CHOOSING A DICTATOR:
BUREAUCRACY AND WELFARE IN
LESS DEVELOPED POLITIES

James E. Rauch

CHOOSING A DICTATOR:
BUREAUCRACY AND WELFARE IN
LESS DEVELOPED POLITIES

ABSTRACT

Recent work in the sociology of economic development has emphasized the establishment
of a professional government bureaucracy in place of political appointees as an important
component of the institutional environment in which private enterprise can flourish. I focus on
the role that internal promotion can play in bringing to power individuals who highly value
(relative to income) imposition of their preferences over collective goods on the public. Such
individuals restrain the corruption of their subordinates as a byproduct of their efforts to
implement their preferences using tax revenue. Within this hierarchical framework I investigate
the effects of varying subordinate compensation levels and of recruiting them meritocratically.

James E. Rauch
Department of Economics
University of California, San Diego
La Jolla, CA 92093
and NBER

I. Introduction

Recent analyses of economic policy-making in less developed countries (LDCs) have stressed that the individuals who make up the state apparatuses can to some extent act independently, rather than responding passively to voters or interest groups as in much of the political economy literature. Such a state might be expected to exhibit the "predatory" behavior predicted by writers such as Lal (1988), as each state functionary seeks to implement regulations on private sector economic activity that will maximize the bribes he can extract. Indeed, we do observe such purely rent-seeking states in LDCs. A good example is Zaire, of which President Mobutu has stated "holding any slice of public power constitutes a veritable exchanged instrument, convertible into illicit acquisition of money or other goods" (Young 1978, p. 172). What is remarkable is that some LDC governments do *not* act as predators. In East Asia, for example, the Korean and Taiwanese states have worked hand in glove with the private sector to promote investment and enhance the capacity of private firms to enter international markets (Amsden 1989, Wade 1990), earning these governments the moniker "developmental states".

In his comparative analysis of the role of the state in the development of several LDCs, Evans (1992) argues that professionalization of the state bureaucracy is a necessary (though not sufficient) condition for a state to be "developmental". The key institutional characteristics of what he calls "Weberian" bureaucracy include meritocratic recruitment through genuinely competitive examinations, Civil Service procedures for hiring and firing rather than political appointments and dismissals, and filling higher levels of the hierarchy through internal promotion. In previous work I studied the potential impact that bureaucratic professionalism could have on the positive role that the state can play in economic development by providing complementary inputs for the private sector. Specifically, in Rauch (forthcoming) I hypothesized that establishment of a professional bureaucracy in place of political appointees will lengthen the period that public decision

makers are willing to wait to realize the benefits of expenditures, leading to allocation of a greater proportion of government resources to long-gestation period projects such as infrastructure. I also hypothesized that this increased government investment in inputs complementary to private capital will increase the rate of economic growth. These hypotheses were tested using data generated by a "natural experiment" in the early part of this century, when a wave of municipal reform transformed the governments of many U. S. cities. Controlling for city and time effects, adoption of Civil Service was found to increase the share of total municipal expenditure allocated to road and sewer investment. This share in turn was found to have a positive effect on growth in city manufacturing employment.

I now wish to turn to the impact of "Weberianism" on the negative effect the state can have on economic development through corruption or "predation". Investigation here is hampered on two fronts: empirical and theoretical. On the empirical front, the problem is that corruption is hard to measure. On the theoretical front, the problem is that we really do not know how the various elements that add up to professional bureaucracy restrain predatory behavior (assuming they in fact do so) and thus do not know what to expect when not all of these elements are present, as might typically be the case.

Some progress is being made on the empirical front. Keefer and Knack (1993) and Mauro (1993) have both collected privately produced measures of bureaucratic performance and related them in cross-country regressions to economic growth. Keefer and Knack use ratings by the International Country Risk Guide (ICRG) and by Business and Environmental Risk Intelligence (BERI) of "corruption in government" and "bureaucratic delays", respectively, while Mauro uses ratings by Business International (BI) of "corruption" and "bureaucracy and red tape". Keefer and Knack (Table 5) find that better performance on both of their variables is positively and significantly associated with growth, and Mauro (Table 8) finds that better performance on both of his variables is positively and significantly associated with the private investment share of GDP.

Unfortunately, while this evidence reinforces the idea that differential governmental performance may have an impact on economic growth, it tells us little about what kind of institutional characteristics are associated with lower levels of corruption or red tape. If the findings of Keefer and Knack and Mauro are meaningful, it is worth identifying which characteristics of government bureaucracies lead to good ratings from the ICRG, BERI, and BI on the variables cited above. But in order to know what to look for we need some additional theoretical guidance. That is the most immediate purpose of this paper. Its longer-term goal is to deepen our understanding of the behavior of government bureaucracies in general.

II. The argument

Bureaucratic corruption is typically addressed using a principal-agent model (see, e.g., Klitgaard 1988), but the standard assumption of such work is that the principal himself is not corrupt, which misses the entire problem of the predatory state. If we are to retain the utility of the principal-agent model without being irrelevant we must therefore model corruption on the part of the principal. This could mean grafting a model of the entire political process onto a model of bureaucratic corruption. I feel, however, that at the present time this would be attempting to do too much. Instead, I abstract from the political process by identifying the state with the bureaucracy. In doing so I am inspired by the example of Soskice, Bates, and Epstein (1992). I also borrow from this paper the assumption of a hierarchical division of labor within the bureaucracy, where decisions can be made only at the top and implemented only at the bottom, and the assumption that individuals may enjoy leadership for its own sake. I believe the identification of the bureaucracy with the state is less restrictive than it seems at first. There exist one-party states where the bureaucracy is very closely identified with the party,[1] military

---

[1] It is worth noting that, for example, Mexico is essentially a one-party state, and that the current and former Presidents of Mexico at the time of writing (Ernesto Zedillo and Carlos Salinas de Gortari) never

dictatorships where the bureaucracy is the military hierarchy, and plenty of executive bureaucracies in various countries with substantial autonomy from political control (in which case the model below would be interpreted to apply only to those aspects of economic affairs over which this bureaucracy has power). In empirical application this modeling strategy amounts to seeing what can be explained by the structure of the bureaucracy, taking the political process as exogenous. One might argue that the political process can negate any incipient effects that bureaucratic structure might generate, but my research cited above offers some hope that this is not always the case.

The model of sections III and IV below contains two key elements. First, individuals are assumed to differ in their desire to exercise effective power, by which I mean their desire to impose their preferences over collective goods on the public.[2] I call the level of this desire power-hunger or ph for short.[3] One can only exercise effective power when one can choose the mix of collective goods the state will supply (or at least the mix of the subset of goods supplied by one's "insulated" bureaucracy), so one's ph can only be satisfied at the top of the bureaucratic hierarchy. Second, there exist different opportunities for corruption at the different levels of the bureaucracy. At the bottom one can engage in "petty corruption", which is defined as stealing tax revenues intended for provision of public goods. An example would be taking kickbacks as a percentage of the value of contracts awarded to collect garbage or build a road. At the top one can engage in "large-scale corruption", which is defined as the use of state regulatory powers to create rents. An example would be establishment of a state trading monopoly in which one has a stake directly or through relatives.

---

ran for elected office before they were selected to run for President.

[2]For simplicity I assume that all individuals in the society are identical in the extent to which they care about income (but see footnote 11 in section IV.A below).

[3]The parallel concept in Soskice et al. (1992) is "ambition". However, they do not allow ambition to vary across individuals, nor can they clearly distinguish it from the rate at which individuals discount the future. The latter limitation is related to the fact that the government in their model does not do anything with the revenue it collects (other than consume it).

To see how these two elements interact I need to specify some more details of the model. The government consists of one *dictator* (so named because he functions as Arrow's Dictator) and a small number of *deputies*. I make the realistic assumption that the dictator needs the deputies to carry out their tasks in order for the government to supply goods and services, but that he can pursue corruption on his own (or with the help of relatives and friends). The deputies allocate their time between their assigned tasks and (petty) corruption. The dictator allocates his time between monitoring the deputies and (large-scale) corruption. A high ph dictator will closely supervise his deputies to force them to implement his will by using the tax revenue under their control to supply the mix of public goods he has chosen, leaving him little time for corrupt pursuits. A low ph dictator is not interested in imposing his preferences over collective goods and hence spends little time supervising his deputies, instead concentrating on creating and appropriating rents while they rob the public till.

Now suppose we institute a rule of internal promotion (the component of Weberian bureaucracy emphasized by Soskice et al.), so that the next dictator can be chosen only from the current deputies rather than from the entire population. This means that deputies have more than a negligible chance of becoming dictator and exercising power. I argue that this will generate an important kind of self-selection among deputies. Any deputy wants to enjoy petty corruption and also wants to be promoted and enjoy large-scale corruption. A high ph deputy, however, wants to be promoted more because he will also enjoy exercise of effective power. It follows that if there is any effective supervision a high ph deputy would respond by reducing his petty corruption more than would a low ph deputy.[4] Since deputies who care about effective power are more likely to

---

[4]The reader might reasonably ask why the deputies do not use tax revenue to bribe the dictator and thus render supervision ineffective. The answer is that the dictator's comparative advantage in large-scale corruption leads him to satisfy his desire for income through this channel and satisfy his desire to exercise effective power using tax revenue. Obviously this answer only works if the dictator's ph is sufficiently high. If it is not, the deputies use all tax revenue under their control either for personal consumption or to bribe the dictator and government supply of collective goods is zero. (See also the discussion in

become dictator, dictators are more likely to care about effective power. A dictator who values exercise of effective power highly will in turn spend more time supervising his deputies to insure that they are carrying out their tasks and less time looking for ways to line his own pockets. Thus internal promotion is a self-reinforcing system that increases the expected ph of dictators, tending to increase the extent to which the bureaucracy as a whole carries out its assigned tasks of public goods provision and decrease the extent to which it implicitly taxes the private sector through large-scale corruption.

Whether or not there is a rule of internal promotion, the model described predicts that the amounts of petty and large-scale corruption will move together. This may typically be true, yet one can certainly imagine that a country could for example have judges on the take (they are not using their salaries to provide the expected level of service) and buses that never work, yet farmers could receive fair prices for their crops and manufacturers could import needed inputs without paying bribes. To allow for this possibility I find that I need to disaggregate the state bureaucracy. This is done at the end of the paper, where I also speculate that extensions of the model can help us understand phenomena such as "pockets of efficiency" (bolsoes de eficiencia--see Geddes 1986, p. 105) and "esprit de corps".

III. The model without internal promotion

There are three kinds of agents in the model: the public, the deputies, and the dictator. There exists a continuum of agents of measure n. As in Soskice et al. (1992), I assume an overlapping generations structure where each agent lives for two periods. Each agent obtains utility from consumption of a public and a private good, the latter serving as

---

section III, p. 15.) I would venture to guess, however, that in real-world bureaucracies where this outcome is observed the proximate cause is not low ph but rather the monopolisation of opportunities for large-scale corruption and/or exercise of effective power by politicians, making the ph of the "dictator" (the top-level bureaucrat) irrelevant. Thus the worst possible outcome in terms of provision of collective goods occurs when bureaucrats are both unsupervised by politicians and powerless.

numeraire. It is convenient to assume that this utility is additively separable not only across time but also across goods. There are no capital markets. The lack of capital markets insures that the government's budget is balanced in every period. The government purchases q units of the private good to produce each unit of the public good. The public good is a differentiated product of which a continuum of types can be produced. An example could be transportation infrastructure, which can be supplied by a mix of roads, rail facilities, airports, etc., so that we can think of the various mixes of equal cost as corresponding to the "types" of public good.

The public

It is assumed that the public is subjected to two forms of taxation, explicit and implicit, which form the bases for two types of corruption, petty and large-scale. Explicit taxation means taxation in the usual sense of the word: trade taxes, sales taxes, income taxes, etc. These are the tax revenues that finance government expenditure on the public good. However, these tax revenues may also fall prey to petty corruption: the deputies charged with procuring the private goods necessary to produce the public good may instead procure some of these private goods for themselves. Implicit taxation is assumed to be the sole prerogative of the dictator and is equivalent to the large-scale corruption he undertakes. Following Shleifer and Vishny (1993), I assume that implicit taxation is more distortionary than explicit taxation. In order to keep the model simple, however, I implement this distinction by assuming that explicit taxation is *non*distortionary (i.e., lump-sum) while implicit taxation takes a very simple distortionary form to be described below.

As in Soskice et al., I assume that the public can only defend itself against the predation of the government by withdrawing its resources from the market. In the language of Hirschman (1970), the public in my model has the option of "exit" but not that

of "voice". This is what I mean by "less developed polities" in the title of this paper.[5] Specifically, I follow Ades and Verdier (1993, p. 5) in assuming that each member of the public is endowed with one unit of labor, which she can allocate to either one of two private good production activities. The first has decreasing returns to scale, with output given by $f(\ell_{1t})$, $f' > 0$, $f'' < 0$, $f'(0) = \infty$, $f'(1) = 0$, where ℓ_{1t} is labor input to activity 1 at time t. The output of this activity cannot be taxed by the dictator and can be interpreted as subsistence agriculture or perhaps informal sector activity. The second productive activity has constant returns to scale, with output given by $a\ell_{2t}$. The output from this activity is taxable by the dictator and can be interpreted as modern or formal sector output. Since $\ell_{1t} + \ell_{2t} = 1$, we can simplify notation by letting $\ell_{2t} \equiv \ell_t$ so that $\ell_{1t} = 1 - \ell_t$, and further simplify without loss of generality by choosing units so that $a = 1$.

Each member of the public chooses ℓ_t and ℓ_{t+1} to maximize her discounted sum of expected utility:

$$\underset{\ell_t, \ell_{t+1}}{\text{Max}} \quad V^P$$

$$= u_1(\ell_t) + u_2(c_t) + \delta E[u_1(\ell_{t+1}) + u_2(c_{t+1})]$$

s.t. $c_t = (1 - \tau_t)\ell_t + f(1 - \ell_t) - T/n$, $c_{t+1} = (1 - \tau_{t+1})\ell_{t+1} + f(1 - \ell_{t+1}) - T/n$,

where

u_1 measures utility from consumption of the public good, $u_1' > 0$, $u_1'' < 0$;

u_2 measures utility from consumption of the private good, $u_2' > 0$, $u_2'' < 0$, $u_2'(0) = \infty$;

c is consumption of the private good;

δ is the rate at which the agent discounts the future $(0 < \delta < 1)$;

E is the expectation operator;

τ is the dictator's tax rate; and

---

[5]The "less developed polity" assumption imposes some important limitations on the ability of my model to explain corruption. Since the public has no say in whether or not the decision-maker remains in power, he has no need to "buy off" the public through patronage jobs, for example, leaving one of the most important forms of corruption unexplained. Avoiding such limitations will require a more comprehensive and ambitious approach than I have taken here.

T is explicit tax revenue (assumed to be constant over time; see discussion below).
Since the public has no choice regarding consumption of the public good, for expositional
simplicity the argument of u_1 is not specified at this time.

It is clear that the public's decision is completely time-separable and that the
first-order condition for a maximum in every time period is given by

$$1 - \tau = f'(1 - \ell),$$

yielding the implicit function $\ell^*(\tau)$. Clearly the value of output is maximized when ℓ is
chosen so that $f'(1 - \ell) = 1$, so any $\tau > 0$ constitutes a distortionary tax.

The deputies

There are N deputies, where N is a small integer. Each deputy chooses what
percentage z of the tax revenue under his control to allocate to procurement of the public
good, the remainder being spent on private goods for himself. If he is caught "shirking" he
is fired.[6]

I make a number of assumptions that simplify the structure of the deputies' problem
without affecting the qualitative nature of the results of the model. First, each deputy
controls tax revenue only in the first period of his life. In the second period of his life, if he
was not fired he receives a fixed retirement compensation (e.g., a pension) R, while if he
was fired he becomes a member of the public. R is assumed to be set at a level such that
the deputy is worse off if he is fired. The purpose of these assumptions is to capture as
simply as possible the tradeoff the deputy might face when engaging in petty corruption.
Second, when he is employed by the government the deputy is assumed to be at his "bliss
point" with regard to consumption of the public good. If we again use the example of
transportation infrastructure for the public good, we might interpret this assumption to
mean that the deputy has access to a private car and chauffeur supplied by the government
for short trips and to a private helicopter or jet for longer trips. Obviously in the real

---

[6]Klitgaard (1988) notes that the penalty for being caught engaging in corrupt activity is typically
dismissal; fines and/or prison terms are very rare.

world this assumption is much more accurate for the dictator himself (to whom it will also apply) than to his deputies. The reason I make it for the latter is to avoid complicating their decisions by letting their choices of z influence their public as well as their private good consumption. Third, I assume that the deputies' maximized expected utility V^{d*} exceeds that of the public V^{P*} so that every agent would prefer to be a deputy rather than a member of the public, and that deputy positions are not rationed on the basis of any characteristic of importance for their behavior in the model. In a later section of this paper I discuss modification of these assumptions to allow for some self-selection into the occupation of deputy.

I denote by λ the probability with which the deputy is fired. λ is a function not only of z but also of e, the percentage of the one unit of time with which the dictator is endowed that he spends supervising his deputies. I assume that the function λ has the following properties:

(1) $\lambda(z,0) = 0 \; \forall \; z$: If the dictator is not supervising the deputies the probability of any deputy being fired is zero.

(2) $\lambda(1,e) = 0 \; \forall \; e$: If the deputy does not engage in any petty corruption his probability of being fired is zero.

(3) $\partial\lambda/\partial z < 0 \; \forall \; e > 0$: The deputy's probability of being fired is decreasing in the amount of tax revenue he allocates to his assigned task (increasing in the amount of petty corruption in which he engages).

(4) $\partial\lambda/\partial e > 0 \; \forall \; z < 1$: The deputy's probability of being fired is increasing in the amount of time the dictator spends supervising him.

(5) $\lambda(0,1) \leq 1$: The maximum probability of being fired is less than or equal to one.

(6) $\partial^2\lambda/\partial z \partial e < 0$: The greater the dictator's supervisory effort, the more discriminating he is in the sense that a given increase in z causes a greater change (decrease) in the probability with which the deputy is fired.

(7) $\partial^2\lambda/\partial z^2 \geq 0$: The marginal effect of z in reducing the probability of being fired is

(weakly) diminishing.

(8) $\partial^2\lambda/\partial e^2 \leq 0$: The marginal effect of e in increasing the probability of being fired is

(weakly) diminishing.

A function $\lambda(z,e)$ that satisfies all of the above properties is $(1 - z)^\alpha e^\beta$, $\alpha \geq 1$, $\beta \leq 1$.

We can now write the problem of any deputy as

$$\underset{z_t}{\text{Max } V^d}$$

$$= \hat{u}_1 + \delta E u_1(_{t+1}) + u_2[(1-z_t)(T-s_t NR)/N - T/n]$$

$$+ \delta\{[1 - \lambda(z_t,e_t)]u_2(R-T/n) + \lambda(z_t,e_t)Eu_2[(1 - \tau_{t+1})\ell^*_{t+1} + f(1 - \ell^*_{t+1}) - T/n]\},$$

where $\hat{u}_1$ is the utility obtained from the bliss level of public good consumption and s is the percentage of deputies from the previous period who "survived" to receive their retirement compensation R (s_t is determined simply by dividing the number of deputies who were not fired in the previous period by N). Note that we have substituted the private goods consumption constraints into u_2, building in the assumption that the amount of tax revenue under each deputy's control is simply one Nth of the tax revenue available after payment of retirement compensation to all eligible former deputies, and incorporated the optimal choice the deputy would make as a member of the public. We also used the fact that in the second period of his life the deputy enjoys the same consumption of the public good regardless of whether he is fired. For now I treat the number of deputies N and their retirement compensation R as institutional givens; we can explore the consequences of varying them later on.

The optimal choice of z_t for every deputy is given by the first-order condition

$$u_2'[(1-z_t)(T-s_t NR)/N - T/n](T-s_t NR)/N$$

$$= -\delta[\partial\lambda(z_t,e_t)/\partial z_t]\{u_2(R-T/n) - Eu_2[(1 - \tau_{t+1})\ell^*_{t+1} + f(1 - \ell^*_{t+1}) - T/n]\},$$

which yields the implicit function $z^*_t(e_t;R,...)$ where e_t is assumed to take on values such that the first-order condition can be satisfied with equality by $z_t \in [0,1]$. This in turn determines the amount of the public good that is actually provided to be $z^*_t(T-s_t NR)/q$.

The utility obtained from consumption of the public good by a representative member of the public is then $u_1[z_t^*(T-s_tNR)/qh(\nu_t)]$, where ν measures the "distance" between this agent's ideal type of public good and the type actually supplied and h is a "compensation function" that is monotonically increasing and satisfies $h(0) = 1$.[7]

The dictator

The dictator is chosen from individuals in the second (last) rather than the first period of their lives. I assume that the maximum utility obtained by the dictator in the second period of his life always exceeds that obtainable by a deputy or the public. Every agent then prefers to be the dictator rather than a deputy or a member of the public in the second period of his life, so that in the absence of internal promotion the dictator is in effect chosen randomly from the population. In particular, he is not selected (by the other agents or by himself) on the basis of his ideal type of public good or his ph. As mentioned above, the dictator chooses how much of his time endowment of one unit to allocate to supervision of his deputies. I assume that the rest of his time is allocated to devising ways to enrich himself at the expense of the public. The most simple formulation is to set $\tau_t = 1 - e_t$, so that we can replace $1 - \tau_t$ with e_t in the problems of the public and the deputies above.

Since the dictator can make society's choices concerning public goods and thus exercise effective power, he is unlike the public or the deputies in that his ph can affect his behavior. In order to know how it does so we need to be able to measure the exercise of effective power, and to do this we need in turn to describe the dictator's preferences over expenditure on the public good. We can think of these preferences as divided into two parts: the amount of expenditure (constrained by explicit tax revenue) and its allocation (the choice of type). I finesse the question of whether a dictator interested in power would

---

[7]Here I have modeled preferences over the differentiated product exactly as in Helpman (1981).

always choose the maximum possible expenditure (the existence of real-world autocrats who enthusiastically implement fiscally conservative programs suggests not) by assuming that explicit tax revenue is fixed exogenously at T (conceivably at a maximum set by administrative capacity). Thus the dictator only has the power to choose the type of public good supplied by the government, and we can conveniently measure the extent to which his preferences over the public good are made effective by the percentage of T that is actually used to produce the type of public good he has chosen rather than winding up in the pockets of the deputies. Any measure of exercise of effective power should therefore be increasing in z. I denote such a measure by $P(z)$, where $P' > 0$ and $P'' \leq 0$.[8]

The dictator's ph is measured by the random variable η, which is drawn from a distribution Ψ that maps $[\underline{\eta}, \infty)$ into $[0,1]$, where $\underline{\eta}$ is defined below. In the absence of internal promotion Ψ is the same as the fixed population distribution ϕ. If we assume that the dictator acts as a Stackelberg leader, optimizing with respect to the functions ℓ^* and z^*, we can write his problem as

$$\underset{e_t}{\text{Max }} V^D = \hat{u}_1 + u_2[n(1-e_t)\ell_t^*] + \eta P(z_t^*).$$

The optimal choice of e_t is given by the first-order condition

$$u_2'[n(1-e_t)\ell_t^*]n[\ell_t^* - (1-e_t)d\ell^*/de_t] = \eta P'(z_t^*)(dz_t^*/de_t),$$

which yields the implicit function $e_t^*(\eta;...)$ where η is assumed to take on values such that the first-order condition can be satisfied with equality. This assumption implicitly defines $\underline{\eta}$ for both this section of the paper and the next.

Let us turn now to the key elements of behavior implied by the first-order conditions derived above. I will stick to intuition and confine all derivations to the Appendix.

---

[8]I have chosen to keep the specification of the function P as simple as possible. Undoubtedly it should contain more arguments. For example, it seems reasonable that P should be increasing in n, the number of people over whom power is being exercised, though this is unimportant in the present context since n is being treated as a constant.

The public's behavior is easiest to analyze. Obviously the amount of labor the public allocates to marketed, taxable production as opposed to subsistence/informal sector, untaxable production will depend negatively on the tax rate and hence positively on the effort the dictator spends supervising his deputies rather than plundering the public. In the Appendix it is shown that $d\ell^*/de > 0$. This behavior gives rise to an inverted-U relationship between the dictator's supervision effort and his income (implicit tax revenue): when effort is zero, so that all marketed output is taxed away, then the public devotes zero labor to taxable output and tax revenue is zero; but when effort is one, the tax rate is zero and tax revenue is again zero. Under a condition derived in the Appendix there will be only a single local (equals global) maximum for tax revenue for $0 < e < 1$. I denote this income-maximizing effort level by $\hat{e}$.

For the deputies, more intense supervision by the dictator means that their probability of being fired increases more steeply with the percentage of tax revenue under their control that they divert into their own pockets ($\partial^2\lambda/\partial z\partial e < 0$). This leads them to decrease this percentage or increase the percentage that they apply to provision of the public good. In the Appendix it is shown that $dz^*/de > 0$.

Turning finally to the behavior of the dictator, we note that given our (implicit) definition of $\underline{\eta}$ it must be at least as great as the value of η for which the dictator's first-order condition is satisfied with equality when $e_t = \hat{e}$. This rules out the existence of a *range* of (low) values of η for which the dictator always chooses effort level $\hat{e}$ (though it allows for non-negligible probability mass at $\underline{\eta}$). Intuitively, in combination with our assumption of a single local (equals global) maximum for $(1 - e)\ell^*$, this insures that the dictator always faces a tradeoff between income and exercise of effective power. In the Appendix it is shown that $de^*/d\eta > 0$ provided that the second-order condition for a maximum is satisfied at e^*. Thus the greater is the dictator's ph, the better off is the public, which is subjected to a lower rate of distortionary taxation and receives a greater supply of the public good (because $dz^*/de > 0$) in return for the tax revenue that is

collected from it in a lump-sum fashion.[9]

Note that since $\hat{e} > 0$ it is never optimal for the dictator to spend zero effort on supervision of his deputies. If for $e_t = \hat{e}$ the deputies' first-order condition is satisfied for $z_t > 0$, this means that the government always supplies a positive amount of the public good. I leave it to the reader to decide whether this would be a realistic feature of the model.

I conclude this section by discussing the effects of two "policies" that might be adopted with the aim of improving bureaucratic performance: higher salaries for deputies and meritocratic recruitment of deputies. The first policy can be modeled as an increase in R. One would expect deputies to devote more resources to their assigned tasks and less to petty corruption if the penalty for being fired is effectively raised by increasing their (retirement) compensation. The only opposing consideration is that increasing R reduces the amount of tax revenue under each deputy's control and thus increases his marginal utility of income and his incentive to be corrupt. The condition under which $dz^*/dR > 0$ is satisfied is stated in the Appendix. Assuming this condition holds, there exists a tradeoff for the public regarding deputies' compensation, holding constant the effort level of the dictator: increasing R reduces the resources *available* to purchase the public good, but increases the share of the remaining resources that are *actually used* to purchase the public good.

Matters become still more complex and ambiguous when we take account of the fact that the dictator's effort level e^* will in general be affected by changes in R. These effects are of two types, which I will call the "concavity effect" and the "responsiveness effect". If there is diminishing marginal utility from exercise of effective power (as there will be in our

---

[9]Since η is a random variable, one could interpret an observation of what appear to be many consecutive low draws (e.g., many decades of predatory rule in Haiti) as evidence against the relevance of the model. A possible response is that such behavior is consistent with the model if dictators are not being drawn from "the public" but rather from an ethnically separate minority for which "the public", and therefore public goods, has no standing.

formulation if $P'' < 0$), then an increase in z_t^* that is exogenous to the dictator will cause him to reduce e_t: his effort is not as urgently needed to insure that his preferences over the public good are being effectively implemented. This concavity effect works to offset any positive effect a change in R might have on public welfare. At the same time, an increase in R not only causes the deputies to increase z_t^* to reduce their risk of being fired, but also causes them to be more responsive to changes in that risk caused by changes in the dictator's effort level. Thus (provided the condition given in the Appendix is satisfied) dz_t^*/de_t is increasing in R, which will cause the dictator to increase e_t: at the margin his supervisory effort is more effective in bending the deputies' behavior to his will. This responsiveness effect reinforces the positive effect of R on z_t^* and also reduces the dictator's implicit taxation of the public. Taking all effects of a change in R into account, we are left with very little predictive power for the effect of increasing deputies' compensation on the welfare of the public, though there remains a presumption that it will reduce petty corruption.[10]

I now sketch out an analysis of the effects of meritocratic recruitment of deputies. It is of course necessary to drop our assumption that all agents have identical ability as measured by their time endowment and productivity. If it is still the case that every agent strictly prefers being a deputy to being a member of the public, then it should in theory be possible to recruit the highest ability agents to government service through, for example, a competitive examination system. Suppose now that the higher the ability of the deputy, the more efficiently he is able to procure the public good, which we can measure by a

---

[10]It is possible, however, to construct a special case in which the effect of increasing deputies' compensation on the welfare of the public is unambiguously positive. First, note that it would be reasonable, given our argument why any measure of exercise of effective power should be increasing in z, to include $(T-s_t NR)$ as a coefficient on z in the function P. Recalling that T, s_t, N, and R are either exogenous by assumption or determined by actions of the previous dictator, we can then write $P[z_t^*(T-s_t NR)]$ in our expression for V^D. Now suppose that R is at a level such that $(d/dR)[z_t^*(T-s_t NR)] = 0$. In this case the concavity effect washes out, as does the effect of R on the welfare of the public through its effect on the deputies' behavior, leaving only the responsiveness effect.

reduction in q. In itself this does not affect the determination of z_t^*, and thus it would seem that, holding e_t constant, the public benefits unambiguously from meritocratic recruitment through reductions in the average q and consequent greater provision of the public good. However, presumably the higher ability of deputy j not only reduces q_j but also increases his income as a member of the public, which weakens his incentive to avoid being fired and thus reduces z_{jt}^*. Thus the effect of meritocratic recruitment on the welfare of the public is ambiguous after all. This ambiguity can be reduced or eliminated if being fired from government service results in a stigma that prevents the former deputy from making full use of his ability in the private sector, or if use of that ability in the private sector would have required accumulation of sector-specific human capital that was foregone during his career as a deputy.

Again, matters become more complex and ambiguous when we account for the fact that the dictator's effort level e^* will in general be affected by the imposition of meritocratic recruitment. There is an effect analogous to the responsiveness effect: there is a greater payoff in terms of exercise of effective power to getting higher ability deputies to allocate more funds to purchase of the public good, causing the dictator to increase e_t. However, this is offset by a negative responsiveness effect if higher ability deputies have better private sector opportunities when they are fired. In short, the only presumption one can establish for the effects of meritocratic recruitment on bureaucratic performance is that whatever funds are actually allocated to provision of the public good will be used more efficiently.

IV. The model with internal promotion

A. *The case of one deputy*

We now suppose that a bureaucratic institutional structure exists such that there is "promotion from within", meaning that the next dictator is chosen from the pool of deputies who are not fired. If all deputies are fired, the next dictator is chosen randomly

from the population as in the previous section. As in Soskice et al. (1992), internal promotion is assumed to be an institutional feature beyond the control of the dictator. The analysis is greatly simplified if we initially assume that there is only one deputy (N = 1). (This assumption is also made by Soskice et al.)

Clearly the existence of internal promotion does not affect the behavior of the public derived in the previous section. For the dictator, since z^* is now a function of the deputy's ph (as we will show immediately below) and is therefore a random variable, we must replace P in the dictator's problem with EP, where the expectation is computed using the density function ϕ associated with the fixed population distribution Φ that is defined in the previous section. Carrying this change through to the dictator's first-order condition, it is clear that the key result $de^*/d\eta > 0$ is unaffected. Internal promotion does, of course, qualitatively change the behavior of the deputy. Let the deputy's ph be measured by the random variable ξ. We can now write the deputy's problem as

$$\operatorname*{Max}_{z_t} V^d$$

$$= \hat{u}_1 + u_2[(1-z_t)(T-s_t NR)/N - T/n]$$

$$+ \delta\{[1 - \lambda(z_t,e_t)]\{\hat{u}_1 + u_2[n(1-e^*_{t+1})\ell^*_{t+1}] + \xi EP(z^*_{t+1})\} +$$

$$\lambda(z_t,e_t)\{Eu_1[z^*_{t+1}(T-s_{t+1}NR)/qh(\nu_{t+1})] + Eu_2[e_{t+1}\ell^*_{t+1} + f(1 - \ell^*_{t+1}) - T/n]\}\},$$

where e^*_{t+1} is the effort the current deputy will expend on supervision of his future deputy in the event that he becomes dictator, and all expectations are computed using the density function ϕ (and with respect to the population distribution of the ideal type of public good where necessary) since the ph of the next deputy is drawn from the fixed population distribution Φ, as is that of the next dictator if the current deputy is fired. We can simplify slightly by noting that N = 1 and that, because there is only one deputy who either becomes dictator or is fired, there are never any retired deputies around to receive compensation so that s = 0.

The deputy's optimal choice of z_t is given by the first-order condition

$$u_2'[(1-z_t)T - T/n]T = -\delta[\partial\lambda(z_t,e_t)/\partial z_t]$$
$$\times \{\hat{u}_1 + u_2[n(1-e_{t+1}^*)\ell_{t+1}^*] + \xi EP(z_{t+1}^*)$$
$$- \{Eu_1[z_{t+1}^*T/qh(\nu_{t+1})] + Eu_2[e_{t+1}\ell_{t+1}^* + f(1 - \ell_{t+1}^*) - T/n]\}\},$$

which yields the implicit function $z_t^*(e_t;\xi,...)$ where $e_t \geq \hat{e}$ insures that the first-order

condition can be satisfied with equality by $z_t \in [0,1]$, as was also assumed in the previous

section. Unlike in the previous section, the deputy's ph now affects his behavior: the

higher his ph, the more he values promotion to dictator, and the more funds he devotes to

implementing the current dictator's preferences over the public good to avoid being fired.

In the Appendix it is shown that $dz^*/d\xi > 0$. Of course $dz^*/de > 0$ as in the previous

section.

We see that the institution of internal promotion has introduced an element of

selectivity into the drawing of the dictator with regard to ph. The deputy is more likely to

become dictator, the higher his ph, and only if he is fired is the dictator drawn randomly

from the population distribution of ph.[11] This selectivity breaks the identity between Ψ,

the dictator's distribution of ph, and ϕ, the population distribution of ph. What can we

say about Ψ now?

To begin, fix the ph of the period t dictator at η. It can then be shown that

$$\psi_{t+1}^{\circ}(\hat{\eta}) = \phi(\hat{\eta})\{\{1 - \lambda\{z_t^*[e_t^*(\eta),\hat{\eta}],e_t^*(\eta)\}\} + \int_{\eta}^{\infty}\phi(a)\lambda\{z_t^*[e_t^*(\eta),a],e_t^*(\eta)\}da\},$$

where $\psi_{t+1}^{\circ}(\hat{\eta})$ is the probability that the dictator in period t+1 will have ph = $\hat{\eta}$.

Intuitively, a dictator of type $\hat{\eta}$ could be chosen in one of two ways: a deputy of type $\hat{\eta}$

could be chosen and promoted, or a deputy of any type could be fired and then a dictator

---

[11]The reader might think that if we allowed agents to differ with respect to "greed" that since the greedy more highly value promotion to dictator they would choose a higher z and be less likely to be fired. This would generate selection in favor of greed when drawing the dictator. However, the same greed also causes these agents to desire current income more intensely, leading them to choose a lower z. It turns out that if we measure greed by a parameter θ that multiplies u_2 and thus shifts the marginal utility of income, it can be shown (see Appendix) that $dz/d\theta < 0$ in the internal promotion case. The reason is that a high θ makes the benefits of dictatorship other than income (the bliss level of public good consumption and the exercise of effective power) relatively less important.

of type $\hat{\eta}$ could be chosen. Given η, the probabilities of the first and second events are given by the first and second terms, respectively, of the equation for $\psi^{\bullet}_{t+1}(\hat{\eta})$. Now we simply relax the assumption of a given ph for the period t dictator and obtain

$$\psi_{t+1}(\hat{\eta}) = \phi(\hat{\eta})\int_{\eta}^{\infty}\psi_t(\eta)\{\{1 - \lambda\{z^*_t[e^*_t(\eta),\hat{\eta}],e^*_t(\eta)\}\} + \int_{\eta}^{\infty}\phi(a)\lambda\{z^*_t[e^*_t(\eta),a],e^*_t(\eta)\}da\}d\eta,$$

where $\int_{\eta}^{\eta}\psi_t(b)db = \Psi_t(\eta)$ and $\psi_0 = \phi$ if we define period 0 as the period in which internal promotion is first instituted. In the Appendix it is shown that Ψ "improves" monotonically with time in the sense of first-order stochastic dominance: $\int_{\eta}^{\eta}\psi_t(b)db \geq \int_{\eta}^{\eta}\psi_{t+1}(b)db$, t = 0,1,.... It is also shown that as t goes to infinity ψ_t must converge to ψ^*, where ψ^* is defined by

$$\psi^*(\hat{\eta}) = \phi(\hat{\eta})\int_{\eta}^{\infty}\psi^*(\eta)\{\{1 - \lambda\{z^*[e^*(\eta),\hat{\eta}],e^*(\eta)\}\} + \int_{\eta}^{\infty}\phi(a)\lambda\{z^*[e^*(\eta),a],e^*(\eta)\}da\}d\eta.$$

The intuition is that a dictator with a higher ph will spend more effort on supervision of his deputy and therefore be more discriminating in his firing (promotion) decision, so that the selection process displays positive feedback. Thus the longer the institutional practice of internal promotion is in place, the greater is the expected ph of the dictator, with all the consequences that entails for the expected welfare of the public.

Apart from the tendency it creates to choose higher ph dictators, the institution of internal promotion clearly creates a greater incentive on the part of the deputy to avoid being fired and thus choose a high z_t. This behavior induces offsetting concavity and responsiveness effects on the part of the dictator of the type analyzed in section III.

I conclude this subsection with a discussion of the impact of meritocratic recruitment in the presence of internal promotion. Provided that every agent, regardless of ability, still strictly prefers to be the deputy rather than a member of the public, there is no *qualitative* change in the analysis of meritocratic recruitment made at the end of section III. The possibility of a qualitative change in the analysis arises if some high ability individuals earn so much in the private sector that they might prefer not to be deputies.

Among this set of individuals, those with higher ph are more likely to choose a career in government service in the hope that they will be promoted to a position in which they can exercise effective power. Restriction of recruitment to this set of agents might thus act to select for what one would conventionally call "idealism" in the deputy, and this selection would then interact with the selection for high ph that we have already shown occurs with internal promotion. A complete analysis of this potential synergy between internal promotion and meritocratic recruitment would require a much more complete modeling of the heterogeneity of agents' abilities, the returns to those abilities in the private and public sectors, and of the recruitment process itself. Rather than pursue this, I turn in the next subsection to the more manageable analysis of selection for ph that might occur *within* the government when we extend the model to allow for more than one deputy.

B. The case of more than one deputy

Extension of the model to allow for more than one deputy complicates the selection mechanism described in the previous subsection. To focus the analysis, note that the strength of selection for ph induced by internal promotion is limited, despite the existence of positive feedback, by the lack of any self-selection to become a deputy. Might we observe stronger selection mechanisms under certain circumstances? The answer is yes, and it turns out that such circumstances may also involve more variance in corruption across state functions.

The most straightforward but least interesting way to include more than one deputy in a model with internal promotion is to suppose that (1) each deputy has the same amount of revenue under his control, (2) each is subject to the same level of supervision by the dictator, and that (3) each of the deputies who is not fired has an equal chance of being promoted to dictator, with those who are neither fired nor promoted receiving the retirement compensation R. What would be the impact of this structure on selection for ph relative to the one-deputy case? On the one hand, selection is weakened because each deputy realizes he may not get an opportunity to exercise effective power even if he is not

fired and thus his ph has a smaller positive impact on his choice of z_t than in the one-deputy case. On the other hand, there will in general be greater variation in ph across deputies than in the one-deputy case, improving the selection process by giving it (so to speak) more material with which to work. Regardless, note that the levels of large-scale and petty corruption continue to move together as in the one-deputy case.

More interesting possibilities arise if we allow for variation across deputies in items (1) - (3) above. In particular, consider the case in which there is one department of the government in which future dictators are "groomed", so that deputies in other departments have no chance of being promoted to dictator. To fix ideas, assume that there is only one deputy in each department and only two departments, A and B, where A is the department from which the deputy can be promoted to dictator. If there are no other differences across departments than every agent will strictly prefer to apply to department A and selection proceeds exactly as in the one-deputy case. But suppose that department A differs from department B in at least one of two respects: (i) Department A handles less funds and thus offers fewer opportunities for petty corruption than department B. For example, department A could be an economic planning department that receives funds mainly for collection and analysis of statistics while department B could be the public works department. (ii) Department A is more effectively supervised than department B because the dictator is always promoted from it and thus understands better both the department's mission and the means of subverting it through petty corruption. In either case it becomes possible that some agents might strictly prefer to apply to department B, and in particular that there will be self-selection for low (high) ph on the part of applicants to department B (A).

The extra round of self-selection by applicants could itself display positive or negative feedback within the context of the entire dictator selection process, depending on how the higher (expected) supervisory effort of the dictator affects the incentives of agents to apply to department A versus department B. To build intuition, consider an artificial

situation in which the dictator's supervisory effort is constrained to be the same across both departments, and in which (contrary to the assumptions of case (ii)) the function $\lambda(z,e)$ is the same across departments. It can be shown that, as (expected) e_t increases, on the one hand the expected utility of being a department B deputy is reduced more because he is choosing a lower z_t and thus his probability of being fired is increased more (by $\partial^2\lambda/\partial z_t \partial e_t < 0$), while on the other hand the expected utility of being a department A deputy is reduced more because he has more to lose from being fired in terms of next period's utility. If the former (latter) effect dominates there is negative (positive) feedback. Now allow for $e^A \neq e^B$, and for case (i) make the natural assumption that because it handles more funds, more power is exercised through department B than A (though clearly this need not hold in general). We should then find that (in expectation) $e_t^{B*} > e_t^{A*}$ (though the responsiveness effect works in the opposite direction), and that this difference will increase absolutely as the dictator's expected ph increases, tending to yield negative feedback. In case (ii) we also allow for $\lambda^A(z,e) \neq \lambda^B(z,e)$, and make the natural assumption that $\partial\lambda^A/\partial e > \partial\lambda^B/\partial e$ for a given z. Now we should find that (in expectation) $e_t^{A*} > e_t^{B*}$, and that this difference will increase absolutely as the dictator's expected ph increases, tending to yield positive feedback.[12]

We can thus predict that in case (ii) department A will become a "pocket of efficiency", and it is likely that both petty corruption in department A and large-scale corruption will be low while petty corruption in department B will be relatively high.[13] If by chance cases (i) and (ii) should be combined, this variation in corruption across departments and state functions could yield a rather poor performance in provision of

---

[12]Note that in the case of positive feedback self-selection by applicants may not occur when internal promotion is first instituted but could still be induced later by the rise in the expected ph of the dictator over time.

[13]Note that much of department A's "efficiency" will survive a bad draw of the dictator's ph that leads to high large-scale corruption because self-selection into department A was on the basis of the dictator's *expected* ph, and because department A remains the only department from which a deputy can be promoted to dictator and is still better supervised than department B.

public goods (given that department B controls the lion's share of tax revenue) in tandem with a private sector that is relatively unburdened by rent-extracting regulation.

In a separate Appendix (available on request) I give a complete formal analysis of case (ii) when department B is a pure sinecure, i.e., $\lambda^B = 0 \; \forall \; e^B$. It is necessary to add a period to the beginning of every agent's life in which he decides to which government department he will apply (or which he will indicate as his first preference), but this is of no consequence provided that his application decision is the only one he can make during this period that affects his future. The condition under which an extra round of self-selection will occur is specified, and it is shown that when this obtains at time t the extra round of self-selection displays positive feedback from time t+1 forward.

In the case of positive feedback, someone who observes department A over time and sees the increasing tendency for the deputy to faithfully carry out the department's mission might report favorably on the evolution of its "corporate culture" (Kreps 1990).[14] Suppose that my analysis of case (ii) still goes through when there is more than one deputy in both department A and department B. An observer of the closely supervised, high ph individuals in department A might comment on its "esprit de corps" or infer the operation of a "norm", but in this case the simulation of a norm is the result of repeated selection rather than the equilibrium of a repeated game.

Taking the point of view of policy, one can ask whether the expected welfare of the public would be improved by designating a department as the one from which the dictator is always promoted where no such tradition exists. Under the extreme assumption on supervision technology where the department from which the dictator is not promoted makes zero contribution to public goods provision, the answer is unambiguously yes. If, on the other hand, there is positive supervision of department B, some ambiguity exists

---

[14]This is not to imply that in my analysis there is no place for "corporate culture" in Kreps's sense of focal principle. One can imagine, for example, that an equilibrium that sustains a favorable "corporate culture" ‮⁓⁓⁓⁓⁓⁓⁓⁓‬ because department A, being the only department to be rank supervision by the dictator, promoted to dictator and is still better supervised than department B.

because one must trade off improved selection against the loss of the promotion incentive in department B (although one also strengthens the promotion incentive in department A).

V. Conclusions

This paper has argued that, in the absence of effective political oversight, internal promotion is a crucial institutional feature generating better performance from powerful state bureaucracies (the most powerful of which could include the nation's chief executive). Internal promotion acts to select for desire to exercise effective power at the top of the hierarchy, with these individuals in turn restraining the corruption of their subordinates as a byproduct of their efforts to implement their preferences over public goods using tax revenue, and also finding less time to carry out large-scale corruption of their own. If one accepts this basic argument, a number of predictions and policy opportunities related to details of institutional design that influence selection present themselves. It should be noted, however, that while internal promotion may lead on average to a greater share of the tax revenue collected to finance provision of public goods being spent for its intended purpose, it does not create any greater tendency for the mix of public goods supplied to match the preferences of the public.

A secondary aim of this paper has been to disaggregate "corruption" into two categories that are intended to be mutually exclusive but not exhaustive: creation and appropriation of rents, and misappropriation of tax revenue. In my model the levels of both types of corruption tend to move together, but in cases where they do not the predictions of the model for the former type (what I called large-scale corruption) are more important in terms of the impact of corruption on economic performance as measured by the usual aggregates. Since the privately produced measures of bureaucratic corruption cited in the introduction are intended to serve the needs of transnational investors, it is likely that they relate more to large-scale than petty corruption. Petty corruption is more likely to influence other privately produced ratings such as "quality of infrastructure".

Future empirical work will show whether bureaucratic characteristics such as the extent of internal promotion will influence these variables in the manner predicted here.

APPENDIX

Proof that $d\ell*/de > 0$: The first-order condition of the public can be written as
$e = f'(1 - \ell)$, implicitly determining $\ell^*(e)$. We then have $de = - f''d\ell^*$, or
$d\ell^*/de = - 1/f'' > 0$.

Condition insuring no more than one local maximum for $(1 - e)\ell*$: A sufficient
condition is that the derivative of $(1 - e)\ell^*$ is always decreasing. This condition can be
expressed in terms of derivatives of the production function f. We have $(d/de)[(1 - e)\ell^*] =$
$(1 - e)d\ell^*/de - \ell^* = (1 - e)(-1/f'') - \ell^*$; $(d/de)^2[(1 - e)\ell^*] = 1/f'' + (1 - e)f'''/(f'')^2 -$
$d\ell^*/de = 2/f'' + (1 - e)f'''/(f'')^2$. The condition can therefore be written as
$2f'' + (1 - e)f''' < 0$ or $2f'' + (1 - f')f''' < 0$.

Proof that $dz_t^*/de_t > 0$: We differentiate the deputy's first-order condition and
rearrange to get $\{- u_2'[(T-s_tNR)/N]^2 +$
$\delta(\partial^2\lambda/\partial z_t^2)\{u_2(R-T/n) - Eu_2[(1 - \tau_{t+1})\ell_{t+1}^* + f(1 - \ell_{t+1}^*) - T/n]\}\}dz_t^* =$
$- \delta(\partial^2\lambda/\partial z_t\partial e_t)\{u_2(R-T/n) - Eu_2[(1 - \tau_{t+1})\ell_{t+1}^* + f(1 - \ell_{t+1}^*) - T/n]\}de_t$. Conditions
(6) and (7) on λ, $u_2' < 0$, and the assumption that the deputy is worse off if fired insure
that the coefficients on both dz_t^* and de_t are positive.

Proof that $de_t^*/d\eta > 0$: We can rewrite the dictator's first-order condition as
$\eta P'(z_t^*)(dz_t^*/de_t) + u_2'n(d/de_t)[(1 - e_t)\ell_t^*] = 0$. Clearly this condition can only be satisfied
if the second term is negative (if implicit tax revenue is increasing in τ_t). Differentiating
and rearranging, we obtain $\{\eta P''(dz_t^*/de_t)^2 + P'(d^2z_t^*/de_t^2) + u_2'\{n(d/de_t)[(1 - e_t)\ell_t^*]\}^2$
$+ u_2'n(d/de_t)^2[(1 - e_t)\ell_t^*]\}de_t = - P'(dz_t^*/de_t)d\eta$. The last two terms in the coefficient on
de_t are negative and the first term is nonpositive, while the second term is unsigned. The
second term must be nonpositive or dominated both for the second-order condition for a
maximum to hold and for $de_t^*/d\eta > 0$ to hold.

Condition under which $dz_t^*/dR > 0$: We again differentiate the deputy's first-order
condition and rearrange to get $\{- u_2'[(T-s_tNR)/N]^2 +$

$\delta(\partial^2\lambda/\partial z_t{}^2)\{u_2(R-T/n) - Eu_2[(1 - \tau_{t+1})\ell^*_{t+1} + f(1 - \ell^*_{t+1}) - T/n]\}\}dz^*_t =$

$[- \delta(\partial\lambda/\partial z_t)u'_2 + u'_2 s_t + u'_2{}'(1 - z_t)s_t(T-s_t NR)/N]dR.$ We have already seen that the coefficient on dz^*_t is positive. For the coefficient on dR to be positive the term with $u'_2{}'$ must be dominated.

Condition under which $(d/dR)(dz_t^*/de_t)$ (the "responsiveness effect") is positive: We have $dz^*_t/de_t =$

$- \delta(\partial^2\lambda/\partial z_t \partial e_t)\{u_2(R-T/n) - Eu_2[(1 - \tau_{t+1})\ell^*_{t+1} + f(1 - \ell^*_{t+1}) - T/n]\} \div$

$- u'_2{}'[(T-s_t NR)/N]^2$

$+ \delta(\partial^2\lambda/\partial z_t{}^2)\{u_2(R-T/n) - Eu_2[(1 - \tau_{t+1})\ell^*_{t+1} + f(1 - \ell^*_{t+1}) - T/n]\}\}.$ Clearly the numerator is increasing in R. The first term in the denominator appears to be decreasing in R but is in fact ambiguous because $u'_2{}'$ is unsigned; the second term in the denominator is nondecreasing in R since $\partial^2\lambda/\partial z_t{}^2 \geq 0$. The increase in the numerator must dominate any increase in the denominator.

Proof that $dz_t^*/d\xi > 0$: We differentiate the deputy's new first-order condition and rearrange to get $\{- u'_2{}'T^2 + \delta(\partial^2\lambda/\partial z_t{}^2)\{\hat{u}_1 + u_2[n(1-e^*_{t+1})\ell^*_{t+1}] + \xi EP(z^*_{t+1})$

$- Eu_1[z^*_{t+1}T/qh(\nu_{t+1})] - Eu_2[e_{t+1}\ell^*_{t+1} + f(1 - \ell^*_{t+1}) - T/n]\}\}dz^*_t =$

$- \delta(\partial\lambda/\partial z_t)EP(z^*_{t+1})d\xi.$ The coefficients on dz^*_t and $d\xi$ are unambiguously positive.

Proof that $dz_t^*/d\theta < 0$: The deputy's first-order condition can now be rewritten as

$u'_2[(1-z_t)T - T/n]T = -\delta[\partial\lambda(z_t,e_t)/\partial z_t]\{\hat{u}_1/\theta + u_2[n(1-e^*_{t+1})\ell^*_{t+1}] + \xi EP(z^*_{t+1})/\theta$

$- \{Eu_1[z^*_{t+1}T/qh(\nu_{t+1})]\}/\theta + Eu_2[e_{t+1}\ell^*_{t+1} + f(1 - \ell^*_{t+1}) - T/n]\}\},$ where we have divided through both sides by θ. It is clear that θ has the opposite effect on z^*_t from ξ.

Proof that $\Psi_t(\eta) = \int_{\underline{\eta}}^{\eta}\psi_t(b)db \geq \int_{\underline{\eta}}^{\eta}\psi_{t+1}(b)db = \Psi_{t+1}(\eta)$, $t = 0,1,...$, and that ψ_t must converge to $\psi*$, where $\psi*$ is defined by

$\psi*(\hat{\eta}) = \phi(\hat{\eta})\int_{\underline{\eta}}^{\infty}\psi*(\eta)\{\{1 - \lambda\{z*[e*(\eta),\hat{\eta}],e*(\eta)\}\} + \int_{\underline{\eta}}^{\infty}\phi(a)\lambda\{z*[e*(\eta),a],e*(\eta)\}da\}d\eta.$[15]

---

[15]I am indebted to Joel Sobel for help with this proof.

We begin by noting that

$$\psi_{t+1}(\hat{\eta}) = \phi(\hat{\eta})\int_{\underline{\eta}}^{\infty}\psi_t(\eta)\{\{1 - \lambda\{z_t^*[e_t^*(\eta),\hat{\eta}],e_t^*(\eta)\}\} + \int_{\underline{\eta}}^{\infty}\phi(a)\lambda\{z_t^*[e_t^*(\eta),a],e_t^*(\eta)\}da\}d\eta,$$

$$= \phi(\hat{\eta}) + \int_{\underline{\eta}}^{\infty}\phi(\hat{\eta})\psi_t(\eta)\{\int_{\underline{\eta}}^{\infty}\phi(a)\lambda\{z^*[e^*(\eta),a],e^*(\eta)\}da - \lambda\{z^*[e^*(\eta),\hat{\eta}],e^*(\eta)\}\}d\eta,$$

$$\tag{A1}$$

where we have dropped the time subscript everywhere except on ψ because all of the other *functions* are invariant with respect to time. Define

$$A(\hat{\eta},\eta) \equiv \int_{\underline{\eta}}^{\hat{\eta}}\phi(b)\{\int_{\underline{\eta}}^{\infty}\phi(a)\lambda\{z^*[e^*(\eta),a],e^*(\eta)\}da - \lambda\{z^*[e^*(\eta),b],e^*(\eta)\}\}db.$$

Integrating both sides of (A1) from $\underline{\eta}$ to $\hat{\eta}$, we have

$$\Psi_{t+1}(\hat{\eta}) = \phi(\hat{\eta}) + \int_{\underline{\eta}}^{\infty}\psi_t(\eta)A(\hat{\eta},\eta)d\eta. \tag{A2}$$

First we want to show that $\Psi_1(\hat{\eta}) \leq \Psi_0(\hat{\eta}) \equiv \phi(\hat{\eta})$, i.e., that the institution of internal promotion improves the distribution of the dictator's ph even without positive feedback. Intuitively, this should hold because a higher ph deputy is more likely to be promoted: $\partial\lambda/\partial\hat{\eta} = (\partial\lambda/\partial z)(dz^*/d\hat{\eta}) < 0$. The proof uses this fact to establish that $A(\hat{\eta},\eta) < 0$ for $\underline{\eta} < \hat{\eta} < \infty$, which is sufficient to show that $\Psi_{t+1}(\hat{\eta}) \leq \phi(\hat{\eta})$ by (A2). First we note that $A(0,\eta) = A(\infty,\eta) = 0$. Then we show that $A(\hat{\eta},\eta)$ cannot have a local maximum:

$$\partial A/\partial\hat{\eta} = \phi(\hat{\eta})\{\int_{\underline{\eta}}^{\infty}\phi(a)\lambda\{z^*[e^*(\eta),a],e^*(\eta)\}da - \lambda\{z^*[e^*(\eta),\hat{\eta}],e^*(\eta)\}\},$$

$$\partial^2 A/\partial\hat{\eta}^2 = \phi'(\hat{\eta})(\partial A/\partial\hat{\eta})/\phi(\hat{\eta}) - \phi(\hat{\eta})\partial\lambda/\partial\hat{\eta} > 0 \text{ when } \partial A/\partial\hat{\eta} = 0.$$

Next, if we can establish that $\Psi_{t+1}(\hat{\eta}) \leq \Psi_t(\hat{\eta})$ implies $\Psi_{t+2}(\hat{\eta}) \leq \Psi_{t+1}(\hat{\eta})$, then since $\Psi_1(\hat{\eta}) \leq \Psi_0(\hat{\eta})$ we will have proved that the distribution of the dictator's ph improves monotonically with time in the sense of first-order stochastic dominance.

First-differencing (A2), we obtain

$$\Psi_{t+2}(\hat{\eta}) - \Psi_{t+1}(\hat{\eta}) = \int_{\underline{\eta}}^{\infty}[\psi_{t+1}(\eta) - \psi_t(\eta)]A(\hat{\eta},\eta)d\eta. \tag{A3}$$

Integration by parts of the right-hand side of (A3) yields

$$[\Psi_{t+1}(\eta) - \Psi_t(\eta)]A(\hat{\eta},\eta)\Big|_{\eta=\underline{\eta}}^{\eta=\infty} - \int_{\underline{\eta}}^{\infty}[\Psi_{t+1}(\eta) - \Psi_t(\eta)][\partial A(\hat{\eta},\eta)/\partial\eta]d\eta, \text{ which reduces to}$$

$$\Psi_{t+2}(\hat{\eta}) - \Psi_{t+1}(\hat{\eta}) = - \int_{\eta}^{\infty}[\Psi_{t+1}(\eta) - \Psi_t(\eta)][\partial A(\hat{\eta},\eta)/\partial \eta]d\eta \qquad (A3')$$

since $\Psi_{t+1}(\underline{\eta}) = \Psi_t(\underline{\eta}) = 0$ and $\Psi_{t+1}(\infty) = \Psi_t(\infty) = 1$. If we can now show that $[\partial A(\hat{\eta},\eta)/\partial \eta] < 0$ for $\underline{\eta} < \hat{\eta} < \infty$, then (A3') yields the desired result. The proof follows *exactly* the same lines as the proof that $A(\hat{\eta},\eta) < 0$ for $\underline{\eta} < \hat{\eta} < \infty$, substituting $\partial^2 \lambda/\partial\hat{\eta}\partial\eta$ for $\partial\lambda/\partial\hat{\eta}$. Thus we need only establish that $\partial^2 \lambda/\partial\hat{\eta}\partial\eta < 0$. We have

$$\partial^2 \lambda/\partial\hat{\eta}\partial\eta = [(\partial^2 \lambda/\partial z^2)(dz^*/de) + \partial^2 \lambda/\partial z \partial e](de^*/d\eta)(dz^*/d\hat{\eta}) +$$

$$(\partial\lambda/\partial z)(d^2 z^*/d\hat{\eta}de)(de^*/d\eta). \qquad (A4)$$

It can be shown that $d^2 z^*/d\hat{\eta}de > 0$. This is just another example of the "responsiveness effect": deputies with higher ph have more to lose by being fired and thus increase z more in response to increases in the dictator's supervisory effort. Thus the responsiveness effect tends to make the dictator's effort more effective in discriminating between high and low ph deputies as it increases with time. We can therefore see that of the three terms on the right-hand side of (A4) only the first term is nonnegative. The intuition behind this ambiguity is that as the increased supervision by the dictator drives the deputy to reduce his corrupt activities, the difference in behavior between high and low ph deputies may narrow (in the most extreme case any deputy chooses $z_t = 1$ regardless of ph) so that supervision has less discriminatory effect. Assuming that the first term is dominated (it does not exist if $\partial^2 \lambda/\partial z^2 = 0$), our proof that Ψ_t improves monotonically with time in the sense of first-order stochastic dominance is complete.

It remains to be shown that $\psi_{t+1}(\hat{\eta})$ converges to $\psi^*(\hat{\eta})$ as defined above. First note that $\Psi_{t+1}(\hat{\eta})$ is a monotonic bounded sequence for each $\hat{\eta}$. It follows that $\Psi_{t+1}(\hat{\eta})$ must converge. Define $\Psi^*(\hat{\eta}) \equiv \lim_{t\to\infty} \Psi_{t+1}(\hat{\eta})$. Now integrate (A2) by parts to obtain

$$\Psi_{t+1}(\hat{\eta}) = A(\hat{\eta},\infty) - \int_{\underline{\eta}}^{\infty}\Psi_t(\eta)[\partial A(\hat{\eta},\eta)/\partial \eta]d\eta. \qquad (A2')$$

Using the monotone convergence theorem, we have

$$\Psi^*(\hat{\eta}) = A(\hat{\eta},\infty) - \int_{\underline{\eta}}^{\infty}\Psi^*(\eta)[\partial A(\hat{\eta},\eta)/\partial \eta]d\eta.$$

Working backwards from (A2') to (A1), we have our result. Q. E. D.

References

Ades, Alberto, and Thierry Verdier. 1993. "The Rise and Fall of Elites: Economic Development and Social Polarization in Rent-Seeking Societies." Harvard University, mimeo (July).

Amsden, Alice. 1989. *Asia's Next Giant: South Korea and Late Industrialization* (New York: Oxford University Press).

Evans, Peter B. 1992. "The State as Problem and Solution: Predation, Embedded Autonomy, and Adjustment." In Stephan Haggard and Robert R. Kaufman, eds., *The Politics of Economic Adjustment* (Princeton, NJ: Princeton University Press).

Geddes, Barbara. 1986. *Economic Development as a Collective Action Problem: Individual Interests and Innovation in Brazil* (Ann Arbor, MI: University Microfilms.)

Helpman, Elhanan. 1981. "International Trade in the Presence of Product Differentiation, Economies of Scale, and Monopolistic Competition: A Chamberlin-Heckscher-Ohlin Approach." *Journal of International Economics* 11: 305-340.

Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty; Responses to Decline in Firms, Organizations, and States* (Cambridge, MA: Harvard University Press).

Kaufman, Herbert. 1960. *The Forest Ranger: A Study in Administrative Behavior* (Baltimore: Johns Hopkins).

Keefer, Philip, and Stephen Knack. 1993. "Why Don't Poor Countries Catch Up? A Cross-National Test of an Institutional Explanation." Center for Institutional Reform and the Informal Sector Working Paper No. 60 (June).

Klitgaard, Robert. 1988. *Controlling Corruption* (Berkeley, CA: University of California Press).

Kreps, David M. 1990. "Corporate Culture and Economic Theory." In J. Alte and K. Shepsle, eds., *Rational Perspectives on Positive Political Economy* (Cambridge: Cambridge University Press).

Lal, Deepak. 1988. *The Hindu Equilibrium, Vol. I: Cultural Stability and Economic Stagnation, India c1500BC-AD1980* (Oxford: Clarendon Press).

Mauro, Paolo. 1993. "Corruption, Country Risk, and Growth." Harvard University, mimeo (November).

Rauch, James E. Forthcoming. "Bureaucracy, Infrastructure, and Economic Growth: Evidence from U. S. Cities During the Progressive Era." *American Economic Review.*

Shleifer, Andrei, and Robert W. Vishny. 1993. "Corruption." *Quarterly Journal of Economics* 108 (August): 599-617.

Soskice, David, Robert H. Bates, and David Epstein. 1992. "Ambition and Constraint: The Stabilizing Role of Institutions." *Journal of Law, Economics, and Organization* (October): 547-560.

Wade, Robert. 1990. *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization* (Princeton, NJ: Princeton University Press).

Weber, Max. 1968 [1904-1911]. *Economy and Society*. Guenther Roth and Claus Wittich, eds. (New York: Bedminster Press).

Young, Crawford. 1978. "Zaire: The Unending Crisis." *Foreign Affairs* 57 (Fall).