Optimal Patronage*

Mikhail Drugov†

May 8, 2018

Abstract

We study 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 and corruption. In an overlapping generations model, juniors compete for promotions. Seniors have two kinds of discretion: direct discretion, which allows an immediate advancement of their cause, and promotion discretion ("patronage"), which allows a biasing of the promotion decision in favour of the juniors from their group. We consider two settings differing in the planner's goal, maximizing juniors' efforts and affecting the steady-state composition of the senior level towards the preferred group, and show that patronage may be strictly positive in both of them. We also apply the second setting to the case of corruption.

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

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 the simple observation that people belong to different groups, and they care about the group to which they belong.[1] Group identity can be exogenous as in the case of ethnicities, tribes, castes and, in most cases, religions. It may also be endogenous and based on values, for example, political parties or political factions.[2]

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.

We build an overlapping generations model in which 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 ability to bias the contest in favour of the juniors they prefer based on their group identity. 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

[1] 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) and Marx, Stoker and Suri (2017) for the most recent (econometric) evidence and references there. Particularly Do, Nguyen and Tran (2017) and Marx, Stoker and Suri (2017) study the favoritism exerted by low-level bureaucrats who do not face any electoral pressure.

[2] There are "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.

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. This is thus a rent-seeking contest for (group) public goods. 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 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 present an example in which the sign of the third effect depends on the difference in motivations, as does the sign of the second effect. Thus, optimal patronage is determined by the size effect and the combined motivation effect. When the planner's preferred group is larger and less motivated, patronage is beneficial through both effects and is set at the maximum level; that is, seniors have full discretion about whom to promote. In the opposite case, when his preferred group is smaller and more motivated, zero patronage is optimal. Otherwise, there is a trade-off. We characterize optimal intermediate patronage. Overall, optimal patronage (weakly) increases with the size of the preferred group and decreases with its relative motivation.

We also present an application of this setting to corruption and investigate if patronage could be useful in combatting it. Some agents are corrupt, and the planner tries to limit the spreading of corrupt agents in the bureaucracy. In other words, his goal is to minimize the number of corrupt agents at the senior level.[3] Allowing for some patronage may then help since the honest seniors use it in order not to promote the corrupt juniors; however, corrupt seniors "sell" the position to corrupt juniors. Even though corrupt agents have no group motivation, the possibility of selling the position creates inter-generational linkage similar to that of group-motivated agents. In particular, the value of the position, and therefore, the bribe that is charged for it increase with the power at that position, that is, with patronage. Thus, formally, the model is very similar to the main model. Corrupt agents are motivated by bribes; honest ones are motivated by the desire not to allow corrupt juniors to be promoted, and the optimal patronage depends on the relative size and motivation of the two groups, as described above.

We then study a number of extensions. First, we allow agents to have warm-glow motivation and impure altruism. Then, we consider the case of antagonistic groups in which group welfare depends negatively on other group favours. Third, we suppose that the planner prefers one group to the other but also cares about the juniors' efforts, combining the two goals studied before in isolation. Fourth, we allow the planner to choose monetary incentives and show that some patronage may still be optimal. Finally, we also briefly discuss a number of interesting directions for future work.

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

[3] The composition of the junior level is exogenous as one cannot observe if a person applying for a governmental job will be corrupt.

juniors' efforts. In Section 4 the planner cares about the steady-state composition of the senior level. Section 4.1 analyzes the application to corruption. A few extensions are analyzed in Section 5. Section 6 discusses the related literature. Section 7 concludes. Appendix A contains the proofs. Appendix B considers two alternative contest models that generate similar results.

2 Model

This is an overlapping generations model in which 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.[4] The senior bureaucrat gets wage w and some discretionary power that we explain below. The junior who is not promoted gets utility normalized to 0.[5]

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.[6] 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.

2.2 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. If the group identity is based on ideology rather than ethnicity, senior

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

[5] He either leaves the bureaucracy or stays there in some low-level position with no discretionary power.

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

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.[7]

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.[8,9] 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.).

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, $c \rightsquigarrow F\left[\underline{c}, \overline{c}\right]$, and are privately informed about it.

[7] We discuss different ways of biasing the contest at the end of Section 3.3 and analyse 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.

[8] 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.

[9] 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.

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. We consider the case of $\delta < 1$ in Section 3.5.

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.

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

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

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.[10] 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.

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 total effort as

$$E = (1 - q) F (w) + q (1 - p) F \left(w + \frac{d}{1 - qp}\right) \tag{2}$$

and the planner maximizes it with respect to $p \in [0, 1]$.

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.

To understand when the optimal patronage is positive, let us compute the two effects at $p = 0$ (and conditional on being in a heterogeneous department). The value

[10] 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.

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. 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.

When the optimal patronage is positive, it is found from the first-order condition

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

We proceed with an example in order to have a simple closed-form solution.

Proposition 1 *Suppose that* $c \rightsquigarrow U[\underline{c}, \overline{c}]$. *Optimal patronage* p^* *is* 0, *if* $d \leq (\underline{c} - w)(1 - q)$ *or* $d \geq \frac{\underline{c} - w}{1 - q}$, *and otherwise it is*

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

Proof. See Appendix A.[11] ■

As we noted above, 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. For the case of the uniform distribution considered in Proposition 1, the optimal patronage (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.

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

[11] Condition $w < \underline{c}$ needed for (4) 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.

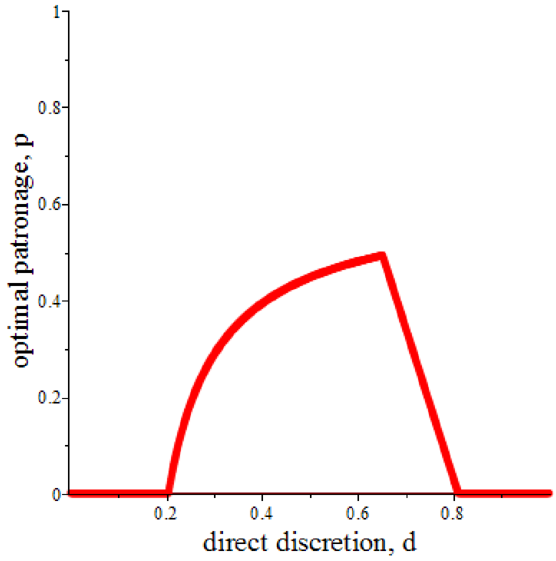

Figure 1: Optimal patronage when the costs are distributed as $Beta\,(5,1)$ $(F\,(c) = c^4)$ and $q = 0.4$, $w = 0.2$.

positive patronage. In general, however, both discouragement and higher stakes effects increase with q. For the uniform distribution of costs, 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.[12] 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 7). To summarize this discussion, introducing patronage as we do in this paper makes it *more* difficult to obtain a strictly positive optimal patronage.

[12] See Meyer (1992) for an early example of an additive bias in a Lazear-Rosen tournament and Franke et al. (2013) for characterization of the multiplicative bias in a general Tullock contest. See Drugov and Ryvkin (2017) for a general condition.

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 questions arise. Should the planner make departments homogenous or heterogenous? What is the optimal composition of the junior level, i.e., λ?

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 $\lambda = \frac{1}{2}$.

3.5 The effect of the discount factor

When the future periods are discounted with the discount factor δ, in a heterogenous department a promoted junior obtains the utility of $\delta\left(w + d + \Delta W^f\right)$, where ΔW^f is found from the equation $\Delta W^f = \delta qp\left(d + \Delta W^f\right)$. The total effort (2) becomes

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

A higher δ increases both the higher stakes effect, $f(\delta(w+d))\,\delta^2 qd$, and the discouragement effect, $F(\delta(w+d))$ (both computed at $p = 0$). Then, the overall effect is ambiguous. For the case of the uniform distribution considered in Proposition 1, the optimal patronage (4) becomes $\frac{1}{\delta q}\left(1 - \sqrt{d\frac{1-\delta q}{\frac{c}{\delta} - w}}\right)$ and it decreases with δ.

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.[13] For

[13] He then probably cares about the overall composition of the bureaucracy, but the composition of the junior level is exogenous. For example, it might be illegal to hire based on the group identity or the group identity may not be observable at the entry stage, as in the case of groups based on values.

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.

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[14]

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

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, 2) heterogenous departments headed by a left senior who uses promotion discretion, and 3) heterogenous departments where promotion is merit-based and the left junior wins it.

The effect of patronage on λ^S can be decomposed into three effects.[15] First, there is the size effect, proportional to $\lambda - \frac{1}{2}$: 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 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.

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

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

[15] See Lemma 3 in the Appendix A for the details.

Expressing λ^S from (5) yields

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

The planner maximizes (6) by choosing promotion discretion $p \in [0, 1]$. As before, we will proceed with a particularly well-behaved example when F is linear, that is, when the junior costs are distributed uniformly. In this case the relative motivation is proportional to the difference in direct discretions, $d_r - d_l$. Then, both motivation effects of patronage mentioned above, 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. See the next proposition and Figure 2.

Proposition 3 *Suppose that* $c \rightsquigarrow U[w, w+1]$ *and* $d_i \leq \frac{1}{2}$, $i = l, r$.[16] *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. ■

This Proposition 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 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

[16] These assumptions mean 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 length of the support equal to one is a normalization.

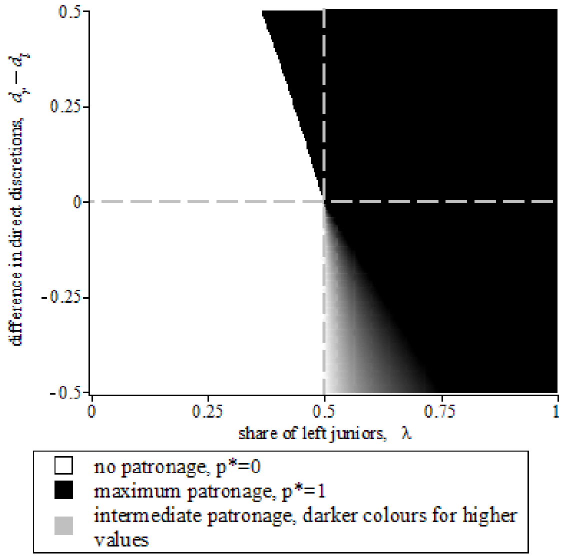

Figure 2: Optimal patronage depending on the share of left juniors, λ, and the difference in direct discretions, $d_r - d_l$.

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 *Optimal patronage p^* (weakly) increases with the share of left juniors, λ, and with the difference in direct discretions, $d_r - d_l$.*

4.1 Corruption

Let us apply the analysis of the previous section to the case of corruption. One group is "honest" and the other is corrupt; the planner is honest and minimizes corruption by minimizing the number of corrupt agents. Since it is impossible to distinguish the corrupt agents at entry level, the planner minimizes the number of corrupt agents at senior level. We thus abstract from the incentive problem of the juniors considered before. As seniors have much more power, juniors' efforts have only a second-order effect on the social welfare as compared to the number of the corrupt agents at the senior level. Another reason is that the previous literature has extensively studied the agency problem when some agents may engage in a corrupt behavior (see, for example, Mishra (2006) for a survey), while looking at the spreading of corrupt agents in an organization is new, to the best of our knowledge.

Corrupt agents take bribes for (not) doing their job, and corrupt senior bureaucrats also "sell" the promotions to corrupt juniors. Other agents are honest in that they dislike corruption. They do not take bribes and they try to prevent corruption if they can. We assume that inside the organization or, at least, inside each department, people know who is corrupt and who is not, but honest agents cannot reveal this to the outside world, either for the lack of hard proof or for the fears for personal safety.[17] Thus, the only way the honest agents can fight corruption is by not promoting corrupt juniors whenever they have such an opportunity. Corrupt seniors use patronage to sell their position, while the honest ones use it to not promote corrupt juniors.

The share of the honest juniors in the bureaucracy is λ. Honest agents derive utility g when an honest junior is promoted. It may come from their moral satisfaction that an honest rather than a corrupt agent obtains the job. It can also be their valuation of the harm for the society that a corrupt senior will do if they have some prosocial or public sector motivation. If corrupt seniors do not provide much effort, g may be the contribution of the honest senior towards the good of the society. In practice, of course, all three reasons might coexist. Proceeding in the same way as in Section 3.2 yields the value of the promotion for the honest juniors as equal to $w + \frac{g}{1-2\lambda(1-\lambda)p}$.

A corrupt senior bureaucrat takes b in bribes using his direct discretion. For example, he can take kickbacks for placing governmental orders, bribes for granting a licence or for not enforcing some rules. He can also literally sell the position. Whenever he exerts his promotion discretion, he can charge a bribe for promotion to a corrupt junior, if there is at least one in his department. This bribe may depend on whether one or both juniors are corrupt in his department; denote it b_1 and b_2 for the cases of one and two corrupt juniors, respectively.

The total expected bribe income of a corrupt senior bureaucrat is then

$$B = b + \left(2\lambda\left(1-\lambda\right)b_1 + \left(1-\lambda\right)^2 b_2\right)p. \tag{7}$$

Suppose that b_1 and b_2 are proportional to B with coefficients k_1 and k_2 which represent the bargaining power of the senior bureaucrat vis-à-vis the junior ones.[18] If $k_1 = k_2 = 1$, the senior bureaucrat has all the bargaining power and extracts all the surplus. However, his bargaining power is likely to be lower if the juniors cannot collect so much in bribes themselves and are credit-constrained. It might also be

[17] The movie *Serpico* (based on the true story of Frank Serpico, a New York policeman) is a good illustration of how corruption may be open and visible inside a department, and yet how difficult and dangerous it is to expose it.

[18] The bribes might be proportional to $B + w$, in which case the senior bureaucrat can sell the promotion even if he cannot take any direct bribes himself, i.e., when $b = 0$. This does not affect the results qualitatively.

reasonable to assume that $k_2 > k_1$ since the senior bureaucrat can essentially auction the promotion when both juniors are corrupt. Plugging $b_i = k_i B$, $i = 1, 2$, into (7) we obtain

$$B = \frac{b}{1 - Kp}, \tag{8}$$

where $K = 2\lambda(1-\lambda)k_1 + (1-\lambda)^2 k_2$ is the average bargaining power of corrupt seniors.

The steady-state composition of the senior level, λ^S, is (6) with $F_l = F\left(w + \frac{g}{1-2\lambda(1-\lambda)p}\right)$ and $F_r = F\left(w + \frac{b}{1-Kp}\right)$.

Consider first

$$K = 2\lambda(1-\lambda), \tag{9}$$

in which case the problem of the planner is exactly as before, that is, to maximize (6) with $d_l = g$ and $d_r = b$. Proposition 3 and Corollary 1 apply. In particular, optimal patronage increases as the honest group becomes larger and relatively less motivated, that is, as $b - g$ increases.

Let us now discuss the effect of the bargaining power of corrupt seniors, k_1 and k_2. When they increase, the motivation of corrupt juniors $\frac{b}{1-Kp}$ is scaled up more for any value of patronage. This affects optimal patronage via two opposed effects. On the one hand, patronage should decrease to counterbalance the scaling up of the motivation of corrupt juniors. On the other hand, a higher motivation of corrupt juniors means that their chances become higher in a fair contest, which calls for a higher patronage. The total effect is ambiguous and therefore a higher bargaining power of corrupt seniors may lead to a higher or lower optimal patronage.

5 Extensions

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 Proposition 1. At zero patronage, α only increases the higher stakes effect, but in general, it also increases the discouragement effect. For the uniform distribution of costs, as in Proposition 1, optimal patronage (4) becomes $\frac{1}{\alpha q}\left(1 - \sqrt{d\frac{1-\alpha q}{\underline{c}-w}}\right)$ and it first increases with α and then decreases.

5.2 Antagonistic and asymmetric groups

In Section 3, the two groups are symmetric and care only about their own direct discretion. In Section 4, the two groups have different direct discretions, d_l and d_r. We also mentioned in the corruption application in Section 4.1 that part of the motivation of the honest agents may come from preventing the harm to the society that corrupt seniors will do. Thus, they are motivated not by the possibility of using their own direct discretion but by the possibility of blocking the direct discretion of the other group. We call this antagonism, and it may be important in a wide range of situations. For example, the effectiveness of left-wing propaganda decreases when there is more right-wing propaganda. This is also the case when the groups care about their relative income or status. 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 $\gamma_i > 0$, $i = l, r$. A higher γ_i corresponds to a group with a higher group altruism. Proceeding in the same way as in Section 3.3, we can find the total output

$$E = (1-q)\, F(w) + \frac{1}{2} q\,(1-p) \left[F\left(w + \gamma_l \frac{d_l + \beta_l d_r}{1-qp} \right) + F\left(w + \gamma_r \frac{d_r + \beta_r d_l}{1-qp} \right) \right] \tag{10}$$

For our example with the uniform distribution of costs, it is the average base motivation which matters; denote

$$\overline{d} = \frac{1}{2} \left[(d_l + \beta_l d_r)\, \gamma_l + (d_r + \beta_r d_l)\, \gamma_r \right].$$

Proposition 4 *Suppose that* $c \rightsquigarrow U[\underline{c}, \underline{c}+1]$. *Optimal patronage* p^* *is* 0, *if* $\overline{d} \leq (\underline{c} - w)(1-q)$ *or* $\overline{d} \geq \frac{\underline{c}-w}{1-q}$, *and otherwise it is*

$$p^* = \frac{1}{q} \left(1 - \sqrt{\overline{d} \frac{1-q}{\underline{c} - w}} \right). \tag{11}$$

Proof. See Appendix A. ■

The optimal patronage (11) is very similar to the one in (4), with the only difference being that d is replaced by $\overline{d}$, which is the average one-period increase in the group welfare from the promotion. Since the altruism towards the group matters only in the heterogenous department, where there is one left junior and one right junior by definition, it is the average altruistic motivation that determines the total effort (also because F is linear).

5.3 The two planner's goals together

We have considered two possible goals of the planner, maximizing juniors' efforts (Section 3) and affecting the composition of the senior level (Section 4), separately. In many cases the planner, however, prefers one group to the other but still cares about the work done by the organization, that is, about juniors' efforts. For example, the planner is the politician for whom the efficiency of the government affects how likely he is to stay in power, but he himself belongs to one of the groups or panders to the bigger group because it contains the median voter. The planner may also have group preferences from the efficiency perspective too if he takes into account the cost of direct discretion and it differs between the two groups. The direct discretion of one group may be less distortionary per se. If one group is richer on average than the other, then a bias in public spending towards this group is more distortionary than an equally sized bias towards the poorer group. Finally, the planner may prefer the group with lower group motivation if agents from this group use the direct discretion less.

The planner may allow for patronage in order to motivate juniors and also to affect the composition of the senior level in favor of his preferred group. Suppose that the planner (dis)likes the direct discretion d_i with the weight $-h_i$, $i = l, r$. This weight is negative, $h_i > 0$, if the direct discretion means favours, corruption, etc. However, if the direct discretion is used by agents with the intrinsic motivation for public sector, then $h_i < 0$.

Denote

$$F_i = F\left(w + \gamma_i \frac{d_i + \beta_i d_{-i}}{1 - qp}\right), i = l, r,$$

the share of juniors of group i that exert effort. This share is increasing in the group motivation, $\gamma_i (d_i + \beta_i d_{-i})$. The planner's objective function is then to maximize (up to a constant)

$$(1 - q) F(w) + \frac{1}{2} q (1 - p) [F_l + F_r] + H\lambda^S, \tag{12}$$

where $H = h_r d_r - h_l d_l$ is the relative harm of the two groups. If it is positive, the planner prefers the seniors from the left group. The steady-state composition of the senior level, λ^S, is (6).

For our example with the uniform distribution of costs, the difference in shares $F_r - F_l$ is proportional to the difference in motivations

$$\Delta = \gamma_r (d_r + \beta_r d_l) - \gamma_l (d_l + \beta_l d_r),$$

which we also call the relative motivation of the right group. As we showed in

Section 4, the influence of patronage on λ^S can be decomposed into the size effect, proportional to $\lambda - \frac{1}{2}$, and the (total) motivation effect, proportional to Δ.[19]

Proposition 5 *Suppose that* $c \rightsquigarrow U\left[\underline{c}, \underline{c}+1\right]$.

(i) Optimal patronage p^* *increases with the relative motivation of the right group,* Δ, *if and only if the planner prefers the left group,* $H > 0$.

(ii) If $\Delta = 0$, *optimal patronage* p^* *is* 0, *if* $\overline{d} - \frac{\lambda - \frac{1}{2}}{1-q} H \leq (\underline{c} - w)(1-q)$ *or* $\overline{d} - \frac{\lambda - \frac{1}{2}}{1-q} H \geq \frac{\underline{c}-w}{1-q}$, *and otherwise it is*

$$p^* = \frac{1}{q}\left(1 - \sqrt{\frac{(1-q)\,\overline{d} - \left(\lambda - \frac{1}{2}\right) H}{\underline{c} - w}}\right).$$

Proof. See Appendix A. ■

When Δ increases, patronage favours the left group more since it becomes more likely to lose the fair contest. If $H > 0$, that is, the right group is relatively more harmful for the planner, the planner counteracts higher chances of the right group in a fair contest by allowing for more patronage. If $H < 0$, that is, the planner prefers the right group, he makes the contest more fair if the right juniors are more likely to win it. This is a generalization of Corollary 1 to the case when the planner cares both about the juniors' efforts and composition of the senior level.

For a closed-form solution we need to assume that the two groups do equally well in the fair contest, that is, $\Delta = 0$. Then, there is only the first effect, as we mentioned above, and the patronage unambiguously benefits a larger group. If the planner prefers the left group, $H > 0$, and it is larger, the optimal patronage is higher than in the baseline model (4) as it is used partly to facilitate the promotions of the left juniors.

5.4 Monetary incentives

We have so far taken the senior wage as given and abstracted from direct monetary incentives for the juniors. The monetary incentives of course come at the cost of public funds. Taking a standard specification of a biased contest, we can show the following.

[19] As we show there, there are two motivation effects: the relative motivation effect proportional to $F_r - F_l$ and the change in the relative motivation, proportional to $\frac{\partial(F_r - F_l)}{\partial p}$. When F is linear, both are proportional to the difference in motivations.

Proposition 6 *Consider the biased Tullock contest from Appendix B. Optimal patronage is positive for any positive costs of public funds.*

Proposition 7 in Appendix B shows that optimal patronage is strictly positive for any senior wage. Even when providing monetary incentives is very cheap and the senior wage is high, at the margin increasing it still has a first-order cost. In this contest specification as well as in many others, patronage has second-order costs at zero but first-order benefits. If juniors can be rewarded not only by the promotion but also by direct monetary incentives for high output, the result still holds for the same reason.[20]

5.5 The two groups caring about the same cause

It is possible that the two groups care about the same cause, but to a different extent. For example, public sector workers may all be motivated by (common) social welfare but to a different extent. One group (say, the left) consists of agents that are highly motivated, while the other group (say, the right) consists of workers that are less motivated. The social welfare is then

$$W = \sum_{t=0}^{+\infty} d_l \delta^t N_l^t + \sum_{t=0}^{+\infty} d_r \delta^t N_r^t,$$

that is, $\beta_l = \beta_r = -1$. Then, if $d_l > d_r$, both groups want to promote the juniors of the left group since the left seniors contribute more towards the common welfare. In a heterogenous department, the value of promotion for the right juniors is less than wage w since their promotion is worse for the social welfare than the promotion of the left juniors.

5.6 Other interesting extensions

There are a number of other interesting extensions for future work. We have assumed throughout the paper that patronage is set once and for all. But what if it can be changed every few periods? When the goal of the planner is to maximize the total effort, a preliminary analysis shows that optimal patronage converges to the stationary one when the number of periods increases. There are end effects since in the first and the last periods there are costs of patronage but no benefits but their effect vanishes as the number of periods increases.

[20] There might be then a question whether monetary incentives should be provided directly or as a senior wage but this does not affect the optimality of patronage.

Another assumption that we maintain elsewhere in the paper is that the power of a senior bureaucrat, the direct discretion and promotion discretion, is independent of what happens in the bureaucracy in other departments. 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 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.

Another promising avenue for future work is to study the situations when the group identity can be changed or hidden, which is of course most relevant when the groups are based on values. For example, a left-wing person may change his convictions when surrounded by right-wing colleagues. He can also hide that he is from the left if his boss is from the right. In the corruption setting both possibilities are particularly relevant. An honest agent may succumb to the temptation of high bribes taken by his colleagues, and a corrupt junior may refrain from taking bribes if his senior is honest in order to get the promotion. In this case, for example, the zero-tolerance policy toward corruption may be counterproductive: it will prevent the corrupt behavior at the junior level where the gains are often low and the risks are high, and hence, not allow potentially corrupt juniors to reveal themselves. In other words, it will reduce corruption at the junior level but increase it at the senior level.

Finally, the entry to the bureaucracy is assumed exogenous. However, since patronage affects the groups differently, unless they are equal in size and motivations, the relative expected utility of joining the bureaucracy also depends on patronage. Patronage then affects the composition of the junior level as well.

6 Related literature

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.[21] 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 completely decides on promotion, in which case the discouragement effect is always of the first order. In Appendix B 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 (see Section 5.4).

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.

Since the group welfare is essentially a (group) public good, the contest for the promotion is similar to the models of rent-seeking for public goods such as Katz, Nitzan and Rosenberg (1990) and Linster (1993). Unlike the usual contest where each participant cares only about winning the contest, here even losers care about the identity of the winner, that is, whether or not he is from the same group.

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

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

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), Auriol and Renault (2008) and Besley and Ghatak (2008).

The political economy models of Roberts (1999) and Acemoglu, Egorov and Sonin (2012) have a similar feature that admitting new members to a club or to a ruling coalition (in our case, promoting) will affect everybody, not only through their direct actions but also via changes in future membership since these new members have voting power.

The application of our model to corruption focuses on the selling of positions, which is completely absent from the corruption literature.[22] Also, very few papers consider organizational design with corrupt agents.

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).

[22] See, for example, the two-volume handbook Rose-Ackerman (2006) and Rose-Ackerman and Søreide (2012).

7 Conclusion

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, and corruption. 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.

We also considered the application to corruption in which some agents are corrupt and others are honest. The corrupt seniors take bribes using their direct discretion and "sell" the promotion to the corrupt juniors. Whenever possible, the honest seniors do not promote corrupt juniors and get a boost in their utility from this action. The planner minimizes the corruption at the senior level (the distribution of junior types is exogenous). Patronage benefits the larger group and the less motivated group. Thus, in some cases the optimal patronage is positive and even becomes maximum, that is, seniors have full discretion in promotions. This is despite the fact that corrupt seniors use patronage to sell promotions to corrupt juniors.

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.

Appendix A. Proofs

Proof of Proposition 1. When $c \rightsquigarrow U\left[\underline{c}, \overline{c}\right]$, the first-order condition (3) becomes

$$-\left(w+\frac{d}{1-qp}-\underline{c}\right)+(1-p)\,\frac{qd}{\left(1-qp\right)^2}=0 \tag{13}$$

The second derivative is $-2dq\frac{1-q}{(1-pq)^3}<0$ and the second-order condition is therefore satisfied.

(13) can be rewritten as

$$q^2p^2-2qp+1-\frac{1-q}{\underline{c}-w}d=0$$

There are two roots, $\frac{1}{q}\left(1 \pm \sqrt{d\frac{1-q}{\underline{c}-w}}\right)$, but the larger is always greater than one. Thus, $p^* = \frac{1}{q}\left(1 - \sqrt{d\frac{1-q}{\underline{c}-w}}\right)$. Condition $p^* > 0$ gives $d < \frac{\underline{c}-w}{1-q}$ and condition $p^* < 1$ gives $d > (\underline{c} - w)(1 - q)$. ■

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, $F_l - F_r$.*

Proof. 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,

$$\frac{\partial(F_r - F_l)}{\partial p} = \frac{2\lambda(1-\lambda)}{(1 - 2\lambda(1-\lambda)p)^2}(d_r f_r - d_l f_l),$$

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

Proof of Proposition 3. When $c \rightsquigarrow U[w, w+1]$ 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

$$\lambda^S = \frac{\lambda}{1 - 2\lambda(1-\lambda)p}\left[\lambda + (1-\lambda)(1-p)\left(1 - \frac{d_r - d_l}{1 - 2\lambda(1-\lambda)p}\right)\right] \quad (14)$$

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 $d_r - d_l \leq 1 - 2\lambda$. ■

Proof of Proposition 4. When $c \rightsquigarrow U[\underline{c}, \underline{c}+1]$,

$$\frac{1}{2}\left(F\left(w + \gamma_l\frac{d_l + \beta_l d_r}{1 - qp}\right) + F\left(w + \gamma_r\frac{d_r + \beta_r d_l}{1 - qp}\right)\right) = w + \frac{\overline{d}}{1 - qp} - \underline{c}$$

and so the planner's problem to maximize (10) is equivalent to maximizing (2) and Proposition 1 applies with $d = \overline{d}$. ■

Proof of Proposition 5. **Part i)** Using (6) write

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

The second cross-derivative of the planner's problem (12) with respect to p and

Δ is equal to

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

Thus, $\text{sgn}\left(\frac{\partial p^*}{\partial\Delta}\right)=\text{sgn}\left(\frac{\partial^2\left(H\lambda^S\right)}{\partial p\partial\Delta}\right)=\text{sgn}\left(H\right)$.

Part ii) When $\Delta=0$, the first-order condition of problem (12) is

$$q\left(-\left(w+\frac{\overline{d}}{1-qp}-\underline{c}\right)+\left(1-p\right)\frac{q\overline{d}}{\left(1-qp\right)^2}\right)+\left(\lambda-\frac{1}{2}\right)\frac{q}{\left(1-qp\right)^2}H=0. \quad (15)$$

The second-order condition is $-\frac{2q}{(1-pq)^3}\left(\overline{d}\left(1-q\right)-\left(\lambda-\frac{1}{2}\right)H\right)<0$. If it is not satisfied, the problem is convex and $p^*=0$ since at $p=1$ the total effort is zero. Assuming the second-order condition is satisfied, rewrite (15) as a quadratic equation in p

$$q^2p^2-2qp+1-\frac{\left(1-q\right)\overline{d}-\left(\lambda-\frac{1}{2}\right)H}{\underline{c}-w}=0.$$

There are two real roots, $\frac{1}{q}\left(1\pm\sqrt{\frac{(1-q)\overline{d}-\left(\lambda-\frac{1}{2}\right)H}{\underline{c}-w}}\right)$, but the larger is always greater than one. Thus, $p^*=\frac{1}{q}\left(1-\sqrt{\frac{(1-q)\overline{d}-\left(\lambda-\frac{1}{2}\right)H}{\underline{c}-w}}\right)$ provided it is between 0 and 1. Note that $p^*<1$ implies the second-order condition. ■

Appendix B. Alternative contest models

Here we present two alternative contest models that generate results very close to the ones obtained before.

B1. Tournament

As before, patronage p means that with probability p the senior bureaucrat has full discretion in deciding whom to promote, while with probability $1-p$ there is a fair contest. Here, the contest is the standard Lazear-Rosen tournament.

Junior i, $i=l,r$, exerts effort e_i at the cost $C\left(e_i\right)=\frac{1}{2}e_i^2$ and is promoted if $e_i-e_{-i}+u\geq 0$, where $u\rightsquigarrow U\left[-\frac{1}{2},\frac{1}{2}\right]$. This is the simplest specification of the Lazear-Rosen tournament and has been used in Meyer (1991), Konrad (2009), Ederer (2010) and Brown and Minor (2014), among others. The probability that junior i is promoted is equal to (assuming interior solution)

$$\Pr\{u \geq e_{-i} - e_i\} = \frac{1}{2} + e_i - e_{-i}.$$

Then, each junior solves

$$\max_{e_i} \left(\frac{1}{2} + e_i - e_{-i}\right) v_i - \frac{1}{2} e_i^2,$$

where v_i is the value of the promotion for junior i. It is equal to w in a homogenous department and to $w + \gamma_i \frac{d_i + \beta_i d_{-i}}{1-qp}$ in a heterogenous department. Then, optimal effort is $e_i^* = v_i$ and the aggregate effort in a heterogenous department is

$$(1-p)\left(e_l^* + e_r^*\right) = (1-p)\left(w + \gamma_l \frac{d_l + \beta_l d_r}{1-qp} + w + \gamma_r \frac{d_r + \beta_r d_l}{1-qp}\right),$$

which is a particular example of the same object in (10).

B2. Tullock contest

The contest for the promotion is now a biased Tullock one in which patronage $p \in [0,1]$ is the bias that takes the following form. If junior i is favoured by the senior, he wins the contest with probability

$$\Pr\{i \text{ is promoted}\} = \frac{(1+p)\, e_i^r}{(1+p)\, e_i^r + (1-p)\, e_{-i}^r},\ r \geq 1. \tag{16}$$

This is a standard specification of a biased Tullock contest as, for example, in Epstein, Mealem and Nitzan (2011) and Franke et al. (2013). 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 (16) 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

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

Solving this system yields the equilibrium efforts

$$e_i^* = e_{-i}^* = \left(r \frac{1-p^2}{4} v\right)^{\frac{1}{\alpha}}. \tag{17}$$

Plugging (17) into (16), 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

$$\max_{p\in[0,1]} E^T = \left(1-p^2\right)\left(w + \frac{d}{1-qp}\right). \tag{18}$$

Proposition 7 *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 (18) with respect to p is

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

and the second is

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

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

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

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

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

■

References

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2012. "Dynamics and Stability of Constitutions, Coalitions, and Clubs." *American Economic Review*, 102(4): 1446–76.

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.

Brown, Jennifer, and Dylan B. Minor. 2014. "Selecting the Best? Spillover and Shadows in Elimination Tournaments." *Management Science*, 60(12): 3087–3102.

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.

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

Ederer, Florian. 2010. "Feedback and Motivation in Dynamic Tournaments." *Journal of Economics & Management Strategy*, 19(3): 733–769.

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.

Katz, Eliakim, Shmuel Nitzan, and Jacob Rosenberg. 1990. "Rent-Seeking for Pure Public Goods." *Public Choice*, 65(1): 49–60.

Konrad, Kai A. 2009. *Strategy and Dynamics in Contests.* Oxford:Oxford University Press.

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

Linster, Bruce G. 1993. "A Generalized Model of Rent-Seeking Behavior." *Public Choice*, 77(2): 421–435.

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://web.mit.edu/tavneet/www/Marx_Stoker_Suri.pdf.

Meyer, Margaret A. 1991. "Learning from Coarse Information: Biased Contests and Career Profiles." *Review of Economic Studies*, 58(1): 15–41.

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.

Mishra, Ajit. 2006. "Corruption, Hierarchies and Bureaucratic Structure." *International Handbook On The Economics Of Corruption*, , ed. Susan Rose-Ackerman Vol. 1, Chapter 7. Edward Elgar Publishing.

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

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.

Roberts, Kevin. 1999. "Dynamic Voting in Clubs." LSE Discussion Paper No. TE/99/367, http://eprints.lse.ac.uk/19349/.

Rose-Ackerman, Susan. 2006. *International Handbook On The Economics Of Corruption.* Vol. 1, Edward Elgar Publishing.

Rose-Ackerman, Susan, and Tina Søreide. 2012. *International Handbook On The Economics Of Corruption.* Vol. 2, Edward Elgar Publishing.

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