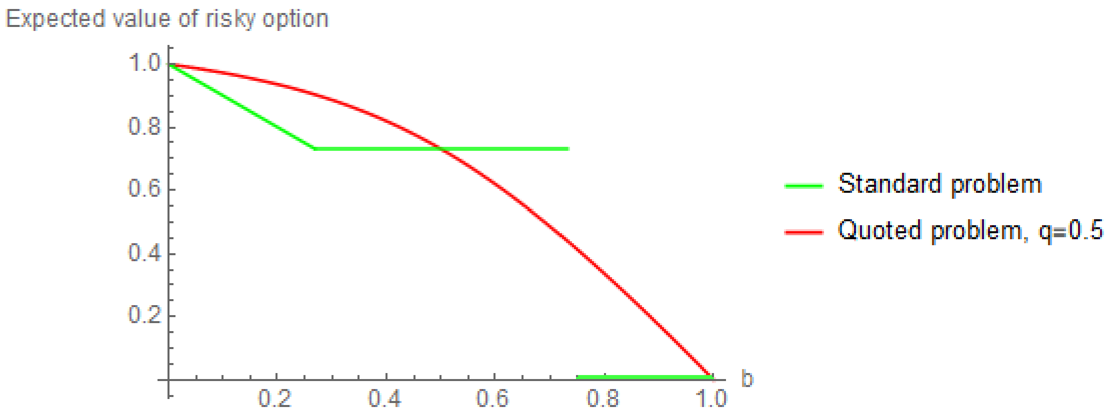

Figure 2: The expected value of the risky candidate conditional on being chosen as a function of b and $\lambda = 0.5$, $q = 0.5$, $C = 0.5$. The green curve is for the standard RI problem and the red curve is for the quoted RI problem.

the expected value of the chosen candidates without taking into account the information costs of the manager. Secondly, we show that quotas could be a mechanism that induce truth-telling by applicants when they are able to falsify their group identity.

6.1 Value maximizing quota

We follow the Lipnowski et al. (2020) and consider the following problem. The social planner maximizes the expected value of the chosen candidate and does not take into account the cost of information. This is reasonable, for instance, if positive production externalities exist. However, while the manager is making hiring decisions she may not take these externalities into account. Thus, the maximization problem of the manager and the social planner (organization, industry as a whole, or government) may differ. However, in contrast to the Lipnowski et al. (2020), in our case the social planner chooses optimal quotas $\mathbf{q}$ and not information structure available to the manager. The maximization problem is

$$\max_{\mathbf{q}} \left\{ \sum_{i=1}^{N} \sum_{\omega \in \Omega} v(i|\omega) \mathcal{P}(i|\omega, \mathbf{q}) \mu(\omega) \right\},$$

where $\mathcal{P}(i|\omega, \mathbf{q})$ is a solution to the RI problem with the quota $\mathbf{q}$.

We refer to the solution of this problem as a *Value maximizing quota*. In contrast to information filtering proposed by Lipnowski et al. (2020), a quota allows the social planner to increase the expected value of the chosen candidate even when the state is binary.

Figure 3 illustrates the solution to the social planner's problem as a function of b, and $\lambda = 0.5$, $C = 0.5$. When the social planner maximizes the expected value of chosen alternatives, there are still non-learning areas, but they are smaller than in the standard RI problem. The reason for the presence of the non-learning areas is as follows. Consider the situation in which b is small, i.e. the probability of the risky candidate being good is high and the manager choose him with certainty without acquiring any information. When the non-trivial quota is implemented, the manager will acquire some information in order to find out whether the risky candidate is good or bad, but the improvement in the expected value of chosen risky options would not compensate for the loss of good risky candidates that were mistakenly rejected. Therefore, the social planner prefers not to constrain the manager, or, in other words, he prefers to implement the quota that would force the manager to always choose a risky candidate – the same action that the manager would take without any constraints. Similar logic applies when the probability of the bad state is high.

Outside these non-learning areas, the social planner induces the manager to acquire more information than in the standard RI problem (Figure 5) and, hence, increases the expected value of the chosen candidate (Figure 4). Thus, in this example, it is optimal to establish a quota that is higher (lower) than the unconditional probability in the standard RI problem when the state is more likely to be bad (good).[9]

[9] In the working paper we provide an example when the social planner has limited information about the distribution of alternatives' values. In this situation, setting a quota may decrease the expected value of the chosen candidates.

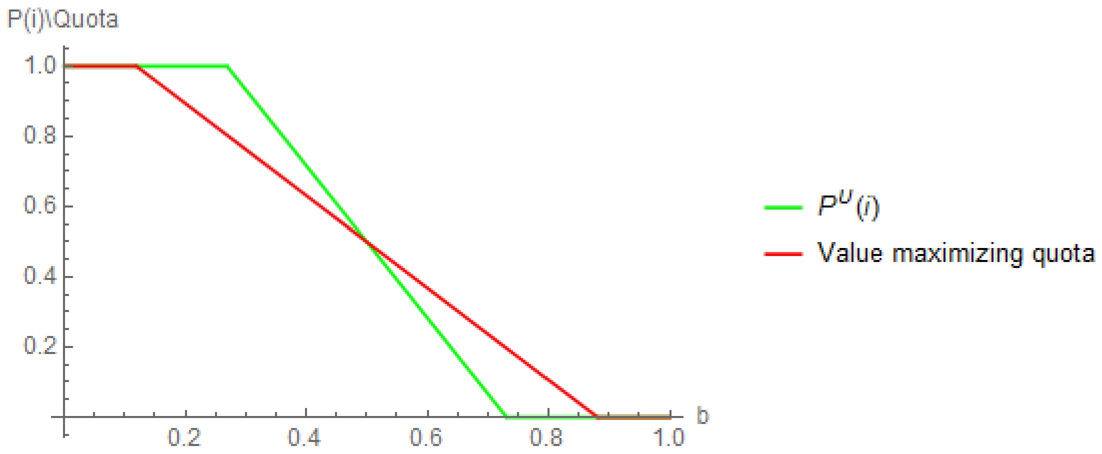

Figure 3: Optimal quota as a function of b and $\lambda = 0.5$, $C = 0.5$. The green curve is for the standard RI problem and the red curve is for the quoted RI problem.

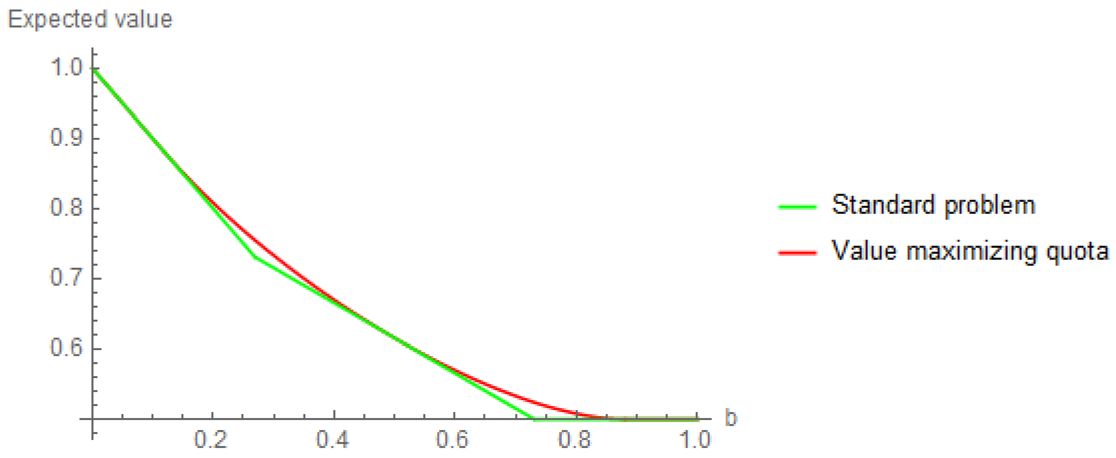

Figure 4: The expected value of the chosen candidates as a function of b and $\lambda = 0.5$, $C = 0.5$. The green curve is for the standard RI problem and the red curve is for the quoted RI problem.

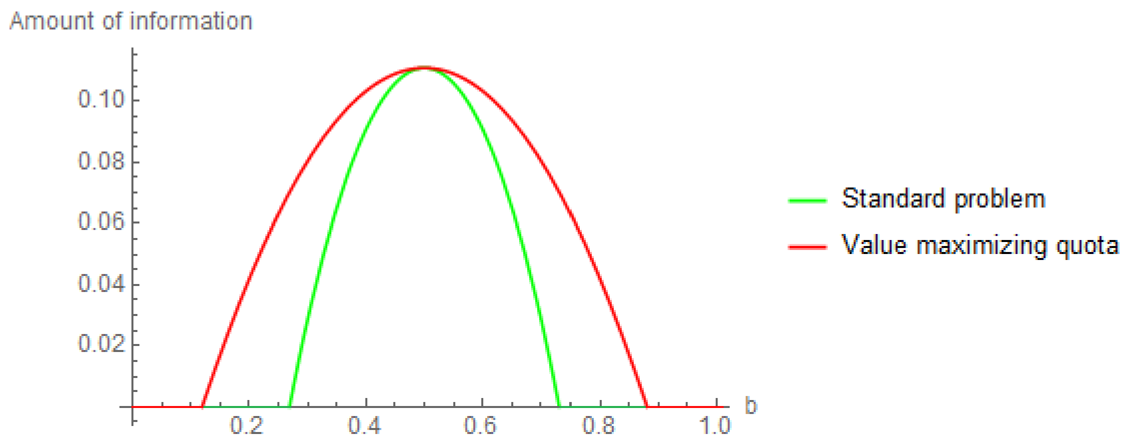

Figure 5: Amount of information as a function of b and $\lambda = 0.5$, $C = 0.5$. The green curve is for the standard RI problem and the red curve is for the quoted RI problem.

6.2 Type Falsification

We extend our model by allowing candidates to control their perceived types (Kim and Loury, 2019). When the HR manager knows candidates group identity (for example, race, nationality, or gender), she can better understand what is the expected productivity of the candidate and, in addition, acquire information more efficiently. For example, for particular groups it can be harder to enter the university because of the discrimination and/or peer effects. Therefore, the university degree for a candidate from the certain discriminated group signals his exceptional quality. At the same time, many group tags are easily falsifiable. Therefore, without the mechanism which incentivizes the applicant to truthfully report his group identity, the employer has to disregard the group identity information and needs additional costly tests in order to define the applicant's quality.

Let us provide an example of how such type falsification can be described in the context of the RI HR manager model. We assume that candidates know beliefs of the manager, but do not know their value for the firm. That is, candidates know their identity, know the distribution of possible productivities, but do not know their productivity. Candidates maximize the unconditional probability of being hired and can affect it by costless lying about their identity. A candidate (he) from group $i \in \{1, ..., N\}$ chooses to present himself to the manager (she) that he is from the group $j \neq i$ if and only if it increases the expected unconditional probability of being chosen.

If any of the applicants falsify their identity, then the manager can not distinguish between them. In such case, candidates are pooled into one category. Therefore, she chooses between identical candidates who take values C with probability 0.5, 1 with probability $0.5(1-b)$ and 0 with probability $0.5b$. We show that in order to counteract the type falsification the manager can commit to certain unconditional choice probabilities, that is, the manager can introduce a quota as a mechanism, which incentivizes applicants to report their true type.

The game has the following timeline:

1. The manager decides if and what quotas she imposes;
2. Applicants decide whether they will report their true type or lie;
3. Values are realized;
4. The manager acquires information and makes choice.

When applicants falsify their identity, the unconditional probabilities of candidates being chosen are

$$\mathcal{P}(1) = \frac{e^{C/\lambda}}{e^{C/\lambda} + e^{1/\lambda}}(1-b) + \frac{e^{C/\lambda}}{e^{C/\lambda} + 1}b,$$

$$\mathcal{P}(2) = 1 - \mathcal{P}(1).$$

If there is no quota, the candidate from group 1 falsifies his group identity when $\mathcal{P}(1) > \mathcal{P}^U(1)$ and the candidate from group 2 when $\mathcal{P}(2) > \mathcal{P}^U(2)$, where $\mathcal{P}^U(i)$ is an unconditional choice probability of choosing a candidate from group i when the types are reported truthfully.

The manager can commit to choose the type 1 applicant with probability of $q^* = \mathcal{P}(1)$ if he reports his type truthfully. Accordingly, she commits to choose the type 2 applicant with unconditional probability $1 - q^* = \mathcal{P}(2)$. In this situation, both candidates do not have incentives to misreport their identity, since they can not increase their probability of being chosen by falsifying their identity. It is important to notice, that any other quota will lead to type falsification: if $q < q^*$ then the safe candidate has incentives to lie; if $q > q^*$ then the risky candidate has incentives to lie. Therefore, there is the unique quota that induces truthful type revelation.

At the same time, if candidates can falsify their identity, the manager always prefers to commit to the quota, which supports type revelation (Figure 6). The intuition for this results is that with a quota the manager still has access to the information about applicant's type and can acquire information more efficiently.

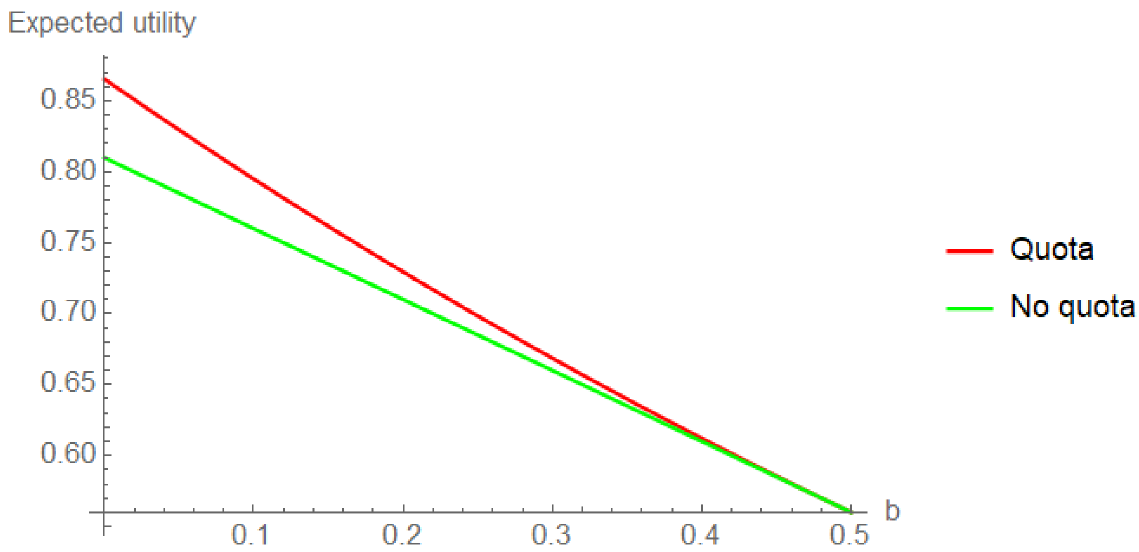

Figure 6: The expected value of the chosen candidates as a function of b and $\lambda = 0.5$, $C = 0.5$. The green curve is for the standard RI problem and the red curve is for the quoted RI problem.

7 Conclusion

In this paper, we study the optimal behavior of the RI manager who is forced to fulfill quotas when making a choice from a discrete menu. While through the paper we have used labor market settings to illustrate how quotas can influence the manager's attention allocation, the proposed model can be also used to analyze the effect of quotas in other areas. For example, the model in Section 5 can be applied to analyze financial interventions. Consider a situation in which the agent manages a financial portfolio and chooses between risky and safe assets. The financial regulator wants to increase the proportion of safe assets in the agent's portfolio. Thus, the regulator imposes a quota. This intervention can lead to higher screening efforts by the agent and, hence, to the more profitable and diversified portfolio.

The model also can be applied to analyze the effect of quotas that are nowadays used in many mobile applications. For example, in many peer-to-peer ride sharing applications, the driver does not know some details of the order before accepting it. In addition, she faces a quota on the number of orders that she can reject. The primary goal of such quotas is to ensure that drivers accept a sufficient number of orders that can be not as profitable for her as some other orders. This restriction forces the driver to calculate the benefits and costs of accepting an order based on the distance, road condition, traffic congestion, etc. At the same time, such a policy can force the driver to switch to a competing platform. Our model can be used to find the optimal quota that will be beneficial for the platform and not too restrictive for the drivers.

References

Acharya, S. and Wee, S. L. (2020). Rational inattention in hiring decisions. *American Economic Journal: Macroeconomics*, 12(1):1–40.

Ahern, K. R. and Dittmar, A. K. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *The Quarterly Journal of Economics*, 127(1):137–197.

Antonovics, K. and Backes, B. (2014). The effect of banning affirmative action on college admissions policies and student quality. *Journal of Human Resources*, 49(2):295–322.

Bartoš, V., Bauer, M., Chytilova, J., and Matějka, F. (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review*, 106(6):1437–1475.

Becker, G. S. (1957). *The economics of discrimination.* University of Chicago press.

Bertrand, M., Black, S. E., Jensen, S., and Lleras-Muney, A. (2019). Breaking the glass ceiling? the effect of board quotas on female labour market outcomes in norway. *The Review of Economic Studies*, 86(1):191–239.

Besley, T., Folke, O., Persson, T., and Rickne, J. (2017). Gender quotas and the crisis of the mediocre man: Theory and evidence from sweden. *American Economic Review*, 107(8):2204–2242.

Blahut, R. (1972). Computation of channel capacity and rate-distortion functions. *IEEE transactions on Information Theory*, 18(4):460–473.

Bohren, A., Haggag, K., Imas, A., and Pope, D. G. (2020). Inaccurate statistical discrimination: An identification problem.

Caplin, A. and Dean, M. (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203.

Caplin, A., Dean, M., and Leahy, J. (2019). Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies*, 86(3):1061–1094.

Caplin, A. and Martin, D. (2017). Defaults and attention: The drop out effect. *Revue économique*, 68(5):747–755.

Card, D., Kluve, J., and Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The economic journal*, 120(548):452–477.

Card, D., Kluve, J., and Weber, A. (2017). What works? a meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3):894–931.

Coate, S. and Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes? *American Economic Review*, 83(5):1220–1240.

Cover, T. M. and Thomas, J. A. (2012). *Elements of information theory*. John Wiley & Sons.

Denti, T., Marinacci, M., and Montrucchio, L. (2019). A note on rational inattention and rate distortion theory. *Decisions in Economics and Finance*, pages 1–15.

Eckbo, B. E., Nygaard, K., and Thorburn, K. S. (2019). Board gender-balancing and firm value.

Fang, H. and Moro, A. (2010). Theories of statistical discrimination and affirmative action: A survey. NBER Working Papers 15860, National Bureau of Economic Research, Inc.

Fosgerau, M., Melo, E., de Palma, A., and Shum, M. (2020a). Discrete choice and rational inattention: A general equivalence result. *International Economic Review*.

Fosgerau, M. and Nielsen, N. (2021). Similarity, discrete choice and rational inattention.

Fosgerau, M., Sethi, R., and Weibull, J. W. (2020b). Categorical screening with rational inattention. *Available at SSRN.*

Goldin, C. and Rouse, C. (2000). Orchestrating impartiality: The impact of" blind" auditions on female musicians. *American economic review*, 90(4):715–741.

Hiscox, M. J., Oliver, T., Ridgway, M., Arcos-Holzinger, L., Warren, A., and Willis, A. (2017). Going blind to see more clearly: Unconscious bias in australian public service shortlisting processes. *Behavioral Economics Team of the Australian Government.*

Ibanez, M. and Riener, G. (2018). Sorting through affirmative action: Three field experiments in colombia. *Journal of Labor Economics*, 36(2):437–478.

John F. Kennedy (1961). *Executive Order 10925 - Establishing the President's Committee on Equal Employment Opportunity.* The American Presidency Project. `https://www.presidency.ucsb.edu/node/237176`. Accessed February 2, 2021.

Kim, Y.-C. and Loury, G. C. (2019). To be, or not to be: Stereotypes, identity choice and group inequality. *Journal of Public Economics*, 174:36–52.

Lindbeck, A. and Weibull, J. (2017). Investment, Rational Inattention, and Delegation. Working Paper Series 1171, Research Institute of Industrial Economics.

Lipnowski, E., Mathevet, L., and Wei, D. (2020). Attention management. *American Economic Review: Insights*, 2(1):17–32.

Maćkowiak, B., Matějka, F., and Wiederholt, M. (2020). Rational inattention: A review.

Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–298.

Moro, A. and Norman, P. (2003). Affirmative action in a competitive economy. *Journal of Public Economics*, 87(3):567 – 594.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review*, 62(4):659–661.

Ravid, D. (2019). Bargaining with rational inattention. *Available at SSRN 2957890.*

Ray, D. and Sethi, R. (2010). A remark on color-blind affirmative action. *Journal of Public Economic Theory*, 12(3):399–406.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423.

Sims, C. A. (1998). Stickiness. In *Carnegie-Rochester Conference Series on Public Policy*, volume 49, pages 317–356. Elsevier.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

U. S. Commission on Civil Rights (1977). *Statement on affirmative action.* Clearinghouse Publication no. 54. Washington, DC.

Weitzman, M. L. (1974). Prices vs. quantities. *The Review of Economic Studies*, 41(4):477–491.

Yang, M. (2019). Optimality of Debt under Flexible Information Acquisition. *The Review of Economic Studies.*

A Main proofs

A.1 Proposition 1

We are using Karush-Kuhn-Tucker theorem in order to find optimal choice probabilities. The Lagrangian of the manager's problem can be written as

$$\sum_{i=1}^{N}\sum_{\omega\in\Omega} v(i|\omega)\mathcal{P}(i|\omega)\mu(\omega) - \lambda\left(-\sum_{i=1}^{N}\mathcal{P}(i)\log\mathcal{P}(i) + \sum_{i=1}^{N}\sum_{\omega\in\Omega}\mathcal{P}(i|\omega)\log\mathcal{P}(i|\omega)\mu(\omega)\right)$$
$$+\sum_{\omega\in\Omega}\xi_i(\omega)\mathcal{P}(i|\omega)\mu(\omega) - \sum_{\omega\in\Omega}\psi(\omega)(\sum_{i=1}^{N}\mathcal{P}(i|\omega)-1)\mu(\omega) - \sum_{i=1}^{N}\varphi_i(\sum_{\omega\in\Omega}\mathcal{P}(i|\omega)\mu(\omega)-q_i).$$

where $\psi(\omega), \xi_i(\omega)$ and $\varphi_i \in \mathbb{R}_+$ are Lagrange multipliers.

The first order condition with respect to $\mathcal{P}(i|\omega)$ is

$$v(i|\omega) + \xi_i(\omega) - \psi(\omega) + \lambda(\log\mathcal{P}(i) - \log\mathcal{P}(i|\omega)) - \varphi_i = 0. \tag{9}$$

Let us first show that $\mathcal{P}(i|\omega) > 0$ for all $i \in \{1, ..., N\}, \omega \in \Omega$. Suppose to the contrary that $P(i|\omega) = 0$. Then the term $-\lambda \log P(i|\omega)$ goes to infinity. The only terms which can balance it in order to make the equation (9) hold are $\psi(\omega)$ and φ_i. Thus, either $\psi(\omega)$ or φ_i goes to infinity. $P(i|\omega)$ cannot be zero for all $\omega \in \Omega$. That is, there exists a state of the world ω' in which $P(i|\omega') > 0$. In this state of the world $\xi_i(\omega') = 0$ and the first order condition is

$$v(i|\omega') + \lambda\log\mathcal{P}(i) - \lambda\log\mathcal{P}(i|\omega') - \psi(\omega') - \varphi_i = 0.$$

Therefore φ_i cannot go to infinity because there is no other term which can balance it. Thus, $\psi(\omega)$ goes to infinity. $P(i|\omega)$ cannot be zero for all $i \in \{1, ..., N\}$. That is, there exists option j such that $P(j|\omega) > 0$. The first order condition for this option is

$$v(j|\omega) + \lambda\log\mathcal{P}(j) - \lambda\log\mathcal{P}(j|\omega) - \psi(\omega) - \varphi_j = 0.$$

The last first order condition cannot hold since there is nothing to balance minus infinity of $-\psi(\omega)$. We arrived to a contradiction, therefore, $\mathcal{P}(i|\omega) > 0$ for all $i \in \{1, ..., N\}$, $\omega \in \Omega$. Hence we have $\xi_i(\omega) = 0$ and the first order condition (9) can be rearranged to:

$$\mathcal{P}(i|\omega) = \mathcal{P}(i)e^{(v(i|\omega)-\psi(\omega)-\varphi_i)/\lambda}. \tag{10}$$

Plugging (10) into (2), we obtain:

$$e^{\psi(\omega)/\lambda} = \sum_{i=1}^{N} \mathcal{P}(i)e^{(v(i|\omega)-\varphi_i)/\lambda},$$

which we again use in (10) and find:

$$\mathcal{P}(i|\omega) = \frac{\mathcal{P}(i)e^{(v(i|\omega)-\varphi_i)/\lambda}}{\sum_{j=1}^{N} \mathcal{P}(j)e^{(v(j|\omega)-\varphi_j)/\lambda}}.$$

Finally, using (4) we obtain:

$$\mathcal{P}(i|\omega) = \frac{q_i e^{(v(i|\omega)-\varphi_i)/\lambda}}{\sum_{j=1}^{N} q_j e^{(v(j|\omega)-\varphi_j)/\lambda}}. \tag{11}$$

The uniqueness of optimal conditional choice probabilities and the sufficiency follow from the fact the the maximand in the RI problem with quotas is strictly concave function, since the payoffs are linear in choice probabilities and the cost is strictly convex (Theorem 2.7.2 in Cover and Thomas (2012)).

A.2 Proposition 2

Let us assume the opposite. If the manager does not acquire information, then $\mathcal{P}(i|\omega) = q_i$ for all $i \in \{1, ..., N\}$. Or

$$\frac{q_i e^{(v(i|\omega)-\varphi_i)/\lambda}}{\sum_{j=1}^{N} q_j e^{(v(j|\omega)-\varphi_j)/\lambda}} = q_i.$$

We use the assumption that $q_i \neq 0$ and divide both parts of the equation above by q_i.

$$\frac{e^{(v(i|\omega)-\varphi_i)/\lambda}}{\sum_{j=1}^{N} q_j e^{(v(j|\omega)-\varphi_j)/\lambda}} = 1.$$

The same holds for all other options. That means that

$$v(i|\omega) - \varphi_i = v(j|\omega) - \varphi_j,$$

or

$$v(i|\omega) - v(j|\omega) = \varphi_i - \varphi_j.$$

The last equation cannot hold for all realizations of v. This is because the LHS of the above equation is state-dependent, while the RHS is state-independent, which contradicts the assumption (ii).

A.3 Proposition 4

We start the proof of the proposition by stating a lemma, which appears to be useful.

Lemma 1. *Vector* $\boldsymbol{\varphi}$, *which appears in the solution for the RI problem with subsidy* $\mathbf{q}$, *is continuous in* $\mathbf{q}$.

Proof. The idea of the proof is the following. We write an equation which provides an implicit function between $\mathbf{q}$ and $\boldsymbol{\varphi}$ and show that the conditions of implicit function theorem are satisfied, so $\boldsymbol{\varphi}$ is continuously differentiable function of $\mathbf{q}$, thus, continuous.

Without loss of generality we normalize $\sum_{i=1}^{N} \varphi_i$ to some constant c and assume that $\lambda = 1$. Let us consider the following function:

$$F(\mathbf{q}, \boldsymbol{\varphi}) = \mathbf{q} - \mathbb{E}\left[\frac{\mathbf{q} \circ e^{\mathbf{v}-\boldsymbol{\varphi}}}{\mathbf{q}^T e^{\mathbf{v}-\boldsymbol{\varphi}}}\right] + \mathbf{i}^T \boldsymbol{\varphi} \mathbf{i} - c\mathbf{i}.$$

Here "$\circ$" denotes element-wise multiplication, $\mathbf{q}^T$ is the transpose of vector

$\mathbf{q}$, $e^{\mathbf{v}-\boldsymbol{\varphi}}$ is a vector with elements $e^{v_i-\varphi_i}$, $\mathbf{i}$ is N-dimensional vector with all elements equal to 1, $\mathbf{q}^T e^{\mathbf{v}-\boldsymbol{\varphi}}$ is a scalar product. This function, if we consider the equality $F(\mathbf{q}, \boldsymbol{\varphi}) = 0$, provides an implicit function between $\mathbf{q}$ and $\boldsymbol{\varphi}$. Equality $\mathbf{q} - \mathbb{E}\left[(\mathbf{q} \circ e^{\mathbf{v}-\boldsymbol{\varphi}})/(\mathbf{q}^T e^{\mathbf{v}-\boldsymbol{\varphi}})\right] = 0$ is derived by plugging equation (5) into equation (4) and $\mathbf{i}^T \boldsymbol{\varphi} \mathbf{i} - c\mathbf{i} = 0$ is the normalization condition $\sum_{i=1}^{N} \varphi_i - c = 0$.

If we show that $\nabla_{\boldsymbol{\varphi}} F$ is invertible, we can use the implicit function theorem, thus, prove that $\mathbf{q} \to \boldsymbol{\varphi}$ is a continuously differentiable correspondence. It is easy to show that

$$\nabla_{\boldsymbol{\varphi}} F = \mathbb{E}\left[diag(Q) - QQ^T\right] + ii^T,$$

where $Q = \frac{\mathbf{q} \circ e^{\mathbf{v}-\boldsymbol{\varphi}}}{\mathbf{q}^T e^{\mathbf{v}-\boldsymbol{\varphi}}}$. $\mathbb{E}\left[diag(Q) - QQ^T\right]$ has rank $N-1$ and is positive semi-definite (PSD)[10]. Then the matrix

$$G = \mathbb{E}\left[diag(Q) - QQ^T\right] + ii^T$$

is PSD. That is so since

$$\mathbf{z}^T G \mathbf{z} = \underbrace{\mathbb{E}\left[\mathbf{z}^T (diag(Q) - QQ^T)\mathbf{z}\right]}_{\geq 0} + \underbrace{(\mathbf{z}^T i)^2}_{\geq 0} \geq 0.$$

A PSD matrix has full rank if and only if it is positive definite (PD). That is, we would like to show that $\mathbf{z}^T G \mathbf{z} \neq 0$. Let us prove it by contradiction. Let us assume the opposite: $\mathbf{z}^T A \mathbf{z} = 0$. Then $\mathbf{z}^T (diag(Q) - QQ^T)\mathbf{z} = 0$. Since the dimensionality of the kernel of $diag(Q) - QQ^T$ is one, thus the solution which has the form $\mathbf{z} = \alpha i$, $\alpha \neq 0$ is the only possible one. But in this case $(\mathbf{z}^T i)^2 = \alpha^2 (i^T i)^2 = \alpha^2 N^2 > 0$, since $\alpha \neq 0$.

Therefore, the matrix G is PD and, thus, it has full rank. We can apply implicit function theorem, and, thus, $\boldsymbol{\varphi}$ is continuous in $\mathbf{q}$. □

[10] See, for example, Fosgerau and Nielsen 2021.

Let us now continue the proof. The idea of the proof is the following. First, we consider a mapping $A : \mathbf{q} \to (\boldsymbol{\beta} \circ e^{\boldsymbol{\varphi}})/(\boldsymbol{\beta}^T e^{\boldsymbol{\varphi}})$, where $\boldsymbol{\varphi}$ is a vector from solution (5) and show that such mapping has a fixed point. Second, we show that the equation which determines the fixed point coincides with a condition which is equivalent to equation (7), thus, prove the existence of quotas $\mathbf{q}$ which induces desired choice probabilities. Throughout the proof, without loss of generality, we assume that $\lambda = 1$.

Let us consider the following mapping $A : [0, 1]^N \to [0, 1]^N$:

$$A : \mathbf{q} \to \frac{\boldsymbol{\beta} \circ e^{\boldsymbol{\varphi}}}{\boldsymbol{\beta}^T e^{\boldsymbol{\varphi}}},$$

where φ is a vector of Lagrange multipliers from equation (5).

Let us show that mapping A has a fixed point. According to Brouwer fixed point theorem, there exists a fixed point of a continuous mapping of compact convex set to itself.[11] The conditions of the theorem are satisfied, since the unit simplex in N-dimensional Euclidean space is compact and convex, $(\boldsymbol{\beta} \circ e^{\boldsymbol{\varphi}})/(\boldsymbol{\beta}^T e^{\boldsymbol{\varphi}})$ clearly belongs to unit simplex for any $\omega \in \Omega$, mapping A is continuous in $\mathbf{q}$, since $\boldsymbol{\varphi}$ is continuous in $\mathbf{q}$ (Lemma 1), and vector of generalized logit choice probabilities is continuous in $\boldsymbol{\varphi}$.

Let us show that the equation which determines the fixed point for mapping A and the equation which determines the quota from Proposition 4 are equivalent. The fixed point for mapping A is

$$A(\mathbf{q}) = \mathbf{q} = \frac{\boldsymbol{\beta} \circ e^{\boldsymbol{\varphi}}}{\boldsymbol{\beta}^T e^{\boldsymbol{\varphi}}},$$

which can be rewritten as

[11] It is important to notice that Brouwer fixed point theorem does not require the mapping to be surjective.

$$\log q_i = \log \beta_i + \varphi_i - \log(\sum_{j=1}^{N} \beta_j e^{\varphi_j}), \quad \forall i \in \{1, ..., N\}. \tag{12}$$

In order to satisfy equation (7), the vector $\mathbf{q}$ should be chosen to satisfy

$$\forall i \in \{1, ..., N\}: \quad \mathcal{P}(i|\omega) = \frac{q_i e^{v(i|\omega)-\varphi_i}}{\sum_j q_j e^{v(j|\omega)-\varphi_j}} = \frac{\beta_i e^{v(i|\omega)}}{\sum_j \beta_j e^{v(j|\omega)}}, \quad \forall \omega \in \Omega.$$

The latter equation does not change if we multiply nominator and denominator of the LHS by $\sum_{j=1}^{N} \beta_j e^{\varphi_j}$. Therefore,

$$\forall i \in \{1, ..., N\}: \quad \frac{q_i e^{v(i|\omega)-\varphi_i} \sum_{j=1}^{N} \beta_j e^{\varphi_j}}{\sum_j q_j e^{v(j|\omega)-\varphi_j} \sum_{j=1}^{N} \beta_j e^{\varphi_j}} = \frac{\beta_i e^{v(i|\omega)}}{\sum_j \beta_j e^{v(j|\omega)}}, \quad \forall \omega \in \Omega.$$

One of the possibilities of satisfaction of the last equation is

$$\forall i \in \{1, ..., N\}: \quad q_i e^{v(i|\omega)-\varphi_i} \sum_{j=1}^{N} \beta_j e^{\varphi_j} = \beta_i e^{v(i|\omega)}, \quad \forall \omega \in \Omega,$$

which implies

$$\forall i \in \{1, ..., N\}: \quad \log q_i = \log \beta_i + \varphi_i - \log(\sum_{j=1}^{N} \beta_j e^{\varphi_j}). \tag{13}$$

The latter equation coincides with equation (12) and, as we showed earlier, equation (12) has a solution. Thus, there exists a quota $\mathbf{q}$ that induces choice probabilities (7).

B Solution for the example with two candidates

In order to find conditional probabilities $\mathcal{P}(i|\omega) = \frac{q_i e^{(v(i|\omega)-\varphi_i)/\lambda}}{\sum_{j=1}^{N} q_j e^{(v(j|\omega)-\varphi_j)/\lambda}}$ we must find φ_i. According to Proposition 3 we can find a vector of subsidies that induces the same behavior and, hence, set $S_1 = \varphi_1 = C$. Therefore, the only parameter that we need to find is φ_2. Probabilities must satisfy the equation 4:

$$q = \frac{qe^{(-\varphi_2)/\lambda}}{qe^{(-\varphi_2)/\lambda} + (1-q)} b + (1-b) \frac{qe^{(1-\varphi_2)/\lambda}}{qe^{(1-\varphi_2)/\lambda} + (1-q)}.$$

Solving this equation for φ_2 and plugging it into $\mathcal{P}(2|0)$ yields two solutions:

$$\mathcal{P}(2|0) \in \{ \frac{-b - q + (b+q-1)e^{\frac{1}{\lambda}} + \sqrt{(b + q - (g+q-1)e^{\frac{1}{\lambda}})^2 + 4q(be^{\frac{1}{\lambda}} - b)}}{2(be^{\frac{1}{\lambda}} - b)}, \\ \frac{-b - q + (b+q-1)e^{\frac{1}{\lambda}} - \sqrt{(b + q - (b+q-1)e^{\frac{1}{\lambda}})^2 + 4q(be^{\frac{1}{\lambda}} - b)}}{2(be^{\frac{1}{\lambda}} - b)} \}.$$

The solution to the manager's problem should be positive. Only the first root is positive. This is so since the denominator $2(be^{\frac{1}{\lambda}} - b)$ is positive. For the root to be positive, the nominator should be positive. The second root is negative since $4q(be^{\frac{1}{\lambda}} - b)$ is positive, so the square root is larger than the term in front of the square root. For a similar reason, the first root is positive.

That is, the solution to the manager's problem is

$$\mathcal{P}(2|0) = \frac{-b - q + (b+q-1)e^{\frac{1}{\lambda}} + \sqrt{(b + q - (b+q-1)e^{\frac{1}{\lambda}})^2 + 4q(be^{\frac{1}{\lambda}} - b)}}{2(be^{\frac{1}{\lambda}} - b)}$$

and

$$\mathcal{P}(2|1) = \frac{q - b\mathcal{P}(2|0)}{1-b}.$$

C Details of the solution for the example with two candidates: subsidies

We use the example from Section 5, but now the social planner sets up a subsidy for the risky candidate: the manager receives extra payment of S if she chooses the risky candidate. In this case, the solution has the standard modified multinomial logit form but with the value of the risky candidate increased by S. Namely,

$$\mathcal{P}(2|0) = \frac{\mathcal{P}(2)e^{S/\lambda}}{\mathcal{P}(2)e^{S/\lambda} + \mathcal{P}(1)e^{C/\lambda}}$$

$$\mathcal{P}(2|1) = \frac{\mathcal{P}(2)e^{(1+S)/\lambda}}{\mathcal{P}(2)e^{(1+S)/\lambda} + \mathcal{P}(1)e^{C/\lambda}}.$$

In order to compare the manager's behavior under both policies we need to find a level of subsidies for which the risky candidate would be chosen by the manager with the required probability q:

$$(1-b)\mathcal{P}(2|1) + b\mathcal{P}(2|0) = q.$$

The unconditional probabilities in the case of the manager's problem with subsidies are as follows:

$$\mathcal{P}(2) = \max\{0, \min\{1, \frac{-e^{C/\lambda}(-e^{(1+S)/\lambda} + e^{C/\lambda} - be^{S/\lambda} + be^{(1+S)/\lambda})}{(e^{(1+S)/\lambda} - e^{C/\lambda})(-e^{S/\lambda} + e^{C/\lambda})}\}\}$$

$$\mathcal{P}(1) = 1 - \mathcal{P}(2).$$

Figure 7 shows the optimal subsidy that is necessary in order to equalize the unconditional probability of choosing the risky candidate to 0.5 as a function of b.

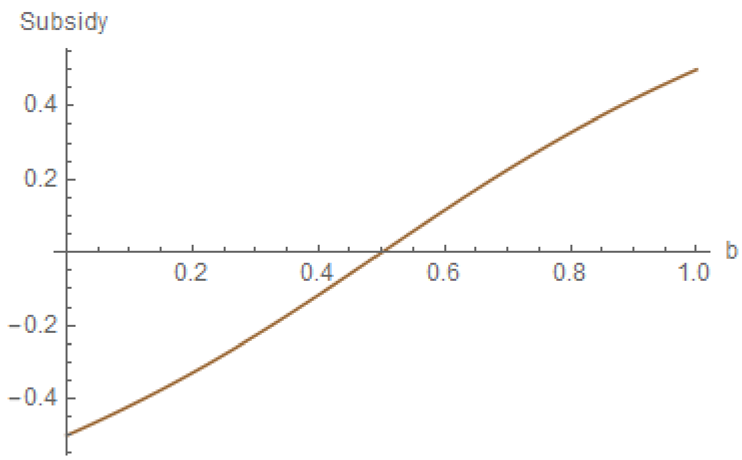

Figure 7: Optimal subsidy as a function of b and $\lambda = 0.5$, $q = 0.5$, $C = 0.5$.

We see that for small b the social planner sets a financial penalty for choosing the risky candidate. That is because the risky candidate is likely to be productive and the manager would prefer to choose it more often than in half of the cases. In contrast, if b is high, the social planner supports the choice of the risky candidate by establishing a positive subsidy.

For high b the utility of the firm in the case of subsidies is higher than in the case of quotas (Figure 8). Therefore, one can speculate that it is impossible to extract all subsidies from the firm afterwards and hence it is more beneficial for firms to lobby for subsidies rather than quotas.

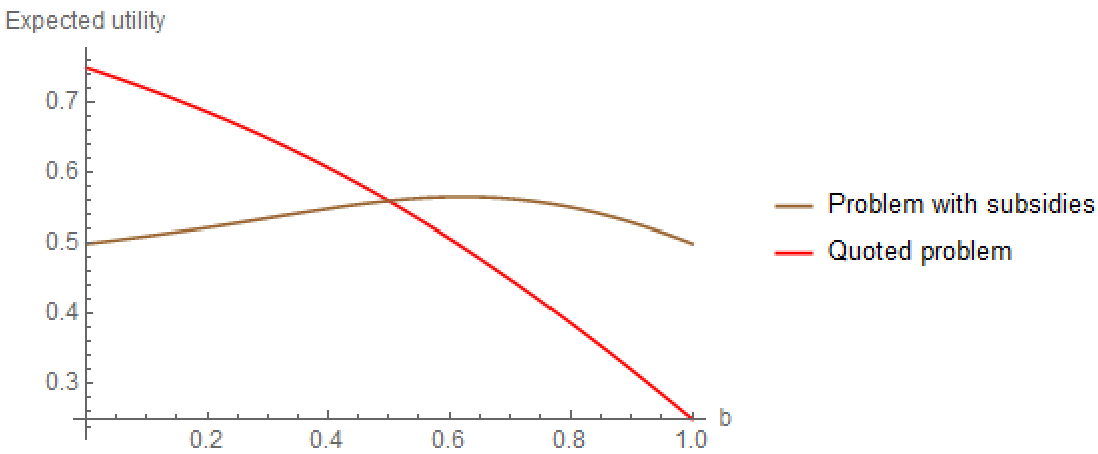

Figure 8: Utility of the manager as a function of b and $\lambda = 0.5$, $q = 0.5$, $C = 0.5$. The red curve is for the quoted RI problem and the brown curve is for the RI problem with subsidies.

D Details of the solution for the model with non-binding quotas

Let us assume that there are N alternatives and there is only one restriction on unconditional probabilities: $\mathcal{P}(1) = q$. Accordingly, this constraint implies that $\sum_{j=2}^{N} \mathcal{P}(j) = 1 - q$. Then the Lagrangian of the manager's problem described in

Section 3.2 is as follows:

$$\sum_{i=1}^{N}\sum_{\omega\in\Omega} v(i|\omega)\mathcal{P}(i|\omega)\mu(\omega)-\lambda\left(\sum_{i=1}^{N}\mathcal{P}(i)\log\mathcal{P}(i)+\sum_{i=1}^{N}\sum_{\omega\in\Omega}\mathcal{P}^U(i|\omega)\log\mathcal{P}^U(i|\omega)\mu(\omega)\right)$$
$$-\sum_{\omega\in\Omega}\psi(\omega)(\sum_{i=1}^{N}\mathcal{P}(i|\omega)-1)\mu(\omega)-\varphi_1(\sum_{\omega\in\Omega}\mathcal{P}(1|\omega)\mu(\omega)-q)-\varphi_2(\sum_{j=2}^{N}\sum_{\omega\in\Omega}\mathcal{P}(j|\omega)\mu(\omega)-1+q),$$

where $\psi(\omega)$, φ_1, and φ_2 are Lagrange multipliers. The first order condition with respect to $\mathcal{P}(1|\omega)$ is:

$$v(1|\omega)-\psi(\omega)+\lambda(\log\mathcal{P}(1)-\log\mathcal{P}(1|\omega))-\varphi_1=0,$$

and with respect to $\mathcal{P}(j|\omega)$ is:

$$v(j|\omega)-\psi(\omega)+\lambda(\log\mathcal{P}(j)-\log\mathcal{P}(j|\omega))-\varphi_2=0.$$

Following the same procedure described in Section 3.2 this can be rearranged to:

$$\mathcal{P}(1|\omega)=\frac{qe^{(v(1|\omega)-\varphi_1)/\lambda}}{\sum_{j=2}^{N}\mathcal{P}(j)e^{(v(j|\omega)-\varphi_2)/\lambda}+qe^{(v(1|\omega)-\varphi_1)/\lambda}},$$

and

$$\mathcal{P}(j|\omega)=\frac{\mathcal{P}(j)e^{(v(j|\omega)-\varphi_2)/\lambda}}{\sum_{j=2}^{N}\mathcal{P}(j)e^{(v(j|\omega)-\varphi_2)/\lambda}+qe^{(v(1|\omega)-\varphi_1)/\lambda}}.$$

Therefore, the solution to the problem will be similar to that described in Section 3.2. The only difference is that now, for all alternatives for which the quota is not binding and for which $\mathcal{P}(j)>0$, the additive state-independent component φ_2 is the same. This logic extends to any situation in which not all quotas are binding.