Optimal Patronage*

Mikhail Drugov†

August 27, 2018

Abstract

This paper studies the design of promotions in an organization where agents belong to different groups. Examples and applications include political groups, ethnicities, agents motivated by the work in the public sector, etc. In an overlapping generations model, juniors compete for promotions. Seniors have two kinds of discretion: direct discretion, which is immediately beneficial for their group, and promotion discretion ("patronage"), which is a bias in the promotion decision. Under two alternative objectives of the planner considered - maximizing juniors' efforts and affecting the steady-state composition of the seniors towards the preferred group - optimal patronage may be strictly positive.

Keywords: motivated agents, contest, promotion, patronage, bureaucracy.

JEL codes: D73, J70, J45, H41.

*I am particularly grateful to Margaret Meyer for many insightful conversations. I also benefited from useful comments of Chaim Fershtman, Guido Friebel, Martina Kirchberger, Rachel Kranton, Anandi Mani, Marta Troya Martinez, Debraj Ray, Silvia Sonderegger, Giancarlo Spagnolo, Anton Suvorov and participants of various seminars and conferences. Any remaining errors are my own.

†New Economic School and CEPR; mdrugov@nes.ru.

1 Introduction

This paper is based on a simple observation that people belong to different groups, and they care about the group to which they belong. The groups can be exogenous as in the case of ethnicities, tribes, castes and, in most cases, religions. They may also be endogenous and based on values, for example, political parties or political factions. Even if the identity is created in the lab it matters. As Tajfel and Turner (1986), the two major figures in social psychology surveying the enormous literature on the issue, wrote "The basic and highly reliable finding is that the trivial, ad hoc intergroup categorization leads to in-group favoritism and discrimination against the out-group."

The main question of this paper is the following: what implications arise for the organizational design when agents belong to and care about their groups? In particular, can we rationally explain some seemingly welfare detrimental phenomena such as patronage? By patronage we mean unfair promotions for which group identity is taken into account rather than only ability or performance. The main result of the paper is that even if the goals of the organization are group-neutral, for example, to maximize the efforts or output of the workers, allowing for some patronage might be optimal. We also study the effectiveness of patronage when one group is preferred to the other in which case the composition of the organization matters.

While patronage occurs in private firms too, we mainly have in mind the design of bureaucracies where agents from different groups inevitably work together and where patronage provokes most public outcry. Indeed, governments usually formally and explicitly do not allow for discrimination, while in reality this is not the case in many countries, especially developing countries.

In an overlapping generations model agents live for two periods. When young, agents work in the organization at junior level. Some will be promoted to senior level and work there when old. Promotions are based on the contest between junior agents, but this contest may be biased. The organizational designer, who we refer to as the planner, may give senior agents the possibility to bias the contest in favour of the juniors they prefer. This is for example the case when the promotions are partially based on the subjective reports of the superiors. As Prendergast and Topel (1996) put it: "...subjectivity opens the door to favoritism, where evaluators use their power to reward preferred subordinates beyond their true performance" (p. 958). When this happens, we say that there is *patronage*.

Agents belong to two different groups and care about the welfare of their group. Senior agents use their discretionary power to contribute to their group welfare in two ways. First, they have *direct discretion*, that is, they can directly increase their group welfare. For example, they can channel public funds towards regions populated by their tribe or they can make public statements and make some decisions that promote their political values. Thus, senior agents prefer to promote juniors of their group because, when they become seniors, they will benefit their group. The second kind of discretion is *promotion discretion*, or patronage as described above. Thus, in our model patronage is valued only when there is direct discretion.

We consider two possible goals of the planner. First, the planner is group-neutral and his goal is to maximize the efforts of the junior agents either because their efforts are productive or, in the case of training, because their efforts increase their ability when they become seniors. When juniors from the two groups compete for promotion, the identity of the winner matters because the promoted junior, becoming senior, will benefit his group. The attractiveness of the senior position increases with both the direct discretion and patronage.

The trade-off faced by the planner is the following: a higher patronage means that the contest for promotion is more biased and, since the juniors are symmetric (except for their group identity), this implies a lower effort; we call this the *discouragement effect*. However, a higher patronage makes the senior position more attractive, and therefore, increases the juniors' efforts; this is the *higher stakes effect*. We find that, when direct discretion is neither too large nor too small, the juniors' efforts are maximized by a strictly positive patronage. In other words, even though the planner can make all the promotions merit-based, he chooses to give senior agents the power to bias them as they please. We also show that in general direct discretion and patronage are neither complements nor substitutes, that is, a higher direct discretion has an ambiguous effect on the optimal patronage. The reason is that both the higher stakes and the discouragement effects increase with the direct discretion.

We then turn to the second possible goal of the planner. The planner might prefer one group to another. For example, the planner is a politician who cares about the preferences of the median voter who is likely to belong to the larger group. Alternatively, the direct discretion may be costly for the planner per se in which case he prefers the group which uses it in a less distortionary way. Suppose that the only goal of the planner is to bias the steady-state composition of the senior level towards his preferred group. There are three effects of patronage on the steady-state composition of the senior level: first, it benefits the larger group because it is more likely to use the patronage; second, it benefits the less motivated group since this group is likely to lose in the fair contest; and third, it changes the values of promotion for the two groups because they increase with patronage, and this effect can go either way. We give conditions when the optimal patronage is zero and when it is maximal, that is, there is effectively no contest between the juniors since senior bureaucrats always promote their preferred junior.

Related literature

There are many papers documenting the existence of the group bias in the decisions of bureaucrats and politicians in various developing countries, see Burgess et al. (2015), Do, Nguyen and Tran (2017), Franck and Rainer (2012), Hodler and Raschky (2014), Iyer and Mani (2012), Kramon and Posner (2016), Marx, Stoker and Suri (2017) and Neggers (2018) for the most recent (econometric) evidence and references there. While sometimes the bias can be explained by the electoral concerns, Do, Nguyen and Tran (2017), Marx, Stoker and Suri (2017) and Neggers (2018) find the evidence of favoritism exerted by low-level bureaucrats who do not face any electoral pressure. In politics, there is a distinction between "factions of principle" based on values and "factions of interest" organized for their own power (Bettcher (2005)); our analysis mainly applies to the former ones. See Persico, Pueblita and Silverman (2011) for a model of the latter ones and Huang (2000) and Shih (2009) for a fascinating analysis of factional politics in China. For the evidence from the lab (in economics), see Chen and Li (2009) and references there.

In terms of the underlying economic forces, this paper is related to several strands of literature. In Athey, Avery and Zemsky (2000), Fryer and Loury (2005) and Morgan, Sisak and Várdy (2012), the planner biases the contests for promotion to reach some further goals, such as promoting more able agents in the first case, diversity in the second case and attracting talent to the organization in the last case. In other words, the planner affects the composition of the organization in the direction he prefers as in this paper when the planner cares about the composition of the senior level. In those papers, it is still the planner who administers the biased contest, while in our model the senior agents use the biased contest to promote the juniors they like.

Meyer (1992) studies a two-period contest between identical agents. Introducing a small additive bias in a Lazear-Rosen tournament has only a second-order effect on efforts.[1] If it is introduced in the second period to reward the winner of the first period, it has a first-order effect on first-period incentives, and therefore, it is optimal to introduce some bias in the second period. In our terms, the discouragement effect is of the second order, while the higher stakes effect is of the first order. We do not rely on this logic since we introduce patronage as the probability that the senior

[1] This is a very general result which holds far beyond the Lazear-Rosen tournament and additive bias, see Drugov and Ryvkin (2017) for details.

completely decides on promotion, in which case the discouragement effect is always of the first order. In Appendix B, however, we consider a standard setup of a Tullock contest with a multiplicative bias as in Epstein, Mealem and Nitzan (2011) and Franke et al. (2013) in which the discouragement effect is of the second order. This fact is useful in showing that optimal patronage is positive even when the costs of public funds are low, and therefore, providing monetary incentives is cheap.

In Ghatak, Morelli and Sjöström (2001), credit market imperfections make current borrowers worse off. However, they increase incentives to work hard and self-finance since the rents to self-financed entrepreneurs also increase. Therefore, reduction in credit market imperfections may reduce welfare. Thus, there is the same very general idea that a certain distortion has some current negative effects but also provides more incentives through higher future rents.

The agents in our model are pure altruists in the sense that they care about their group welfare but not how it is achieved. A few papers, such as Francois (2000), Francois (2007) and Engers and Gans (1998), have considered implications of such agents for organizational design. However, none of these papers is concerned with the promotion policy. In models where agents have public sector motivation, such as Besley and Ghatak (2005), Delfgaauw and Dur (2008), Macchiavello (2008) and Delfgaauw and Dur (2010) agents have a "warm glow" motive. They value their contribution to the welfare irrespective of what happens if they do not contribute. We can easily incorporate the "warm glow" into our model (it is equivalent to a higher senior wage). We also consider an intermediate case in which the agents discount their effect on the group welfare depending on how far their action is from the eventual increase in their group welfare. This can be seen as a generalization of impure altruism, see Andreoni (2006) for the definitions and discussion.

Prendergast and Topel (1996) consider an agency model where a supervisor intrinsically cares about his junior being promoted and biases his evaluation report to the principal. The model and the questions there are very different from the ones in this paper, but the same broad lesson emerges. While favoritism creates distortions, completely eliminating it might not be optimal since the agents value exercising it. In Prendergast and Topel (1996) they then agree to a lower wage while in our model they work harder.

As one interpretation of the group welfare is the status of its members, this paper is also related to the small literature on the role of status for incentives, including Auriol and Renault (2001, 2008) and Besley and Ghatak (2008).

Finally, from the modelling perspective, using an overlapping generations model to study organizations has been used in the past. For example, it is used in Ghatak, Morelli and Sjöström (2001) described above. In Meyer (1994), the organization decides how to organize teams in order to learn the most about the workers' abilities. In Carrillo (2000), the focus is on fighting corruption with various tools (but not patronage).

The rest of the paper is organized as follows. The model is introduced in Section 2. In Section 3 the optimal patronage is characterized when the planner cares about juniors' efforts. In Section 4 the planner cares about the steady-state composition of the senior level. A few extensions are analyzed in Section 5. Section 6 concludes. Appendix A contains the proofs. Appendix B considers the standard Tullock contest which generates similar results.

2 Model

In an overlapping generations model each agent lives for two periods. While young, agents work in the organization, which we call a bureaucracy, at the junior level. Some of them will be promoted to the senior level and work there when old. The bureaucracy is organized in departments, each consisting of two junior bureaucrats and one senior bureaucrat. Every period the senior bureaucrat retires and one (and only one) junior of his department is promoted to replace him.[2] The senior bureaucrat gets wage w and some discretionary power that we explain below. The junior who is not promoted either leaves the bureaucracy or stays there in some low-level position with no discretionary power and gets utility normalized to 0.

2.1 Types and utilities of agents

There are two groups, left (l) and right (r), and each agent belongs to one of them. The type of an agent is the group to which he belongs. The probability that a junior is of type l is λ. The composition of the departments is random, that is, the types of juniors are independent.[3] The type of agent matters because agents care about the welfare of their group. That is, the agents' utility has two components: the standard "private" part that depends on their wage and effort costs and an "altruistic" part that depends on the welfare of their group.

[2] It does not matter if the promoted junior stays in the same department.

[3] We discuss the preferences of the planner over the composition of the junior level in Section 3.4.

2.2 Seniors' discretion and group welfare

The discretionary power of the senior bureaucrats takes two forms. First, they can directly benefit their group by amount $d \geq 0$; we call this *direct discretion*. For example, they administer some funds and can disburse them to the members of their group. Or, they can choose to implement public projects in ways that benefit their group.[4] If the group identity is based on ideology rather than ethnicity, senior bureaucrats can effectively promote their values among the general public since they are highly visible. If the senior position confers status, senior bureaucrats benefit their group by increasing the average status of their group members.

The second form of the seniors' discretionary power is *promotion discretion* or *patronage*. Senior bureaucrats administer the promotion of the juniors in the department and they can bias it in favour of the junior from their group. The size of the promotion discretion is the focus of this paper. Even if it is possible to eliminate all promotion discretion and make promotions entirely merit-based, the planner may not find it optimal. We formalize promotion discretion in the simplest way: with probability p a senior bureaucrat has complete discretion about which junior from his department to promote, while with probability $1 - p$ the promotion is entirely merit-based.[5]

The welfare of each group is equal to the (discounted) sum of the direct discretions exerted by its seniors, $W_i = d \sum_{t=0}^{+\infty} \delta^t N_i^t$, $i = l, r$, where δ is the discount factor and N_i^t is the number of seniors of group i in period t.[6,7] Note that patronage increases the group welfare only indirectly. A group benefits from its juniors being promoted because they will use their direct discretion when senior (and also promote juniors of the group in the future who will benefit the group when senior, etc.).

[4] In the U.S. terminology, d can be seen as an *earmark* which is a provision that allocates funds to a specific recipient or a project.

[5] We discuss different ways of biasing the contest at the end of Section 3.3 and analyze two different contest models in Appendix B. Introducing the bias in this way makes it more difficult to obtain a positive optimal patronage as compared to the standard additive or multiplicative handicaps for one of the players.

We also restrict attention to constant patronage policies. Studying dynamics of patronage is an exciting avenue for future work.

[6] In some cases the welfare of each group may decrease with the direct discretion used by the seniors of the other group. For example, agents may care about the relative income or status of their group. Promoting your values is harder when other people promote different (or opposite) values. See Section 5.2 for such an extension.

[7] The group welfare does not include the "private" part of the agents' utilities, that is, their wages and effort costs. This is done so that the different interpretations of group welfare (income, values, status) map into exactly the same model. Also, in the case of income, one can assume that the direct discretion d is much larger than the wage w and omitting w (and effort costs) does not greatly affect the results. Modifying the model to include the "private" part into group welfare is straightforward.

2.3 Promotion contest

When the promotion is merit-based, the two juniors of the department engage in the contest by exerting effort equal to 0 or 1. If a junior exerts effort 1, he generates a high output, while exerting effort 0 results in a low output. The junior with a higher output is promoted; in the case of equal outputs each junior is promoted with probability $\frac{1}{2}$. The cost of effort 1 is $\frac{c}{2}$ (and 0 for effort 0) and juniors differ in the cost parameter which is distributed according to some distribution F on $[\underline{c}, \overline{c}]$, and are privately informed about it.

3 Maximizing juniors' efforts

In this section, the planner maximizes the (expected) output at the junior level and therefore chooses the promotion discretion p to maximize the juniors' efforts. Interpreting the model literally, the senior bureaucrats do not exert any effort since they will be retiring afterwards. Alternatively, their effort may be subject to another (unmodeled) moral hazard problem and is independent of the direct discretion and promotion discretion which are the focus of this paper.

We now solve the model and find the optimal patronage. Set $\delta = 1$. While this makes the welfare of both groups infinite, what matters for the decisions of the agents is the impact they make on the group welfare, which is always finite. The case of $\delta < 1$ is straightforward, see the working paper Drugov (2015).[8]

The first step is to solve the promotion contest. There are two cases depending on whether the two juniors in a department belong to the same group. We call the first case the "homogeneous department" and the second case the "heterogeneous department".

3.1 The contest in a homogeneous department

When both juniors belong to the same group, the welfare of their group does not depend on who gets promoted. The value of the promotion for each of them is only the senior's wage w. The senior bureaucrat does not use his promotion discretion, as he cannot change the group of the promoted junior.

[8] We assume that the agents care about the welfare of their group up to the infinite horizon. This is done only for analytical tractability - qualitative insights survive when they care only about a few future periods.

A junior with cost parameter c exerts an effort if and only if

$$\left(\frac{1}{2}F(\widehat{c})+1-F(\widehat{c})\right)w-\frac{c}{2}\geq\frac{1}{2}(1-F(\widehat{c}))\,w, \tag{1}$$

where $\widehat{c}$ is the cost threshold of the other junior. Simplifying this inequality gives rise to the following Lemma.

Lemma 1 *In a homogeneous department a junior exerts an effort if and only if $c \leq w$, that is, with probability $F(w)$.*

Note that $\widehat{c}$ does not matter. By exerting an effort a junior increases his promotion probability by $\frac{1}{2}$ independent of what the other junior is doing. Indeed, if the other junior does not exert an effort, exerting an effort changes the promotion probability from $\frac{1}{2}$ to 1. If he exerts an effort, exerting an effort changes the promotion probability from 0 to $\frac{1}{2}$.

3.2 The contest in a heterogeneous department

In a heterogenous department, the two juniors belong to different groups. Then, being promoted not only results in the senior wage w but also impacts the group welfare. Indeed, a senior bureaucrat increases the welfare of his group by d directly and by ΔW^f from possibly biasing future promotion. The latter occurs in a heterogenous department and with probability p and, when it occurs, the group welfare changes by $d+\Delta W^f$. Solving the equation

$$\Delta W^f = 2\lambda(1-\lambda)\,p\left(d+\Delta W^f\right)$$

yields the total impact on the group welfare, $d+\Delta W^f = \frac{d}{1-2\lambda(1-\lambda)p}$.

Suppose that juniors know when the patronage will be used in which case they do not exert any effort.[9] When the patronage is not used, the contest is merit-based and, writing the condition for exerting an effort similar to (1), gives the following Lemma.

Lemma 2 *In a heterogeneous department when patronage is not used, a junior exerts an effort if and only if $c \leq w+\frac{d}{1-2\lambda(1-\lambda)p}$, that is, with probability $F\left(w+\frac{d}{1-2\lambda(1-\lambda)p}\right)$.*

When the contest is merit-based, the juniors exert a higher effort than in a homogenous department, and this effort is increasing in the size of patronage p.

[9] Making the opposite assumption does not change the results qualitatively. See also the discussion at the end of Section 3.3 on the different ways of introducing the bias.

3.3 Characterizing the optimal patronage

Denote $q = 2\lambda(1-\lambda)$, the probability of having a heterogenous department. Using Lemmas 1 and 2 we can now write the planner's problem of maximizing the total effort as

$$E = (1-q)\,F(w) + q(1-p)\,F\left(w + \frac{d}{1-qp}\right) \to \max_{p\in[0,1]} \tag{2}$$

Promotion discretion has two opposite effects on the total effort (2). First, there is a *higher stakes effect*: promotion becomes more valuable since senior bureaucrats have more say in future promotions. Second, there is a *discouragement effect*: there is no effort when the senior promotes the junior of his group for certain. Comparing these effects at $p = 0$ leads to the following proposition.

Proposition 1 *If $f(w+d)\,qd > F(w+d)$, the optimal patronage p^* is strictly positive and is found from the condition*

$$(1-p)\,f\left(w + \frac{d}{1-qp}\right)\frac{qd}{(1-qp)^2} = F\left(w + \frac{d}{1-qp}\right) \tag{3}$$

When $p = 0$ and conditional on being in a heterogeneous department the value of the promotion is $w + d$. The higher stakes effect is then equal to $f(w+d)\,qd$, that is, the probability of the junior marginal type times the increase in the value of promotion. The discouragement effect is equal to $F(w+d)$ since each junior provides effort with probability $F(w+d)$ in a merit-based contest. When the higher stakes effect dominates the discouragement effect at zero, the optimal patronage is positive. It is then given by (3), which is the first-order condition for problem (2), that the two effects are equal. Since no effort is exerted in a heterogeneous department when $p = 1$, optimal patronage is always strictly below 1.

Next, we present a particularly well-behaved example.

Example 1 *Suppose that c is distributed uniformly on $[\underline{c}, c]$ and $\underline{c} > w$.*[10] *Optimal patronage p^* is 0, if $d \le (\underline{c} - w)(1-q)$ or $d \ge \frac{\underline{c}-w}{1-q}$, and otherwise it is*

$$p^* = \frac{1}{q}\left(1 - \sqrt{d\frac{1-q}{\underline{c}-w}}\right). \tag{4}$$

[10] Condition $\underline{c} > w$ may seem restrictive. However, since the utility of the non-promoted juniors is normalized to zero, senior wage w is in fact the difference between the wages of promoted and non-promoted juniors. In many developing countries public servants, including senior ones, are badly paid and the benefits of the job come mainly from the power associated with it.

The discouragement effect dominates when the direct discretion d is either small or large. When it is small, patronage does not increase the value of promotion by a lot. When it is large, the value of promotion with no patronage, $w + d$, is already large enough to incentivize all or almost all juniors, and there is not much to gain from increasing this value further, while the loss due to discouraging effort is large.

Patronage is not used if direct discretion is either too small or too large. Thus, overall, the two kinds of discretion are neither substitutes nor complements. Optimal patronage in (4) decreases with d. In general, a higher promotion discretion always increases the discouragement effect and increases the higher stakes effect if $f' > 0$. See Figure 1 for an example of where the optimal patronage first increases with d and then decreases while being strictly positive.

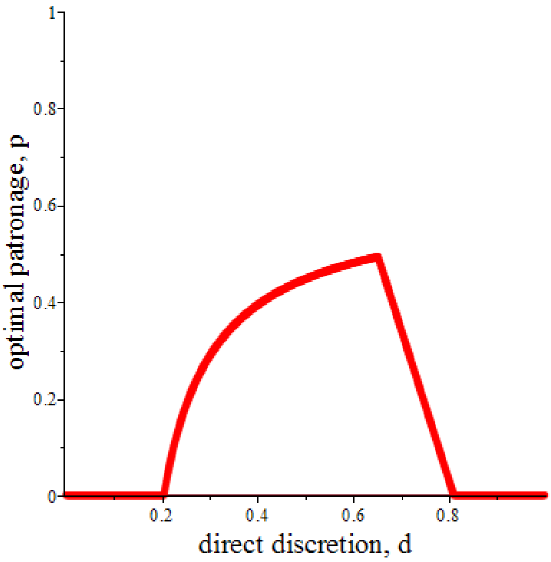

Figure 1: Optimal patronage when the costs are distributed as $Beta\,(5,1)$ $(F\,(c) = c^4)$ and $q = 0.4$, $w = 0.2$.

The comparative statics of the optimal patronage with respect to other parameters also depends on the cost probability density function f and its derivative f'. The optimal patronage decreases with wage w if $f' \leq 0$. The effect of the probability of the heterogenous department q is more ambiguous. At zero patronage, q only increases the higher stakes effect and hence makes a stronger case for a strictly positive patronage. In general, however, both discouragement and higher stakes effects increase with q. In Example 1 the effect is of inverted U-shape: optimal patronage first increases with q and then decreases.

Finally, let us comment on different ways of biasing the contest for promotion and the resulting discouragement effect. Introducing patronage as a probability that the efforts do not matter means that the discouragement effect is always of the first order. This is true for both when the juniors know if patronage will be used, as we assume throughout the paper, and when they do not, and therefore, exert effort that probably will not matter. Introducing the bias in a more standard way as is done in the contest literature makes the discouragement effect of the second order at zero bias.[11] Since the higher stakes effect is always of the first order, optimal patronage is then strictly positive for any positive direct discretion. In Appendix B we consider the Tullock contest with the multiplicative bias and show that the optimal patronage $p^* > 0$ for any $d > 0$ (see Proposition 4). To summarize this discussion, introducing patronage as we do in this paper makes it *more* difficult to obtain a strictly positive optimal patronage.

3.4 Optimal composition of the departments

Whenever group identity is observable, which is the case of groups based on ethnicity, caste, religion, etc., two related questions arise. First, should the planner make departments homogenous or heterogenous? In other words, is diversity good for performance? Second, what is the optimal composition of the junior level, i.e., λ?

Next proposition answers these questions.

Proposition 2 *The optimal composition of the junior level is balanced, that is, $\lambda = \frac{1}{2}$, and all the departments are heterogenous.*

The efforts are strictly higher in a heterogenous department since the planner can always set the patronage to zero, $p = 0$, in which case the juniors always compete and have higher incentives than in the homogenous department (see Lemmas 1 and 2). Thus, the planner composes heterogenous departments whenever possible, that is, he sets $q = 2\min\{\lambda, 1-\lambda\}$. The optimal composition of the junior level is then to have $\lambda = \frac{1}{2}$.

At the macro level diversity is typically associated with negative outcomes such as lower growth, worse policies, or a higher likelihood of a conflict (see, e.g. Alesina and La Ferrara (2005) and Easterly and Levine (1997)). At the micro level, however, the findings are more nuanced, see Shore et al. (2009) for a survey of the evidence from private firms. The evidence from bureaucracies is very limited. In the context of Nigeria, a highly ethnically fractionalized country, Rasul and Rogger (2015) find a positive impact of diversity on bureaucratic performance.

[11] See Meyer (1992) for an early example of an additive bias in a Lazear-Rosen tournament, Franke et al. (2013) for optimal multiplicative bias in the Tullock contest and Drugov and Ryvkin (2017) for a general analysis of biased contests with symmetric players.

4 Affecting the senior level

We now turn to a scenario which is in some ways opposite to the one in Section 3 and in which the planner cares only about the composition of the senior level.[12] For example, the planner is a politician who cares about the preferences of the median voter who is likely to belong to the larger group. Alternatively, the direct discretion may be costly for the planner per se, in which case he prefers the group which uses it in a less distortionary way. The planner's trade-off with respect to the patronage is simple: the seniors from his preferred group use it as he would like while the seniors from the other group use it against his preferences. Hence, the aim of this section is to understand which effect is likely to dominate and under which circumstances.

As we will see, the effect of patronage depends on how relatively motivated the two groups are. Thus, we allow for the direct discretion to be different between the two groups, d_l and d_r. For example, diverting funds of a given size is more valuable for a poorer group. Alternatively, exerting the direct discretion may be costly for the agents if they need to exert an effort or can be caught, and groups differ in how much the agents are motivated.

Suppose that the planner prefers the left group to the right one, and therefore, maximizes the steady-state share of left seniors, λ^S. It is found from the equation[13]

$$\lambda^S = \lambda^2 + 2\lambda(1-\lambda)\left[p\lambda^S + (1-p)\frac{1}{2}(1+F_l-F_r)\right], \tag{5}$$

where $F_i = F\left(w + \frac{d_i}{1-2\lambda(1-\lambda)p}\right)$, $i = l, r$. In what follows, we will sometimes refer to $\frac{d_i}{1-2\lambda(1-\lambda)p}$ as the *motivation* of group i. The left seniors come from 1) homogenous departments where both juniors are left (the first term on the right-hand side of (5)), 2) heterogenous departments headed by a left senior who uses promotion discretion (the first term in the square brackets in (5)), and 3) heterogenous departments where promotion is merit-based and the left junior wins it (the second term in the square brackets in (5)).

The patronage affects λ^S via three channels listed in the next Lemma.

[12]He then probably cares about the overall composition of the bureaucracy, but it is assumed that the composition of the junior level cannot be manipulated. For example, in many cases there are civil service exams to enter bureaucracy. Also, the group identity may not be observable at the entry stage, as in the case of groups based on values.

[13]When the contest is merit-based, the probability that a left junior is promoted in the heterogenous deparment is equal to

$$\frac{1}{2}(F_lF_r + (1-F_l)(1-F_r)) + F_l(1-F_r) = \frac{1}{2}(1+F_l-F_r).$$

Lemma 3 *Patronage p affects the steady-state composition of the senior level via three effects: 1) by benefiting the larger group; 2) by benefiting the less motivated group and 3) by changing the difference in shares of juniors that exert the effort, $F_l - F_r$.*

The first effect is the size effect which has the sign of $\lambda - \frac{1}{2}$: the promotion discretion benefits the larger group because it is more likely to use it. The second and the third effects arise because patronage changes the likely winner of the fair contest. The second effect is the relative motivation effect proportional to $F_r - F_l$: the patronage benefits the less motivated group because on average this group loses in the fair contest. Finally, the third effect is the change in the relative motivation, proportional to $\frac{\partial(F_r - F_l)}{\partial p}$ since the patronage changes the motivations. The sign of this effect depends on the cost distribution F and group motivations. An immediate corollary of Lemma 3 is that when direct discretions are equal, $d_l = d_r$, the optimal patronage is determined by the relative size of the two groups: if the left group is bigger (smaller), the optimal patronage is maximal, $p^* = 1$ (minimal, $p^* = 0$).

Expressing λ^S from (5) yields

$$\lambda^S = \frac{\lambda}{1 - 2\lambda(1-\lambda)p}\left[\lambda + (1-\lambda)(1-p)(1 + F_l - F_r)\right]. \quad (6)$$

The planner maximizes (6) by choosing promotion discretion $p \in [0, 1]$. Next proposition gives sufficient conditions for an extreme policy, $p^* = 0$ or $p^* = 1$, to be optimal.

Proposition 3 *When the planner maximizes the steady-state share of left seniors (6),*

(i) Optimal patronage is maximal, $p^* = 1$, if $\lambda \geq \frac{1}{2}$, $d_r \geq d_l$ and $\frac{d_r}{d_l} \geq \frac{f_l}{f_r}$.

(ii) Optimal patronage is minimal, $p^* = 0$, if $\lambda \leq \frac{1}{2}$, $d_r \leq d_l$ and $\frac{d_r}{d_l} \leq \frac{f_l}{f_r}$.

The three conditions listed in parts (i) and (ii) of Proposition 3 correspond to the three effects of the patronage on the steady-state share of left seniors in Lemma 3. When $\lambda \geq \frac{1}{2}$, the left group is bigger. When $d_r \geq d_l$, the left group is less motivated. When $\frac{d_r}{d_l} \geq \frac{f_l}{f_r}$, the third effect of the patronage - the change in relative motivations - is also positive. It is satisfied when the density f is decreasing slowly enough or increasing, that is, there are less people with lower costs. Even if the cost distribution is initially interior unimodal like the normal one, civil service exams present in many countries may select only the left tail of sufficiently talented or educated people who then join the bureaucracy; in this case f is increasing. If smaller groups are more tight-knit, this means that there is a positive correlation between $\lambda - \frac{1}{2}$ and $d_r - d_l$.

Conditions in part (ii) of Proposition 3 are reversed and make sure that all the three effects of the patronage are negative. When these effects are in the opposite directions, a general characterization is difficult. We then proceed with the case of the uniform cost distribution started in Example 1.

Example 2 *Suppose that c is distributed uniformly on $[w, w+1]$ and $d_i \leq \frac{1}{2}$, $i = l, r$.*[14] *When the planner maximizes the steady-state share of left seniors, the optimal patronage is*

- *Maximum, $p^* = 1$, if $d_r - d_l \geq \max\{1 - 2\lambda, \frac{1-2\lambda}{1-2\lambda(1-\lambda)}\}$;*
- *Intermediate, $p^* = \frac{2\lambda-1}{2\lambda(1-\lambda)} \frac{1+(2\lambda-1)(d_r-d_l)}{2\lambda-1-(d_r-d_l)}$ if $\lambda > \frac{1}{2}$ and $d_r - d_l < 1 - 2\lambda$;*
- *No patronage, $p^* = 0$, otherwise.*

Proof. See Appendix A. ■

In the uniform case the relative motivation $F_r - F_l$ is proportional to the difference in direct discretions, $d_r - d_l$. Then, both motivation effects of patronage mentioned in Lemma 3, of relative motivation and of the change in the relative motivation, are proportional to $d_r - d_l$; they can be jointly labelled as the motivation effect. Therefore, there are only two parameters in the planner's problem, λ and $d_r - d_l$, which simplifies the characterization of the optimal patronage.

Example 2 is illustrated in Figure 2. Consider the upper right quadrant. The left group is larger, $\lambda > \frac{1}{2}$, and less motivated, $d_l < d_r$, that is, both the size and motivation effects of a higher patronage are positive. The optimal patronage is then maximum, $p^* = 1$. The lower left quadrant in Figure 2 is the opposite case: the left group is smaller and more motivated. A higher patronage decreases λ^S via both effects and it is optimal to set patronage to zero, $p^* = 0$.

The two effects are opposed in the other two quadrants. In the lower right quadrant the left group is larger, $\lambda > \frac{1}{2}$, but also more motivated, $d_l > d_r$. When the motivations are close, the first effect dominates and optimal patronage is at the maximum, $p^* = 1$. As the gap in motivations increases, the second effect becomes more important and the optimal patronage becomes less than maximum and then further

[14] These assumptions imply that $w + \frac{d_i}{1-2\lambda(1-\lambda)p} \in [w, w+1]$ for any λ and p which is the most interesting case. The unit length of the support is a normalization.

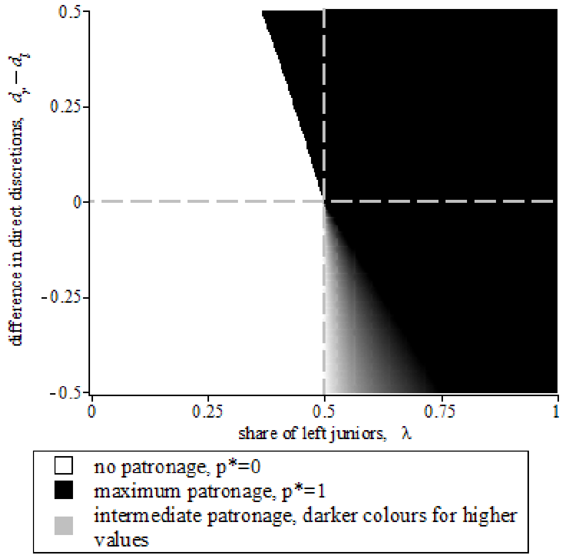

Figure 2: Optimal patronage depending on the share of left juniors, λ, and the difference in direct discretions, $d_r - d_l$.

decreases. Increasing λ makes the larger left group even larger, and therefore, the optimal patronage increases. In the opposite, upper left quadrant the two effects are reversed: now the left group is smaller, $\lambda < \frac{1}{2}$, but also less motivated, $d_r > d_l$. However, in this case λ^S is U-shaped in patronage and therefore the optimal patronage is either zero or the maximum one.

The comparative statics just discussed leads to the following corollary.

Corollary 1 *In Example 2 optimal patronage p^* (weakly) increases with the share of left juniors, λ, and with the difference in direct discretions, $d_r - d_l$.*

Let us conclude this section by coming back to the planner's trade-off: discretion in promotions is used "well" by some agents and is "abused" by others, both from the planner's perspective. This is a general trade-off when deciding on how much discretion to give to bureaucrats. In the procurement context, discretion can be used to create dynamic incentives, that is, to reward good performance by future contracts. It can also be abused by giving contracts to "friends". In the U.S. following the work of Steven Kelman (Kelman, 1990) and his efforts when working for the Office of Federal Procurement Policy in the early 1990s there has been a significant increase in flexibility and discretion in procurement. In the European Union there is an active debate on this topic, see Coviello, Guglielmo and Spagnolo (2018) for evidence from Italy and further references. In the regulation context, discretion can be used to target more problematic firms increasing the effectiveness of the regulation or it can be used to skip "friends" and target other firms to extract bribes, see Duflo et al.

(forthcoming) for a field experiment on environmental inspections in India. In all these cases, the key question is when the positive effect of more discretion dominates the negative one.

5 Extensions

The model we study is very flexible and allows for a number of extensions. Here we present two extensions. First, we consider the case when the agents are impure altruists. Second, we allow agents to care differentially about their group and to value the relative welfare. In the working paper Drugov (2015) we also studied the planner caring about both goals - maximizing juniors' efforts and affecting the senior level - simultaneously. In Appendix B we consider the standard Tullock contest success function (Tullock, 1980) and show the robustness of the patronage to the monetary incentives.

5.1 Warm glow and impure altruism

People often value their own contribution to a public good irrespective of what others do or would do if they do not contribute. This is called "warm glow" (see Andreoni (2006)). Impure altruists combine pure altruism (that is, the total amount of the public good enters the utility function) and warm-glow motivation (that is, their contribution directly enters the utility function). Introducing the warm glow or impure altruism in our model is straightforward: the own direct discretion d has a positive weight in the agents' utility function. Hence, it is equivalent to increasing the senior wage w.

Another question arises, however, when an agent is not a pure altruist. How should he care about the actions of the junior he promoted? How should this agent care about the actions of the junior who is promoted by the junior he promoted? What about the junior promoted in his department ten generations later? It seems natural that an agent cares more about the actions of the junior he promoted than of the one a few generations later, even though his decision is necessary for both. One of the reasons is that in the latter case there are other seniors that contribute to the promotion. In other words, the distance between the promotion decision and the eventual increase in the group welfare affects how the agent values this increase. We can then introduce an "altruism" factor to reflect this imperfect altruism. The difference with the time discount factor is that imperfect altruism does not discount the wage but only group welfare gains.

More specifically, suppose that a senior agent assigns an "altruism" factor $\alpha \leq 1$ to the increase in the group welfare brought about by a junior he promoted, α^2 to the increase in the group welfare brought about by a junior promoted by a junior he promoted, etc. In a heterogenous department a promoted junior then obtains the utility of $w + d + \Delta W^f$, where ΔW^f is found from the equation $\Delta W^f = \alpha qp \left(d + \Delta W^f\right)$. The total effort becomes

$$E = (1-q) F(w) + q(1-p) F\left(w + \frac{d}{1-\alpha qp}\right).$$

The effect of the altruism factor α is then the same as the one of the probability of a heterogenous department q, see the discussion after Example 1. At zero patronage, α increases only the higher stakes effect, but in general it also increases the discouragement effect and its effect is ambiguous.

When the planner maximizes the share of left juniors as in Section 4, the altruism factor dampens the two motivation effects mentioned Lemma 3. When α is close to zero, that is, the agents are almost entirely driven by the warm glow, the third effect - the change in the relative motivation - almost disappears. Hence, Proposition 3 goes through with only two conditions - on the relative size, $\lambda \gtrless \frac{1}{2}$, and motivation, $d_r \gtrless d_l$, - instead of three.

5.2 Antagonistic and asymmetric groups

It has been assumed throughout the paper that the agents care only about their own group welfare. They may care instead about their *relative* well-being or status. They might be motivated not only by the possibility of using their own direct discretion but by the possibility of blocking the direct discretion of the other group. The utility of a given income depends on how much the other group earns while the effectiveness of the left-wing propaganda decreases when there is more right-wing propaganda. This antagonism can be captured by parameter $\beta_i \geq 0$, $i = l, r$, such that the welfare of group i decreases by factor β_i when the senior from the other group exerts direct discretion.

The welfare of the left group becomes

$$W_l = d_l \sum_{t=0}^{+\infty} \delta^t N_l^t - \beta_l d_r \sum_{t=0}^{+\infty} \delta^t N_r^t$$

and analogously for the right group. Then, each time a junior of group i is promoted instead of a junior from group $-i$, the direct impact on the welfare of group i is

$d_i + \beta_i d_{-i}$, $i = l, r$.

The groups may also differ in the weight with which the group welfare enters the agents' utility function. We have implicitly assumed throughout the paper that this weight is 1 for both groups. Here the weights are generalized to $\gamma_i > 0$, $i = l, r$. A higher γ_i corresponds to a group with a higher group altruism. Then, the immediate value of the promotion of a junior from group i instead of the one from the other group is $\tilde{d}_i \equiv \gamma_i (d_i + \beta_i d_{-i})$, $i = l, r$, for the members of group i.

The results of both Sections 3 and 4 go through with the "effective" direct discretions $\tilde{d}_i$. For example, instead of (2) the total output becomes

$$E = (1 - q) F(w) + \frac{1}{2} q (1 - p) [F_l + F_r], \tag{7}$$

where $F_i = F\left(w + \frac{\tilde{d}}{1-qp}\right)$, $i = l, r$. When maximizing the juniors' efforts as in Section 3, the planner is now comparing the discouragement and the higher stakes effect averaging over the two groups. In the case of the uniform cost distribution considered in Example 1 it is the average of the "effective" direct discretions that matters. In particular, in the expression for the optimal patronage p^* (4) one should use $\frac{1}{2}\left(\tilde{d}_l + \tilde{d}_r\right)$ instead of d.

6 Discussion and conclusions

We studied the design of promotions in an organization where agents belong to groups that advance their cause. Examples and applications include political groups, ethnicities, agents motivated by the work in the public sector, etc. Under either of two goals of the organizational designer considered, to maximize the efforts of junior agents and to maximize the number of the senior agents from a certain group, we showed that optimal patronage can be positive. The planner allows the senior agents to favor the juniors from their group in the contest for promotion even though these favours can be removed at no cost.

There are a number of interesting and promising extensions and alternative assumptions, some of which we outlined in Section 5. We hope that the rich but relatively simple framework proposed in this paper will be applied and used to generate many other interesting results. For example, in the paper the power of a senior bureaucrat - the direct discretion and promotion discretion - is assumed to be independent of what happens in other departments of the bureaucracy. Hence, the power of the group at the senior level is proportional to the number of its senior bureaucrats.

There are at least two reasons, however, why a larger group might have disproportionately more power. First, some decisions on the allocation of public funds, say, which regions to develop, require a joint decision of the senior bureaucrats. When the larger group has the required majority for the decision, it will of course bias the decision in its favor. The second reason is that promotions often require the agreement of more than just the head of the department. They are often decided by committees and might be vetoed by the "very" senior bureaucrats. Again, the larger group will then acquire more power that its share suggests.

We have also assumed that the entry to the bureaucracy is exogenous. But, since patronage affects the groups differently, the relative expected utility of joining the bureaucracy also depends on patronage. If people can choose to work in the private sector, where there are group homogenous firms as is the case in many developing countries, patronage affects the self-sorting of agents between the private and the public sectors and thus the composition of the junior level. Studying the labour market equilibrium is an exciting direction for future research.

Finally, in the working paper Drugov (2015) we considered the application to corruption in which some agents are corrupt and others are honest. Since screening for honesty at the entry level is impossible, and a few immoral agents will always be present, the main concern of the planner is to limit the spread of the corruption to upper ranks where it is much more damaging. The corrupt seniors take bribes using their direct discretion and "sell" the promotion to the corrupt juniors when the patronage is allowed. The honest seniors try not to promote corrupt juniors and get a boost in their utility from this action. As in Section 4, patronage benefits the larger group and the less motivated group. Thus, in some cases the optimal patronage is positive and even becomes maximal, that is, seniors have full discretion in promotions. In other words, the optimal policy is to fight fire with fire - to use patronage against corruption.

Appendix A. Proofs

Proof of Lemma 3. Express λ^S from (5) as

$$\lambda^S = \frac{\lambda}{1 - 2\lambda (1 - \lambda) p} \left[\lambda + (1 - \lambda)(1 - p)(1 + F_l - F_r)\right].$$

Its derivative with respect to p is equal to $\frac{\lambda(1-\lambda)}{1-2\lambda(1-\lambda)p}$ multiplied by

$$\frac{2\lambda-1}{1-2\lambda(1-\lambda)p}+\frac{1-2\lambda(1-\lambda)}{1-2\lambda(1-\lambda)p}(F_r-F_l)+(1-p)\frac{\partial(F_r-F_l)}{\partial p}.$$

The first term has the sign of $\lambda-\frac{1}{2}$ and it is thus positive when the left group is larger. The second term has the sign of F_r-F_l and it is positive when the left group is less motivated. The third term has the sign of $\frac{\partial(F_r-F_l)}{\partial p}$ which is ambiguous. Indeed,

$$\frac{\partial(F_r-F_l)}{\partial p}=\frac{2\lambda(1-\lambda)}{(1-2\lambda(1-\lambda)p)^2}(d_rf_r-d_lf_l),$$

where $f_i=f\left(w+\frac{d_i}{1-2\lambda(1-\lambda)p}\right)$, $i=l,r$. ■

Proof for Example 2. When c is distributed uniformly on $[w,w+1]$ and $d_i\in\left[0,\frac{1}{2}\right]$, $w+\frac{d_i}{1-2\lambda(1-\lambda)p}\in[w,w+1]$ for any $\lambda\in[0,1]$ and $p\in[0,1]$ and hence $F_i\left(w+\frac{d_i}{1-2\lambda(1-\lambda)p}\right)=\frac{d_i}{1-2\lambda(1-\lambda)p}$. Rewrite (6) as

$$\lambda^S=\frac{\lambda}{1-2\lambda(1-\lambda)p}\left[\lambda+(1-\lambda)(1-p)\left(1-\frac{d_r-d_l}{1-2\lambda(1-\lambda)p}\right)\right] \tag{8}$$

and take the first derivative with respect to p

$$\frac{\partial\lambda^S}{\partial p}=\frac{\lambda(1-\lambda)}{(1-2\lambda(1-\lambda)p)^2}\left((2\lambda-1)+(d_r-d_l)\frac{(1-2\lambda)^2+2\lambda(1-\lambda)p}{1-2\lambda(1-\lambda)p}\right).$$

When $\lambda\geq\frac{1}{2}$ and $d_r\geq d_l$ (the left group is larger and less motivated), both terms in brackets are positive, and therefore, $p^*=1$.

When $\lambda<\frac{1}{2}$ and $d_r<d_l$ (the left group is smaller and more motivated), both terms in brackets are negative, and therefore, $p^*=0$.

When the two terms have opposite signs, an interior value of p might be optimal. Solve the first-order condition $\frac{\partial\lambda^S}{\partial p}=0$ to obtain

$$p^{FOC}=\frac{1}{2\lambda(1-\lambda)}(2\lambda-1)\frac{1+(2\lambda-1)(d_r-d_l)}{2\lambda-1-(d_r-d_l)}.$$

Compute the second derivative of λ^S with respect to p

$$\frac{\partial^2\lambda^S}{\partial p^2}=\frac{8\lambda^2(1-\lambda)^2}{(1-2\lambda(1-\lambda)p)^3}\left(\lambda-\frac{1}{2}+(d_r-d_l)\frac{1-\lambda(1-\lambda)(3-p)}{1-2\lambda(1-\lambda)p}\right).$$

Plug in p^{FOC} to obtain

$$\frac{\partial^2 \lambda^S}{\partial p^2} \mid_{p=p^{FOC}} \propto 1 - 2\lambda + d_r - d_l.$$

When $\lambda < \frac{1}{2}$ and $d_r > d_l$ (the left group is smaller and less motivated), $\frac{\partial^2 \lambda^S}{\partial p^2} \mid_{p=p^{FOC}} > 0$ and therefore the optimal patronage is either 0 or 1. Comparing $\lambda^S \mid_{p=0}$ and $\lambda^S \mid_{p=1}$ we obtain that $p^* = 1$ if and only if $d_r - d_l \geq \frac{1-2\lambda}{1-2\lambda(1-\lambda)}$.

When $\lambda > \frac{1}{2}$ and $d_r < d_l$ (the left group is larger and more motivated), $\frac{\partial^2 \lambda^S}{\partial p^2} \mid_{p=p^{FOC}} < 0$ and therefore $p^* = p^{FOC}$ provided it is between 0 and 1. Since $d_r - d_l \geq -\frac{1}{2}$, $p^{FOC} > 0$. Solving $p^{FOC} \leq 1$ we obtain $d_r - d_l \leq 1 - 2\lambda$. ■

Appendix B. The Tullock contest

Consider now the usual Tullock contest success function (Tullock, 1980). The patronage introduces a multiplicative bias which is the standard specification of a biased Tullock contest as, for example, in Epstein, Mealem and Nitzan (2011) and Franke et al. (2013). In particular, if junior i is favoured by the senior, he wins the contest with probability

$$\Pr\{i \text{ is promoted}\} = \frac{(1+p)\, e_i^r}{(1+p)\, e_i^r + (1-p)\, e_{-i}^r},\ r \geq 1. \tag{9}$$

This specification keeps patronage between 0 and 1 as in the rest of the paper.

The effort cost is $C(e_i) = \frac{1}{\alpha} e_i^\alpha$, $\alpha \geq 1$. As in Section 3, the group motivation is the same in the two groups and is equal to d.

Lemma 4 *Bias p in the contest success function (9) results in the difference in promotion probabilities of the favoured and non-favoured juniors equal to p. The equilibrium efforts are $e_i^* = e_{-i}^* = \left(r \frac{1-p^2}{4} \left(w + \frac{d}{1-qp} \right) \right)^{\frac{1}{\alpha}}$.*

Proof. Denote the value of the promotion in a heterogenous department as v. The favoured junior i maximizes

$$\max_{e_i} \frac{(1+p)\, e_i^r}{(1+p)\, e_i^r + (1-p)\, e_{-i}^r} v - \frac{1}{\alpha} e_i^\alpha,$$

while the other junior maximizes

$$\max_{e_{-i}} \frac{(1-p)\, e_{-i}^r}{(1+p)\, e_i^r + (1-p)\, e_{-i}^r} v - \frac{1}{\alpha} e_{-i}^\alpha.$$

The two first-order conditions are

$$\frac{re_i^{r-1}e_{-i}^r\left(1-p^2\right)}{\left(e_i^r\left(1+p\right)+e_{-i}^r\left(1-p\right)\right)^2}v=e_i^{\alpha-1},\qquad\frac{re_{-i}^{r-1}e_i^r\left(1-p^2\right)}{\left(e_i^r\left(1+p\right)+e_{-i}^r\left(1-p\right)\right)^2}v=e_{-i}^{\alpha-1}.$$

Solving this system yields the equilibrium efforts

$$e_i^*=e_{-i}^*=\left(r\frac{1-p^2}{4}v\right)^{\frac{1}{\alpha}}. \tag{10}$$

Plugging (10) into (9), compute the difference in winning probabilities between the favoured junior i and the non-favoured one $-i$

$$\Pr\{i\text{ is promoted}\}-\Pr\{-i\text{ is promoted}\}=\frac{1+p}{2}-\frac{1-p}{2}=p.$$

The value of the promotion v is then $w+\frac{d}{1-qp}$. ■

The planner maximizes the total effort (up to a monotonic transformation) in a heterogenous department

$$\max_{p\in[0,1]}E^T=\left(1-p^2\right)\left(w+\frac{d}{1-qp}\right). \tag{11}$$

Proposition 4 *Optimal patronage p^* is intermediate, that is, $p^*\in(0,1)$ if and only if direct discretion is strictly positive, $d>0$. It is increasing in direct discretion d and decreasing in senior wage w. When $w=0$, $p^*=\frac{1-\sqrt{1-q^2}}{q}$.*

Proof. The first derivative of (11) with respect to p is

$$\frac{\partial E^T}{\partial p}=d\frac{qp^2-2p+q}{\left(1-pq\right)^2}-2pw$$

and the second is

$$\frac{\partial^2E^T}{\partial p^2}=-2d\left(1+q\right)\frac{1-q}{\left(1-qp\right)^3}-2w<0.$$

A solution to the first-order condition $\frac{\partial E^T}{\partial p}=0$ then gives the optimal patronage p^*. Since $\frac{\partial E^T}{\partial p}\mid_{p=0}=dq>0$, $p^*>0$. At $p=1$, $E^T=0$ while $E^T>0$ at $p<1$; thus, $p^*<1$. To get the comparative statics of p^*, note that

$$\frac{\partial^2E^T}{\partial p\partial w}<0,\qquad\frac{\partial^2E^T}{\partial p\partial d}>0.$$

Finally, take $w = 0$. Then, $\frac{\partial E^T}{\partial p} \mid_{w=0} \propto qp^2 - 2p + q$ and so

$$p^* \mid_{w=0} = \frac{1 - \sqrt{1 - q^2}}{q}.$$

■

Monetary incentives We have so far taken the senior wage as given and abstracted from direct monetary incentives for the juniors. The monetary incentives come at the cost of public funds μ. Hence, the planner maximizes the total effort (11) minus the wage costs

$$\max_{p \in [0,1], w \geq 0} (1 - p^2) \left(w + \frac{d}{1 - qp} \right) - \mu w.$$

Then Proposition 4 directly leads to the following corollary.

Corollary 2 *Optimal patronage p^* is positive for any positive costs of public funds μ.*

Indeed, Proposition 4 shows that optimal patronage is strictly positive for any senior wage. Even when providing monetary incentives is very cheap, at the margin a higher wage still has a first-order cost. In this contest specification as well as in many others (see Drugov and Ryvkin, 2017) biasing the contest has second-order costs at zero but first-order benefits. If juniors can be rewarded not only by the promotion but also by monetary bonuses for high output, the result still holds for the same reason.

References

Alesina, Alberto, and Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature*, 43(3): 762–800.

Andreoni, James. 2006. "Philanthropy." *Handbook of the Economics of Giving, Altruism and Reciprocity*, , ed. Kolm Serge-Christophe and Ythier Jean Mercier Vol. 2, 1201–1269. Elsevier.

Athey, Susan, Christopher Avery, and Peter Zemsky. 2000. "Mentoring and Diversity." *American Economic Review*, 90(4): 765–786.

Auriol, Emmanuelle, and Régis Renault. 2001. "Incentive Hierarchies." *Annals of Economics and Statistics / Annales d'Économie et de Statistique*, 63/64: 261–282.

Auriol, Emmanuelle, and Régis Renault. 2008. "Status and Incentives." *RAND Journal of Economics*, 39(1): 305–326.

Besley, Timothy, and Maitreesh Ghatak. 2005. "Competition and Incentives with Motivated Agents." *American Economic Review*, 95(3): 616–636.

Besley, Timothy, and Maitreesh Ghatak. 2008. "Status Incentives." *American Economic Review*, 98(2): 206–211.

Bettcher, Kim Eric. 2005. "Factions of Interest in Japan and Italy: The Organizational and Motivational Dimensions of Factionalism." *Party Politics*, 11(3): 339–358.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padró i Miquel. 2015. "The Value of Democracy: Evidence from Road Building in Kenya." *American Economic Review*, 105(6): 1817–51.

Carrillo, Juan D. 2000. "Corruption in Hierarchies." *Annals of Economics and Statistics / Annales d'Économie et de Statistique*, 59: 37–61.

Chen, Yan, and Sherry Xin Li. 2009. "Group Identity and Social Preferences." *American Economic Review*, 99(1): 431–457.

Coviello, Decio, Andrea Guglielmo, and Giancarlo Spagnolo. 2018. "The Effect of Discretion on Procurement Performance." *Management Science*, 64(2): 715–738.

Delfgaauw, Josse, and Robert Dur. 2008. "Incentives and Workers' Motivation in the Public Sector." *Economic Journal*, 118(525): 171–191.

Delfgaauw, Josse, and Robert Dur. 2010. "Managerial Talent, Motivation, and Self-Selection into Public Management." *Journal of Public Economics*, 94(9-10): 654–660.

Do, Quoc-Anh, Kieu-Trang Nguyen, and Anh N. Tran. 2017. "One Mandarin Benefits the Whole Clan: Hometown Favoritism in an Authoritarian Regime." *American Economic Journal: Applied Economics*, 9(4): 1–29.

Drugov, Mikhail. 2015. "Optimal Patronage." CEPR Discussion Paper No. 10343, https://www.cepr.org/active/publications/discussion_papers/dp.php?dpno=10343.

Drugov, Mikhail, and Dmitry Ryvkin. 2017. "Biased contests for symmetric players." *Games and Economic Behavior*, 103(Supplement C): 116–144.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. forthcoming. "The Value of Regulatory Discretion: Estimates from Environmental Inspections in India." *Econometrica*. https://economics.mit.edu/faculty/eduflo/papers.

Easterly, William, and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions*." *Quarterly Journal of Economics*, 112(4): 1203–1250.

Engers, Maxim, and Joshua S. Gans. 1998. "Why Referees Are Not Paid (Enough)." *American Economic Review*, 88(5): 1341–1349.

Epstein, Gil S., Yosef Mealem, and Shmuel Nitzan. 2011. "Political culture and discrimination in contests." *Journal of Public Economics*, 95(1-2): 88–93.

Franck, Rafaël, and Ilia Rainer. 2012. "Does the Leader's Ethnicity Matter? Ethnic Favoritism, Education, and Health in Sub-Saharan Africa." *American Political Science Review*, 106(02): 294–325.

Francois, Patrick. 2000. "'Public Service Motivation' As an Argument for Government Provision." *Journal of Public Economics*, 78(3): 275–299.

Francois, Patrick. 2007. "Making a Difference." *RAND Journal of Economics*, 38(3): 714–732.

Franke, Jörg, Christian Kanzow, Wolfgang Leininger, and Alexandra Schwartz. 2013. "Effort maximization in asymmetric contest games with heterogeneous contestants." *Economic Theory*, 52(2): 589–630.

Fryer, Roland, and Glenn Loury. 2005. "Affirmative Action in Winner-Take-All Markets." *Journal of Economic Inequality*, 3(3): 263–280.

Ghatak, Maitreesh, Massimo Morelli, and Tomas Sjöström. 2001. "Occupational Choice and Dynamic Incentives." *Review of Economic Studies*, 68(4): 781–810.

Hodler, Roland, and Paul A. Raschky. 2014. "Regional Favoritism." *Quarterly Journal of Economics*, 129(2).

Huang, Jing. 2000. *Factionalism in Chinese Communist Politics.* Cambridge University Press.

Iyer, Lakshmi, and Anandi Mani. 2012. "Traveling Agents: Political Change and Bureaucratic Turnover in India." *Review of Economics and Statistics*, 94(3): 723–739.

Kelman, Steven. 1990. *Procurement and Public Management: The Fear of Discretion and the Quality of Government Performance.* Washington, DC:AEI Press.

Kramon, Eric, and Daniel N. Posner. 2016. "Ethnic Favoritism in Education in Kenya." *Quarterly Journal of Political Science*, 11(1): 1–58.

Macchiavello, Rocco. 2008. "Public Sector Motivation and Development Failures." *Journal of Development Economics*, 86(1): 201–213.

Marx, Benjamin, Thomas Stoker, and Tavneet Suri. 2017. "There Is No Free House: Ethnic Patronage and Property Rights in a Kenyan Slum." http://economics.mit.edu/files/10461.

Meyer, Margaret A. 1992. "Biased Contests and Moral Hazard: Implications for Career Profiles." *Annals of Economics and Statistics / Annales d'Économie et de Statistique*, 25/26: 165–187.

Meyer, Margaret A. 1994. "The Dynamics of Learning with Team Production: Implications for Task Assignment." *Quarterly Journal of Economics*, 109(4): 1157–1184.

Morgan, John, Dana Sisak, and Felix Várdy. 2012. "On the Merits of Meritocracy." http://faculty.haas.berkeley.edu/rjmorgan/Meritocracy.pdf.

Neggers, Yusuf. 2018. "Enfranchising Your Own? Experimental Evidence on Bureaucrat Diversity and Election Bias in India." *American Economic Review*, 108(6): 1288–1321.

Persico, Nicola, José C. R. Pueblita, and Dan Silverman. 2011. "Factions and Political Competition." *Journal of Political Economy*, 119(2): 242–288.

Prendergast, Canice, and Robert H. Topel. 1996. "Favoritism in Organizations." *Journal of Political Economy*, 104(5): 958–978.

Rasul, Imran, and Daniel Rogger. 2015. "The Impact of Ethnic Diversity in Bureaucracies: Evidence from the Nigerian Civil Service." *American Economic Review*, 105(5): 457–61.

Shih, Victor C. 2009. *Factions and Finance in China: Elite Conflict and Inflation.* Cambridge University Press.

Shore, Lynn M., Beth G. Chung-Herrera, Michelle A. Dean, Karen Holcombe Ehrhart, Don I. Jung, Amy E. Randel, and Gangaram Singh. 2009. "Diversity in organizations: Where are we now and where are we going?" *Human Resource Management Review*, 19(2): 117–133.

Tajfel, Henri, and John Turner. 1986. "The Social Identity Theory of Intergroup Behavior." *The Psychology of Intergroup Relations*, , ed. Stephen Worchel and William Austin, 7–24. Chicago: Nelson-Hall.

Tullock, Gordon. 1980. "Efficient rent seeking." *Toward a Theory of the Rent-Seeking Society*, , ed. James M. Buchanan, Robert D. Tollison and Gordon Tullock, 97–112. College Station:Texas A&M University Press.