A dynamic theory of competence, loyalty and stability in dictatorships

Nils-Petter Lagerlöf*

Abstract: This paper presents a dynamic model of power competition in a nondemocracy. In each period, a ruler from an incumbent dynasty is challenged. If he survives, he hands over power to a (biological or ideological) offspring. He can control his offspring's chances of surviving future threats, by choosing how competent and loyal administrators to hire, and how many. The society can stay for a long time on a volatile path where subsequent dynasties of rulers regularly replace one another, each purging the preceding dynasty's competent administrators, replacing them with loyalists, thus keeping competence levels bouncing around a low level. This may be followed by an endogenous transition to a path without such purges and a simultaneous rise in competence.

*Department of Economics, York University, 4700 Keele Street, M3J 1P3 Toronto, ON, Canada. E-mail: lagerlof@econ.yorku.ca. This project has benefitted from discussions with Daron Acemoglu, Ahmet Akyol, Yoram Barzel, Fahad Khalil, Oksana Leukhina, Yinan Li, Torsten Persson, and Kevin Tsui. I thank seminar participants at the University of Washington, Seattle, for comments. I also thank Mir Kabir and Quazi Fidia Farah for research assistance. Finally, I am grateful for financial support from the Social Sciences and Humanities Research Council of Canada.

1 Introduction

Autocratic regimes are not directly accountable to any electorate and therefore there is no simple mechanism through which incompetent or corrupt autocrats are weeded out. However, such regimes are not invulnerable. According to Tullock (1987) there are three ways in which an autocrat can be ousted: by someone within the own government or administration; by an external power; or by a public uprising. The last one of these, Tullock argued, is exceedingly rare (perhaps because he was writing in 1987).

Here we analyze a dynamic model of power competition in a nondemocracy. In a Tullockian spirit, we assume that the regime faces either an external or an internal threat. We link this threat structure to recent work by Egorov and Sonin (2011) on the loyalty-competence trade-off faced by dictators: in a situation where the threat to the ruler is internal his survival chances depend on how loyal his administration is; when the threat is external, the ruler is dependent on his administration's competence. A real-world example could be when the Soviet Union was invaded in 1941. In that moment almost everyone rallied around Stalin, including many of his enemies, and the regime's survival hinged on the competence of its military leadership. In times of peace, however, Stalin instead worried about competition from within the party elite, and therefore promoted incompetent but loyal administrators to important posts.

In our model, rulers belong to different dynasties. In each period the incumbent ruler, if he survives a challenge to his own power, hands power over to what we call an "offspring." This could be the ruler's own son, or a younger follower of the same political movement (or even the ruler himself in a liberal interpretation, as discussed later). The ruler cares about his offspring's chances of surviving in the next period, which he can influence by adjusting the composition of the administration. When making those changes, he faces an Egerov-Sonin type of trade-off between loyalty and competence. Moreover, big and rapid changes in the administration are here costlier than gradual changes: the more administrators are replaced, the less competence can be afforded at a given level of loyalty.

The crucial assumption driving the dynamics is that loyalty is dynasty-specific, but competence is not. When a new dynasty seizes power, it inherits the competence of the existing administration, but starts off with fewer loyalists.

How big changes a new ruler chooses to make depends on the inherited administration's composition. He may choose to purge the previous administration completely, filling the positions with incompetent loyalists. This is optimal if the previous administration's competence level is low, making a purge less costly is terms of destroyed "competence capital." If the inherited competence level is higher, the new ruler instead chooses to maintain that competence and even build it further.

This gives rise to some interesting path dependence. A society can stay for a long time on a path where subsequent dynasties purge the administrators hired by the previous dynasties, making competence levels bounce around a low level. At some point, some dynasty manages by chance to stay in power long enough to accumulate sufficiently many competent administrators, so that, when the dynasty is eventually ousted, the new dynasty does not find it optimal to purge the inherited administration.

The result is a transition away from repeated purges to a more stable path and a simultaneous rise in the levels of administrative competence. This may capture something important for understanding development, if we interpret a larger fraction competent administrators as representing broad institutional quality, like the organizational structures of police, courts, civil service, etc. Having better such institutions makes it less worthwhile for a new dynasty to destroy them.

This paper seeks to contribute to a vast literature on the link between political institutions and economic development. There is a great deal of work in economic history and political science on how and why democracies tend to be more conducive to economic development, although many of these do not apply formal models (see e.g. Lipset 1959, Moore 1966, North and Weingast 1989, Olson 1993).

Empirically, however, the causal link between democracy and development is not as clear as one might think. There is little evidence of cross-country correlation between per-capita income levels and democracy, when controlling for country fixed effects (Acemoglu and Robinson 2006, Ch. 3; Acemoglu, Johnson, Robinson and Yared 2008). This might motivate the search for theories of how development depends on other political or institutional factors than democracy, such as administrative competence and political stability.

A large body of research in political economy, going back at least to Downs (1957), has analyzed how democracies function (see e.g. Osborne and Slivinski 1996, Besley and Coate 1997, Persson and Tabellini 2000). A different strand of that literature examines nondemocracies (see e.g. Tullock 1987; Grossman 2000; Grossman and Noh 1994; Acemoglu, Egorov and Sonin 2008, 2009, 2010; Acemoglu, Ticchi and Vindigni 2010).

There are also models of transitions from dictatorship to democracy in the form extensions of the franchise; see e.g. Acemoglu and Robinson (2000, 2001, 2006). Gradstein (2007, 2008) and Cervellati et al. (2008) set up multiple-equilibria models that can endogenously generate different institutional outcomes (democracy or dictatorship), each associated with different economic outcomes.

Acemoglu, Ticchi and Vindigni (2011) explain how an inefficient state organization can persist in an (emerging) democracy, where a rich elite maintain power by, prior to democratization, employing inefficient bureaucrats, who earn rents and become pivotal in the voting process.

Acemoglu, Egorov and Sonin (2010) study a dynamic environment where some fraction of the members of an incumbent government can block changes to its composition. Such incumbency veto power enables incompetent governments to survive, and leads to less flexibility in the presence of exogenous shocks. In a similar spirit, we explain how administrative incompetence can be sustained over time, but our mechanism is linked to purges of competent administrators, rather than any veto powers of incumbent administrators.

A similar loyalty-competence trade-off is examined in great detail by Egorov and Sonin (2011). Whereas our model has a collective administration they have one single administrator (or vizier), who chooses whether to be loyal, or not. In our model, loyalty is not a choice, but rather depends on how far away, e.g. ideologically or ethnically, the ruler recruits administrators in the search for competent candidates. Although less detailed in many dimensions, our framework can be used to analyze the dynamic paths of competence and loyalty, in particular transitions from a state of continual purges to one with no purges. We also let the fraction of the administrators who are replaced in any given period be endogenous, which enables the model to capture the degree of gradualism in such adjustments, i.e., a measure of stability.

Besley and Persson (2008) study the incentives for rulers to invest in fiscal and legal state capacity, i.e. the state's ability to collect taxes and enforce contracts (see also Besley and Persson 2009, 2010). One result in their model, also supported by the data, is that external threats promote investments in state capacity, whereas civil war tends to do the opposite. This resembles the link in our model between external and internal threats and the relative value of competence and loyalty.

Other related papers include Mueller (2009), who studies politicians' choices of appointment procedures for bureaucrats. Mulligan and Tsui (2008) examine a dynamic model where an incumbent ruler can use taxation (or rents) and entry barriers to influence his competitors' incentives and abilities to challenge him. They do not allow for within-regime dynamics, or transitions.

Finally, our model has a connection to theories on endogenous transitions from stagnation to sustained growth in per-capita incomes (see e.g. Galor and Weil 2000, Hansen and Prescott 2002, Lagerlöf 2003). Although we do not even have any explicit variable measuring income, our model can generate transitions in the form of a rise in administrative competence and political stability, and we cite evidence that such factors are correlated with (and, possibly, impact) economic growth.

The rest of this paper is organized as follows. Next Section 2 motivates why the mechanisms that we model could be relevant when thinking broadly about the links between institutions and development. Section 3 sets up the model, first deriving the competence-loyalty trade-off in a static environment (Section 3.1) and then examining a dynamic framework (Section 3.2). Section 4 ends with a concluding discussion.

2 Background

2.1 Examples of purges

There are many examples from nondemocracies of new rulers eliminating administrators or politicians from the previous regime. Consider first two examples from Iraq. In 1979, after forcing his predecessor Ahmad Hassan al-Bakr out of power, Saddam Hussein imprisoned and killed several of the ruling Baath party's top brass, some of whom were publicly escorted out from a special party congress (Marr 2004, pp. 178-181).

Then in 2003 after Saddam Hussein's regime had been ousted by a U.S. led invasion, the new provisional government under Paul J. Bremer's leadership decided to disband the whole army and fire all Baathist high-ranking civil servants (Packer 2005, pp. 189-196).

Another example comes from the Soviet Union. Although leader of the Communist Party since 1924, it was not until after the 17th party congress in 1934 that Stalin became the unopposed dictator after surviving a failed attempt from some delegates to oust him. Following that he put in motion the famous Terror of 1937-38, during which the vast majority of the congress delegates were arrested or killed, as well as 5 of the 15 members of the Politburo and 98 of the 139 members of the Central Committee (Montefiore 2003, p. 237).

Yet another example comes from South Korea, where President Park after seizing power in 1961 embarked on a "purification" campaign, blacklisting a whole generation of politicians who had served under the previous regime (Oh 1968, pp. 138-140).

2.2 Stability, competence and development

In our model, two changes occur when the economy transits to a path with no purges: administrative competence increases, and the economy becomes more stable. (This holds whether stability is measured by the fraction of the administrators being replaced, or the frequency with which dynasties are ousted.)

Even though our model does not say anything explicitly about incomes, we think such transitions capture something about takeoffs from stagnation to sustained economic growth. Arguably, improved administrative competence should be growth enhancing. In planned economies administrators could be making detailed decisions on production quotas and need to organize great amounts of information.[1] Such tasks may require high skill levels to be performed well.

As argued already, we may also interpret competence more broadly, capturing e.g. absence of corruption. There is a large literature documenting a negative correlation between various measures of corruption and economic growth (see e.g. Mauro 1995, Svensson 2005).

Similarly, various measures of political stability show high correlation with economic growth in the cross-country data (see e.g. Mauro 1995, Alesina et al. 1996, Alesina and Perotti 1996). This is particularly true for nondemocracies (Przeworski et al. 2000, pp. 211-213).

Our model predicts that, once a purge is over, competence grows as long as the dynasty is not ousted. Moreover, when a dynasty is ousted, the effect is less likely to be a competence destroying purge if the dynasty has been in power longer, and thus accumulated more competence. This seems consistent with Clague, Keefer, Knack and Olson (1996), who document that good governance in some broad sense increases with an autocratic groups's tenure: more long-lived autocratic regimes are associated with e.g. smaller black market exchange premiums and better credit ratings than short-lived ones. Also consistent with our model, Clague et al. (1996) find that the probability of a regime being ousted in a coup is highest in the earlier years's of its tenure.

3 The model

3.1 The competence-loyalty trade-off

In this section, we set up a static model of a ruler's trade-off between loyalty and competence among administrators. These traits are not mutually exclusive, so administrators can be competent or loyal, neither, or both.

The idea we seek to capture is that an administrator is more likely to be loyal if he is close to the ruler in some sense, e.g. ethnically, ideologically, or

[1] See e.g. Weitzman (1970) for a model of the interaction between a planning ministry and several industries.

through family ties, whereas competence does not depend on such closeness. To model this we let the positions of the ruler, as well as the pool of potential administrators, be represented by points on a circle with circumference B, on which the pool of potential administrators are uniformly distributed. (This interval need not be a circle, but that assumption may make it easier to see why the ruler's own position does not matter for the results below.)

The ruler has a mass of $r \in [0, 1]$ positions to fill. This will later be interpreted as the fraction of the initial unit-sized administration that he chooses to replace, but in this section we treat r as exogenous.

The ruler selects his candidates by first letting a mass $d \in [r, B]$ of the potential candidates do a test that informs him about their competence, and then hires $r \leq d$ of those d candidates.

Let the ruler's position on the circle be normalized to 0 and let $x \in [-B/2, B/2]$ be the position of some candidate, as illustrated in Figure 1. The distance between the ruler and the candidate is thus $|x| \in [0, B/2]$.

Let the stochastic variable $\widetilde{c}(x)$ denote whether the candidate at location x is competent $[\widetilde{c}(x) = 1]$, or not $[\widetilde{c}(x) = 0]$. Similarly, let the stochastic variable $\widetilde{l}(x)$ denote whether the candidate at location x is loyal $[\widetilde{l}(x) = 1]$, or not $[\widetilde{l}(x) = 0]$.

The first crucial assumption we make is that, at any location x, the stochastic variables $\widetilde{c}(x)$ and $\widetilde{l}(x)$ are independent. More precisely, the probability that a candidate is competent does not depend on the distance from the ruler:

$$\Pr\left[\widetilde{c}(x) = 1\right] = \rho \in (0, 1) \tag{1}$$

for all $x \in [-B/2, B/2]$. We assume that $\rho B \geq 1$, so that the total mass of competent candidates in the pool does not fall below the maximum mass of positions needed to be filled; recall that $r \in [0, 1]$.

Different from competence, loyalty is a trait that cannot be tested for, but it can be inferred from the candidate's distance from the ruler. More precisely, the probability that a candidate is loyal is given by:

$$\Pr\left[\widetilde{l}(x) = 1\right] = f(x) = 1 - \left(\frac{2\kappa}{B}\right)|x|, \tag{2}$$

where $\kappa \in [0, 1]$ is an exogenous parameter that determines the fraction loyal candidates at the maximum distance from the ruler, i.e., at distance $x = B/2$ or $x = -B/2$. That is, $f(B/2) = f(-B/2) = 1 - \kappa$. If $\kappa < (=)1$, then a strictly positive (zero) fraction of the most distant candidates are loyal. All candidates at zero distance are loyal, since $f(0) = 1$.

Consider a ruler who tests candidates over an interval of length $d \in [r, B]$

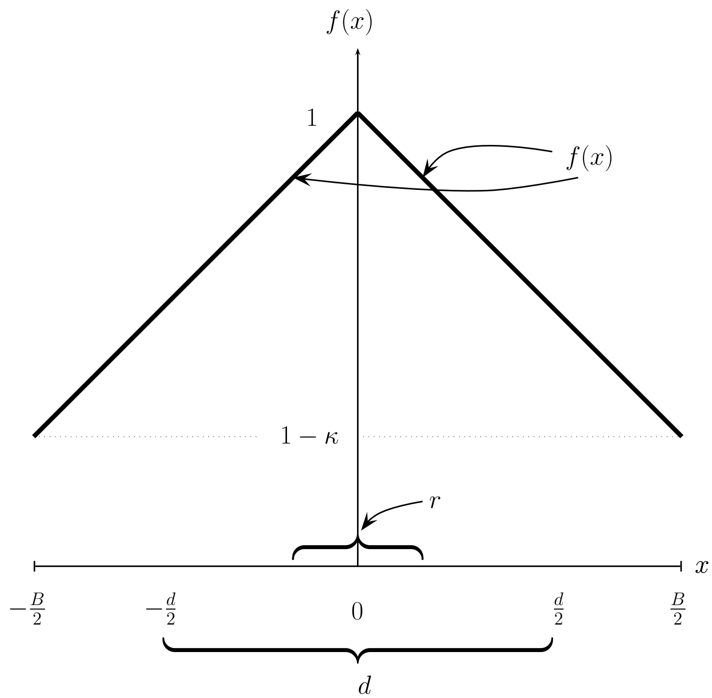

Figure 1: *Illustration of the competence-loyalty trade-off. The ruler sets the recruiting interval, d, taking the mass of positions to be filled, r, as given.*

centered on his own position, 0, and hires only those he finds competent. The recruitment interval thus equals $[-d/2, d/2]$; see Figure 1. The total mass of candidates tested as competent is ρd. The fraction of the r hired administrators who are competent, denoted q, becomes

$$q = \begin{cases} \frac{\rho d}{r} & \text{if } d \in [r, \frac{r}{\rho}], \\ 1 & \text{if } d > \frac{r}{\rho}. \end{cases} \tag{3}$$

The lowest value q can take is ρ, which happens if $d = r$, i.e., if the ruler tests only as many candidates as he has positions to fill. Only by testing more candidates ($d > r$) can he ensure above-average competence among the recruits ($q > \rho$). If $d > r/\rho$, then there are more competent candidates among the d tested than positions to fill, implying $q = 1$. However, as will be seen soon, the ruler will not hire over a larger interval than needed, because that would be costly in terms of loyalty.

Let $d_{\min}$ be the *minimum* distance over which the ruler must test and recruit r candidates, ensuring that a fraction $q \in [\rho, 1]$ of them are competent. From (3) we get:

$$d_{\min} = \frac{rq}{\rho}. \tag{4}$$

Next let p denote the average fraction candidates over any interval $[-d/2, d/2]$ who are loyal. Since the distance to the ruler decreases symmetrically on both sides of the ruler's position, p is given by the average fraction loyal on $[0, d/2]$. Using (2) gives

$$p = \frac{1}{(d/2)} \int_0^{d/2} f(x)dx = \frac{1}{(d/2)} \int_0^{d/2} \left[1 - \left(\frac{2\kappa}{B}\right) x\right] dx. \tag{5}$$

Solving the integral in (5) gives p as a decreasing function of d:

$$p = 1 - \left(\frac{\kappa}{2B}\right) d. \tag{6}$$

Since competence is independent of the candidates' positions, the mass of ρd competent candidates is also distributed uniformly on $[-d/2, d/2]$. Therefore the average fraction loyal among the ρd competent is also given by p. We can now derive the *maximum* fraction loyal candidates the ruler can recruit at any given fraction competent, q, by applying (6) and (4), and setting $d = d_{\min}$. This gives:

$$p = 1 - \phi q r, \tag{7}$$

where

$$\phi = \frac{\kappa}{2B\rho}. \tag{8}$$

This describes the trade-off between competence (q) and loyalty (p), at any given r. Recall that the minimum level of q in (7) is ρ, so the maximum level of p the ruler can choose is $1 - \phi\rho r$.[2]

For simplicity, in the rest of this paper we let ρ be very close to zero and B very large, holding the product ρB constant (and, recall, greater than one; see below). In other words, we let competent candidates be extremely rare, but the pool to choose from very large. This means that (as a good approximation) we can let the ruler choose q on $[0, 1]$, and thus be able to ensure full loyalty ($p = 1$) by setting $q = 0$. In terms of Figure 1, loyalty falls very slowly when expanding the recruiting interval [because B is large; see (6)], but the

[2]More precisely, if $d = r$, then $p = 1 - \phi\rho r$, which follows from either setting $d = r$ in (6) and using the notation in (8); or from setting $d = r$ in (3), to get $q = \rho$, and then using (7).

ruler must recruit over a very large distance to find the very few competent candidates (because ρ is small).

By assuming that $B\rho \geq 1$, it follows that, if $r = 1$, then the ruler can fill the positions with only competent candidates ($q = 1$) if he so chooses; see (3) and note that $d \leq B$.[3] Since $\kappa \in [0, 1]$, it then also follows that $0 \leq \kappa \leq B\rho$, which in turn implies from (8) that

$$0 \leq \phi \leq \frac{1}{2}. \tag{9}$$

The trade-off in (7) has the interesting property that the higher is the fraction competent among new hires, q, the costlier is an increase in r in terms of reduced loyalty, p. In that sense, improving competence is best done gradually, by replacing small segments of the administration, whereas improving loyalty can be done in one sweep.

Having modeled this trade-off, we next examine the ruler's choices in a dynamic setting.

3.2 A dynamic setting

Consider next a discrete-time non-overlapping generations setting, where different dynasties of rulers compete for power. In each period, some ruler starts off in power, is challenged, reshuffles the administration, and then passes power on to a (biological or ideological) offspring in the next period. The reshuffling is done with the aim to maximize the probability that the offspring survives in the next period, and is made subject the competence-loyalty trade-off in (7).

More precisely, the timing of events in each period is as follows:

1. A ruler enters office, taking as given the composition of a unit-sized administration. A fraction $c \in [0, 1]$ of the administrators are competent, and a fraction $l \in [0, 1]$ loyal. (As discussed already, and further below, these categories are not mutually exclusive.)

2. The state of the world is realized, determining if the threat to the ruler is external or internal. If the threat is external the ruler is ousted with a probability that depends on the fraction competent, c; if the threat is internal the ruler is ousted with a probability that depends on the fraction loyal, l.

[3] Relaxing this assumption, there would be some upper limit on the fraction competent recruits the ruler could hire, q, and that upper limit would depend on how many are replaced, r.

3. If the incumbent ruler is not ousted he can fire some, or all, administrators and hire new ones, endogenous fractions of whom are loyal and competent. He then hands power over to his offspring, who enters stage 1 in the next period. If the incumbent ruler is ousted, the new ruler first inherits the previous ruler's administration, with the same fraction competent administrators (same c), but with fewer loyalists (lower l). The new ruler then decides how many of the administrators to fire, hires new ones, and hands over to his offspring who enters stage 1 in the next period.

Most of the analysis in this section will refer to the ruler's choices at stage 3 about how to update c and l, taking the initial levels of these state variables as given. This could refer to either an incumbent ruler who has survived a threat of being ousted, or a new ruler who has just ousted his predecessor. In Section 3.2.5 we analyze the effects of a change in leadership, associated with a drop in l prior to that update.

Assuming that the ruler seeks to maximizing his offspring's chances of survival in the next period seems like a useful starting point, although one can consider different setups. In particular, the ruler here does not care about the survival of the dynasty in the period after next, i.e., his (biological or ideological) "grandchildren."

Note also that the individual administrators do not live in generations like the rulers; those who are not fired stay on as administrators in the next period. One interpretation is that the administrators are infinitely lived (when not replaced). Another is that the administrative positions are inherited by default, and that the traits (competence and loyalty) are also passed on among administrators from father to son.[4] Yet another interpretation abstracts from generations altogether, letting the ruler himself stay in power (if not ousted), rather than handing power to an offspring, but that assumes that the ruler is myopic and cares about his own survival chances only in the very next period (perhaps not unreasonable in this context).

3.2.1 The ruler's survival function

Let the incumbent ruler's survival probability at stage 2 be S. In the next period, his offspring's probability of surviving is S'. Let z denote the probability that the threat is external; with residual probability the threat is internal. In each respective state, we let the survival chances be linear in the fraction competent and loyal, respectively, so that the survival probability of the ruler

[4] This may be a good description of many preindustrial societies.

becomes:

$$S = zc + (1 - z)l. \tag{10}$$

Some remarks are in order. First note that (10) assumes that only c matters when the threat is external, and only l when the threat is internal. We could allow c and l to matter in both states of the world. As long as the survival probability in each state is linear in c and l, we can still write the overall survival probability as in (10), but with z being a function of the underlying probability of an external threat.[5]

As noted already, competence and loyalty are not mutually exclusive: an administrator can be, e.g., both loyal and competent, or both disloyal and incompetent. Moreover, the formulation in (10) implies that if all administrators are loyal and competent ($c = l = 1$), then the ruler is fully safe ($S = 1$). This assumption is made only to economize on notation; Section 3.2.7 considers an alternative formulation, where the maximum survival probability falls strictly below one.

The model is silent about the precise way in which competence matters in the face of an external threat. Arguably, the competence level of the military should play a role for success in war. Diplomatic skills and the ability to negotiate could also matter, e.g. by enabling a ruler to avoid a war he would likely lose.

We could also allow for a non-Tullockian interpretation, by letting the state where competence matters correspond to one where the threat comes from a public uprising (rather than from an external power). The implicit assumption could then be that a more competent (less corrupt) administration would make the public less inclined to rebel.[6] However, for the rest of the presentation we shall stick to the interpretation of an external threat.

3.2.2 The ruler's optimal choices

Consider now a ruler at stage 3 (either an incumbent who was not ousted or a new ruler who has ousted a predecessor) seeking to maximize his offspring's survival chances in the next period, S'. Taking as given c and l, the ruler replaces a fraction $r \in [0, 1]$ of the (unit-mass) administration before handing

---

[5]For example, let $\widetilde{z}$ be the probability that the threat is external and let

$$S = \widetilde{z}[\sigma c + (1 - \sigma)l] + (1 - \widetilde{z})[\eta c + (1 - \eta)l],$$

where $\widetilde{z}$, σ and η all lie on $[0, 1]$. This boils down to the formulation in (10) with $z = \widetilde{z}\sigma + (1 - \widetilde{z})\eta$. If $\sigma > \eta$, so that competence matters more when the threat is external, then z is increasing in $\widetilde{z}$.

[6]I am grateful to Fahad Khalil for this suggestion.

power over to his offspring.

We restrict the ruler's powers by assuming that r is a random sample of the administration. The ruler thus cannot make replacements contingent on the administrators' characteristics, and is unable to replace primarily the disloyal and incompetent. This simplifies the analysis a great deal. Ideally, the model should allow the ruler to choose who to replace, but depending on context and interpretation, imposing that he cannot need not be completely unrealistic. It could capture limitations on the information and management abilities of the ruler, especially when replacing large segments of the administration.[7] See also Section 4 for some further discussion of this assumption.

Let c' and l' be the fractions competent and loyal when the offspring enters stage 1 in the next period, and recall from Section 3.1 that we let q and p be the fractions of the r new hires who are competent and loyal, respectively. Since r is a random sample of the initial administration, it follows that:

$$c' = c(1-r) + rq, \tag{11}$$

and

$$l' = l(1-r) + rp = l(1-r) + r - \phi q r^2, \tag{12}$$

where we have used the loyalty-competence trade-off among new hires in (7).

Since the offspring's survival probability in the next period is $S' = zc' + (1-z)l'$, using (11), (12) and some algebra, we get

$$S' = q[z - (1-z)\phi r]r + zc(1-r) + (1-z)[l(1-r) + r] \equiv S(r, q; c, l). \tag{13}$$

The problem thus boils down to maximizing (13) with respect to r and q. Since (13) is linear in q, the ruler sets $q = 0$ $(q = 1)$ if $S_q(r, q; c, l) < (\geq) 0$, which in turn depends on the sign of the coefficient in front of q in (13), i.e., $z - (1-z)\phi r$. It follows that

$$q = \begin{cases} 0 & \text{if } r > \widehat{r}, \\ 1 & \text{if } r \leq \widehat{r}, \end{cases} \tag{14}$$

where

$$\widehat{r} = \frac{z}{\phi(1-z)}. \tag{15}$$

In other words, if the ruler chooses to make big changes in the administration

[7]For example, under Stalin's terror precise quotas were given by the Politburo that specified how *many* were to be arrested and executed in each region of the Soviet Union. Exactly *who* fell victim was then decided by local authorities, often based on long-forgotten personal conflicts or sheer coincidence (Montefiore 2003, Ch. 20).

($r > \widehat{r}$), then he also chooses to hire only incompetent candidates ($q = 0$).

We assume that $z < \phi/(1+\phi)$, implying that $\widehat{r} < 1$. This allows for the possibility that the ruler finds it optimal to set $r > \widehat{r}$ and thus hire only incompetent candidates.

The optimal choice of r can now be derived by maximizing (13) over r after substituting q for (14). Define $\widehat{S}(r; c, l)$ as follows:

$$\widehat{S}(r; c, l) = \begin{cases} zc + (1-z)l + r[(1-z)(1-l) - zc] & \text{if } r > \widehat{r}, \\ zc + (1-z)l + r[z(1-c) + (1-z)(1-l)] - (1-z)\phi r^2 & \text{if } r \leq \widehat{r}. \end{cases} \tag{16}$$

That is, $\widehat{S}(r; c, l) = S(r, 0; c, l)$ for $r > \widehat{r}$; and $\widehat{S}(r; c, l) = S(r, 1; c, l)$ for $r \leq \widehat{r}$. Figure 2 shows an example of how $\widehat{S}(r; c, l)$ may be shaped. Generally, it holds that $\widehat{S}(0; c, l) = zc + (1-z)l$; with no changes in the administration the initial levels of c and l determine the offspring's survival probabilities in the next period. Also, $\widehat{S}(1; c, l) = 1 - z$; if replacing the whole administration with incompetent loyalists ($r = 1$ and $q = 0$) the ruler's offspring survives if, and only if, the threat in the next period is internal, which happens with probability $1 - z$.

Where $\widehat{S}(r; c, l)$ peaks depends on c and l. The details are shown in Section A of the appendix, but the idea is quite intuitive. First define these two threshold levels of c (both of which depend on l):

$$\overline{c} = \left(\frac{1-z}{z}\right)(1-l), \tag{17}$$

and

$$\underline{c} = \left(\frac{1-z}{z}\right)(1-l) - 1 = \overline{c} - 1. \tag{18}$$

There are three cases to consider: (A) $c > \overline{c}$, (B) $c < \underline{c}$, and (C) $c \in [\underline{c}, \overline{c}]$. Figure 2 applies to Case (C) but the other cases can be understood intuitively from the same figure.

In **Case (A)**, $\widehat{S}(r; c, l)$ has negative slope for $r > \widehat{r}$. The ruler's optimal r is here given by an interior maximum on $[0, \widehat{r}]$, denoted $r_{\max}$ in Figure 2, and given by

$$r_{\max} = \frac{z(1-c) + (1-z)(1-l)}{2\phi(1-z)}. \tag{19}$$

In this case, it holds that $r_{\max} < \widehat{r}$, so the ruler's optimal q equals 1; see (14).

In **Case (B)**, $\widehat{S}(r; c, l)$ has positive slope both for $r > \widehat{r}$ and $r \leq \widehat{r}$. We thus have a corner solution, where the ruler sets $r = 1$ and $q = 0$, which is what we call a purge.

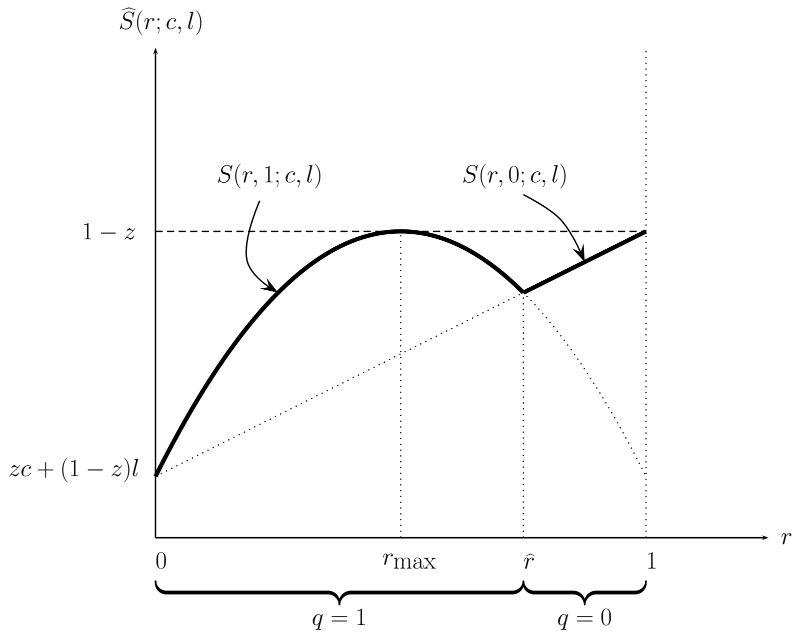

Figure 2: *Illustration of the ruler's optimal choices of r and q. In this case the ruler is indifferent between setting either $r = 1$ and $q = 0$ or $r = r_{max}$ and $q = 1$.*

In **Case (C)**, which could be the one illustrated in Figure 2, there are two local maximum points: $r = 1$ and $r = r_{\max}$. To determine which one of these is the global maximum we compare the value of the objective function, $\widehat{S}(r; c, l)$, at these two points. The analysis becomes quite cumbersome and is suppressed to the appendix, but the solution is relatively easy to state. Section A.3 of the appendix shows that the ruler chooses $r = r_{\max}$ ($r = 1$) if $c > (<)\Psi(l)$, where

$$\Psi(l) = \left(\frac{1-z}{z}\right)(1-l) - \Omega, \tag{20}$$

and

$$\Omega = \frac{2}{\widehat{r}}\left[1 - \sqrt{1-\widehat{r}}\right] - 1 \in [0, 1], \tag{21}$$

where $\widehat{r} = z/[\phi(1-z)]$ is given by (15).

That $\Omega \in [0,1]$ is shown in Section B of the appendix. This can also be shown to imply that $\underline{c} \leq \Psi(l) \leq \overline{c}$; see (17), (18) and (20). Thus, Cases (A) and (B) become special cases of $c > \Psi(l)$ and $c < \Psi(l)$, respectively. (See also Section A.4 of the appendix.)

Figure 2 illustrates the knife-edge case where $c = \Psi(l)$, so that the ruler is indifferent between $r = r_{\max}$ and $r = 1$; in this case we assume that he chooses $r = r_{\max}$.

To sum up, the ruler's optimal choice of r is given by

$$r = \begin{cases} r_{\max} = \frac{z(1-c)+(1-z)(1-l)}{2\phi(1-z)} & \text{if } c \geq \Psi(l), \\ 1 & \text{if } c < \Psi(l), \end{cases} \tag{22}$$

and the fraction competent among the new hires, q, is given by

$$q = \begin{cases} 1 & \text{if } c \geq \Psi(l), \\ 0 & \text{if } c < \Psi(l). \end{cases} \tag{23}$$

Using (7), (22) and (23) we can also derive an expression for the fraction loyal among new hires:

$$p = 1 - \phi q r = \begin{cases} 1 - \frac{z(1-c)+(1-z)(1-l)}{2(1-z)} = \frac{(1-z)(1+l)-z(1-c)}{2(1-z)} & \text{if } c \geq \Psi(l), \\ 1 & \text{if } c < \Psi(l). \end{cases} \tag{24}$$

The expressions for r and q in (22) and (23) can be understood intuitively by thinking of competence as a form of accumulated capital. When the fraction competent is low, $c < \Psi(l)$, the ruler loses little if he replaces all administrators with incompetent loyalists. The incentives to keep some administrators are stronger when there is something to lose. Similarly, if the administration is already relatively loyal – implying that l is large and thus $\Psi(l)$ small – then the incentives to improve loyalty further, thus losing the existing competence, are weak.

We have now derived the incumbent ruler's behavior, as a function of the initial levels of c and l in his administration. Next we shall examine the dynamics of these state variables.

3.2.3 The dynamical system

To a ruler taking c and l as given, the fraction competent in the next period, c', can be derived from (11), (22) and (23). If $c < \Psi(l)$, the ruler purges the administration and the fraction competent in the next period drops to zero, $c' = 0$. If $c \geq \Psi(l)$, we see from (11) that the fraction competent in the next period is given by $c' = c(1 - r_{\max}) + r_{\max} = c + (1 - c)r_{\max}$ where (recall) $r_{\max}$ is a function of c and l given by (19). This gives

$$c' = F(c,l) \equiv \begin{cases} c + (1-c)\left[\frac{z(1-c)+(1-z)(1-l)}{2\phi(1-z)}\right] & \text{if } c \geq \Psi(l), \\ 0 & \text{if } c < \Psi(l). \end{cases} \tag{25}$$

Similarly, the fraction loyalists in the next period can be derived from (12), (22) and (23). When $c < \Psi(l)$, the ruler replaces the whole administration with incompetent loyalists so $l' = 1$. When $c \geq \Psi(l)$, and thus $r = r_{\max}$ and $q = 1$, then $l' = l(1 - r_{\max}) + r_{\max} - \phi(r_{\max})^2$, where $r_{\max}$ is given by (19); the algebra for this case is sorted out in Section C of the appendix. The result is:

$$l' = G(c,l) \equiv \begin{cases} l + \frac{1}{4\phi}\left[(1-l)^2 - \left(\frac{z}{1-z}\right)^2(1-c)^2\right] & \text{if } c \geq \Psi(l), \\ 1 & \text{if } c < \Psi(l). \end{cases} \tag{26}$$

Next recall that the ruler may be ousted at stage 2, in which case the new ruler inherits the fraction competent in the previous administration, c, but faces a lower fraction loyalists. Now let l_0 denote the fraction administrators loyal to the new ruler's dynasty, whereas l is the fraction who were loyal to the previous dynasty. We here treat l_0 as exogenous, which simplifies the analysis a great deal. An alternative approach would be to try to endogenize l_0 by examining how the current administrators are positioned on the circle of circumference B relative to the new ruler.[8]

Recall that the incumbent ruler survives with probability $S = zc + (1-z)l$. Thus, the stochastic difference equation for c can be written

$$c' = \begin{cases} F(c,l) & \text{with probability } S = zc + (1-z)l, \\ F(c,l_0) & \text{with probability } 1 - S = 1 - zc - (1-z)l, \end{cases} \tag{27}$$

where $F(c,l)$ is given by (25). Similarly, the stochastic difference equation for l can be written

$$l' = \begin{cases} G(c,l) & \text{with probability } S = zc + (1-z)l, \\ G(c,l_0) & \text{with probability } 1 - S = 1 - zc - (1-z)l, \end{cases} \tag{28}$$

where $G(c,l)$ is given by (26).

Together with initial conditions for c and l, (27) and (28) characterize the stochastic time paths of c and l. Since this system is stochastic the exact paths

[8]This approach would require a theory of what the new ruler's position on the circle is, as well as that of the current administrators (who, recall, have been hired at potentially different points in time).

depend on what shocks are realized in each period. This is easiest to illustrate numerically, as shown in Section 3.2.6 below.

3.2.4 Phase diagram

Before studying the full stochastic dynamical system, consider a dynasty that is never ousted (although every generation still behaves so as to minimize that risk), so that the dynamics are deterministic and described by the (25) and (26).[9]

The dynamics for c then become trivial to analyze: if $c \geq \Psi(l)$, then c is growing over time (or constant if $c = 1$); if $c < \Psi(l)$, then c drops to (or stays at) zero in the next period. From (25), the locus along which c is constant can be written

$$\mathcal{L}_c = \left\{(c, l) \in [0, 1]^2 : (c < \Psi(l), c = 0) \text{ or } (c \geq \Psi(l), c = 1)\right\}. \tag{29}$$

The dynamics for l are almost as easy to see: if $c < \Psi(l)$, then l jumps to one in the next period; if $c \geq \Psi(l)$, then l can be increasing, decreasing or constant. From (26) we can define the locus along which l is constant as

$$\mathcal{L}_l = \left\{(c, l) \in [0, 1]^2 : c \geq \Psi(l), c = \Gamma(l)\right\} \tag{30}$$

where

$$\Gamma(l) = 1 - \left(\frac{1 - z}{z}\right)(1 - l). \tag{31}$$

Deriving $\Gamma(l)$ amounts to setting $l' = l$ in (26), considering the case when $c \geq \Psi(l)$; note from (20) that $\Psi(1) = -\Omega < 0$, so $l' = l$ cannot hold for $c < \Psi(l)$.[10]

The joint dynamics for c and l are illustrated in the phase diagram in Figure 3. The thick solid lines are the loci along which $c' = c$ and $l' = l$, defined by (29) and (30). The steady state (point SS) is given by their intersection, $\mathcal{L}_c \cap \mathcal{L}_l = (1, 1)$. That is, all administrators are both competent and loyal in steady state. It is straightforward to see that this steady state is unique.[11]

[9] Equivalently, this amounts to imposing $S = 1$ in (27) and (28).

[10] One can also derive $c = \Gamma(l)$ by setting the fraction loyal among new hires, p, equal to that among the existing administration, l. Using (24), and considering the case when $c \geq \Psi(l)$, it is seen that $p = l$ gives $c = \Gamma(l)$.

[11] More precisely, a steady state of the system in (25) and (26) is defined as a combination of c and l, such that $c' = c$ and $l' = l$. From (25) we see that $c' = c$ can hold only if either $c = 1 \geq \Psi(l)$, or $c = 0 < \Psi(l)$. From (26) we see that $l' = l$ cannot hold if $c = 0 < \Psi(l)$, since then $l' = l = 1$ must hold but $\Psi(1) = -\Omega < 0$, contradicting $c = 0 < \Psi(l)$. If $c = 1 \geq \Psi(l)$, then $l' = l$ implies $l = 1$.

Note also that the dynasty survives with certainty in steady state, since $S = 1$ when $c = l = 1$; recall (10).

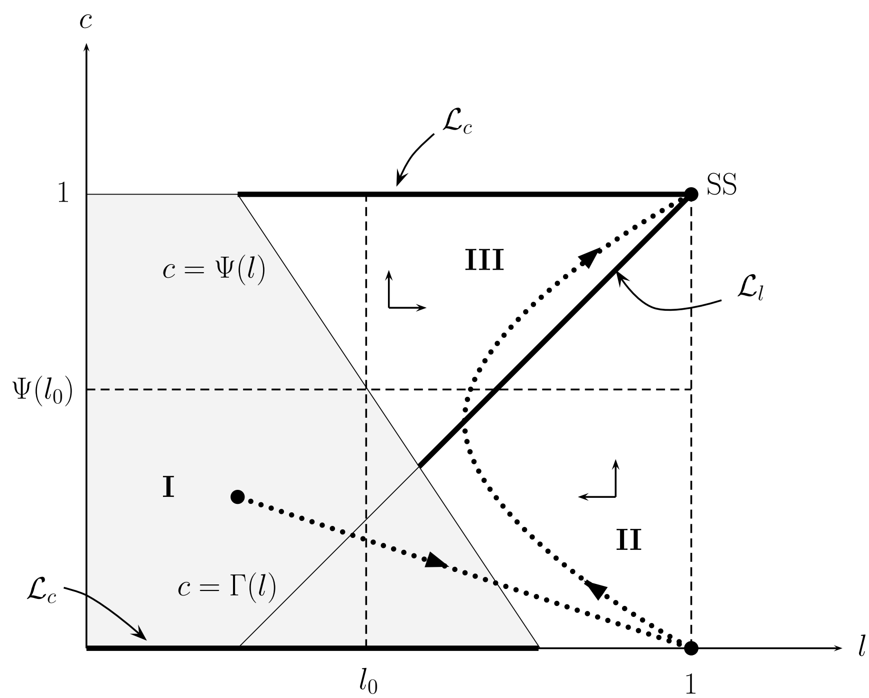

Figure 3: *Phase diagram. The thick solid lines show where c and l are constant. The dotted line shows a dynasty's path if not ousted. Starting in Region I (shaded) the initial ruler undertakes a purge, replacing the whole administration with incompetent loyalists, implying a discrete jump to $c = 0$ and $l = 1$ in the next period. Thereafter c increases monotonically over time, while l declines temporarily (in Region II), and then rises again, as the path enters Region III and converges to the unique steady state (point SS). A dynasty being ousted amounts to a horizontal jump to l_0.*

How c and l evolve over time off steady state depends on initial values. In Figure 3 we distinguish three regions: I, II, and III. Consider first **Region I** (the shaded area), where $c < \Psi(l)$. This is the "purge region." A ruler starting off here sets $c' = 0$ and $l' = 1$, replacing the whole administration with incompetent loyalists. In the next period the offspring survives with certainty if the threat is internal, and is ousted with certainty if the threat is external (had we not imposed certain survival).

Regions II and III constitute a "no-purge region," where $c \geq \Psi(l)$. Consider first **Region III**, where $c \geq \Psi(l)$ and $c \geq \Gamma(l)$. From (25) we see that $c' \geq c$, and from (26) that $l' \geq l$. That is, both competence and loyalty are increasing over time. Since $\Psi'(l) < 0$ it also follows that $\Psi(l') \leq \Psi(l)$. Given $c \geq \Psi(l)$, it must thus hold that $c' \geq \Psi(l')$, implying that no purge takes place in the next period. In terms of Figure 3, the dynasty's path leads away from Region I.

Consider finally **Region II**, where $c \geq \Psi(l)$ and $c < \Gamma(l)$. Here it follows from (25) that $c' \geq c$, and from (26) that $l' < l$, so that competence increases but loyalty declines from one period to the next. Intuitively, starting off with a very loyal and incompetent administration (c small, l large) the ruler can only raise competence by firing loyalists. In Figure 3 it looks as if this "northwesterly" path could lead into Region I, but Section D of the appendix shows that $c' \geq \Psi(l')$ always holds in the next period if starting in Region II. However, the path eventually enters Region III, at which point l starts to grow again.

3.2.5 Effects of a change in power

In terms of Figure 3 a dynasty's overthrow is interpreted as a horizontal leap away from the dotted path, i.e. a drop to l_0 on the l-axis. We assume that l_0 is such that $0 < \Psi(l_0) < 1$.[12] This means that for low enough inherited competence, $c < \Psi(l_0)$, a change in power leads to a purge, as the new ruler starts in Region I; for higher inherited competence, $c \geq \Psi(l_0)$, there is no purge.

Whether a purge follows a change in power, or not, thus depends on the inherited competence level, c; on the new loyalty level, l_0; and on the parameters z and ϕ, which determine the boundary of Region I, i.e. $\Psi(l)$ in (20). We next sum up how these variables determine a new ruler's choice whether, or not, to purge the administration.

A higher l_0 makes a purge less likely. This follows because a higher l_0 implies a lower purge threshold, $\Psi(l_0)$; note that $\Psi'(l) < 0$. Intuitively, the need for a purge is lower when the new ruler enjoys greater initial loyalty from the inherited administration.

[12]Using (20) this holds if

$$1 - \frac{z(1+\Omega)}{1-z} < l_0 < 1 - \frac{z\Omega}{1-z}.$$

It is easy to see that the lower boundary is less than the upper boundary (and also less than one). The upper boundary is positive, and greater than 1/2, because $\Omega < 1$ and $z/(1-z) \leq 1/2$; the latter follows from the assumption that $\widehat{r} = z/[\phi(1-z)] < 1$, and $\phi \leq 1/2$ [recall (9)]. Thus, there exists some interval on which l_0 can fall, such that $0 < \Psi(l_0) < 1$ holds.

A higher c makes a purge less likely. As described earlier, c constitutes competence capital inherited from the previous dynasty, which is costlier to destroy the more of it has been accumulated. Interestingly, since c grows over time while a dynasty stays in power, purges are thus less likely if the ousted dynasty had been in power longer. Note also that the probability of the incumbent dynasty's survival (S) is increasing along the path, and equals one in steady state.

This fits with Clague et al. (1996), who document both that good governance (which we can interpret as competence in our model) increases with the ruling group's tenure, and that the probability of a regime being ousted in a coup is highest in the earlier years of its tenure.

A higher z makes a purge less likely. This follows because an increase in z shifts down $\Psi(l)$, thus shrinking Region I in Figure 3. (This is shown in Section E of the appendix.) The intuition is straightforward. A higher z implies a higher value of competence, making a new ruler more reluctant to destroy existing competence capital.

A higher ϕ makes a purge more likely. In Figure 3, an increase in ϕ shifts up $\Psi(l)$, thus expanding Region I. (This is also shown in Section E of the appendix.) Intuitively, a higher ϕ implies a less favorable competence-loyalty trade-off, making a purge more attractive to the ruler.

To interpret this result, recall from (8) that $\phi = \kappa/(2B\rho)$, where ρ is the fraction competent candidates on each point on the circle of circumference B and $1 - \kappa$ measures the fraction loyal candidates at the maximum distance from the ruler, $B/2$. A higher ϕ can thus be interpreted as a lower ρ, due e.g. to a less educated pool of candidates. This could suggest a causal link from education to political stability.

3.2.6 A quantitative illustration

Parameter values The model is extremely stylized, so we cannot pursue any serious calibration exercise. However, some of the qualitative features of the stochastic paths are easier to understand with the help of some simple numerical simulations. To that end, we need values for three parameters, ϕ, z and l_0, and initial conditions for two state variables, c and l. The values we choose are somewhat arbitrary but we provide some rationales below. These values are summed up in Table 1.

We set ϕ at its highest feasible level, given the permissible range in (9), i.e., $\phi = 1/2$. This imposes the least favorable loyalty-competence trade-off to the ruler that is consistent with the assumptions imposed on κ, B and ρ in Section 3.1.

We set the probability of an external threat, z, to 0.2. We could in-

Parameter	Value	Comment
z	0.2	Probability of external threat 20%
ϕ	0.5	At upper bound; loyalty-competence trade-off tightest possible
l_0	0.7248	Fraction loyal to new ruler; 2.5% above no-purge threshold when $c = 1$

Initial condition	Value	Comment
Initial l	0.7248	Same as l_0
Initial c	0.25	Arbitrary but below $\Psi(l_0)$, ensuring initial ruler purges administration

Table 1: *Parameter values and initial conditions for the baseline case.*

terpret this as the frequency of inter-state war, which seems to have varied greatly across countries/empires, epochs and regimes. For example, in European preindustrial history the war frequency among the Great Powers seems to have been around 35-90%, then declining around the beginning the 19th century, followed by a short rise around the two 20th-century world wars, and sustained peace after that (Lagerlöf 2010). Setting $z = 0.2$ might be a reasonable compromise.[13]

We set l_0 to approximately 0.7248, which is 2.5% above the threshold level that a new ruler's loyalty level must exceed for him to choose not to purge the administration if all administrators are competent, $c = 1$.[14] If l_0 falls below that threshold, new rulers would always purge the inherited administration and the model would not be able to generate any transition to a no-purging path.

When choosing initial conditions, we set the initial l to l_0. Thus, the very first ruler has the same loyalty base as one who has just ousted the incumbent.

The initial level of c does not matter, as long as it falls below $\Psi(l_0)$. (We set it arbitrarily to 0.25.) Thus, the initial ruler starts in Region I in Figure 3, with a purge, allowing a subsequent transition to a no-purging path to take place.

Simulations Figure 4 shows the time paths for some dynasty that is never ousted, corresponding to the case illustrated in the phase diagram in Figure 3.

Figure 5 shows a simulation when, in each period, power changes hands

[13] Recall also that we assumed $\widehat{r} = z/[\phi(1-z)] < 1$ in order to ensure that some ruler will choose to purge the administration; see (15). Having set $\phi = 1/2$, we must thus set $z < 1/3$.

[14] More precisely, let $\widetilde{l}$ be that threshold, defined from $1 = \Psi(\widetilde{l})$. Then $l_0 = 1.025 \cdot \widetilde{l}$

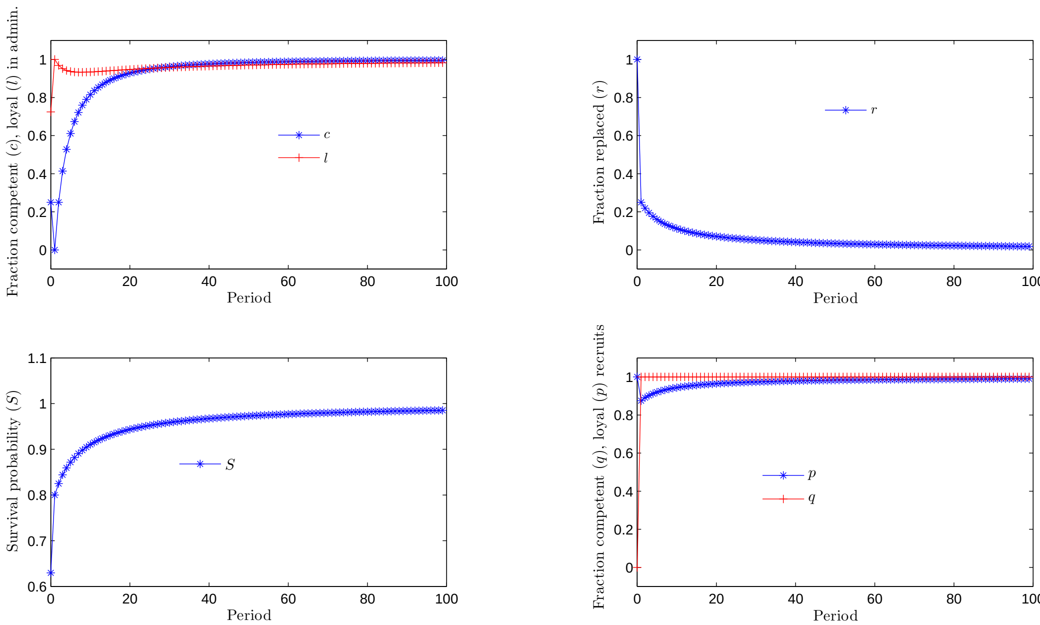

Figure 4: *Simulated time paths when artificially letting one dynasty survive forever ($S = 1$).*

with the endogenously determined probability, $1 - S$, at which point l drops to $l_0 = 0.7248$. The economy first goes through a phase with several purges where competence is destroyed as the incumbent dynasty is ousted, eventually followed by a transition to a no-purging path when c exceeds $\Psi(l_0)$. The length of each dynasty's reign is random, as each faces the state-dependent probability S of surviving. Over each dynasty's reign competence is increasing over time, confirming the results in Section 3.2.4, and consistent with e.g. Clague at al. (1996). We also see that the transition comes in the wake of an exceptionally long-lived dynasty.

Note also that the time paths, although quite volatile overall, become less volatile after the transition. Also, the level around which r fluctuates falls, and S tends to peak at higher levels. That is, the economy transits to a more stable path.

Figure 6 shows some results from a Monte Carlo simulation, based on 1,000 runs. These show a continuous rise in the mean level of competence across those 1,000 runs, and a decline in mean loyalty. The decline in loyalty is driven by a larger fraction of the runs having made transitions to the no-purging phase; recall that in those purges l jumps to one. The upper-right panel displays the fraction of the runs in which the shares of competent and loyal administrators exceed 99%, showing a continual rise in competence and

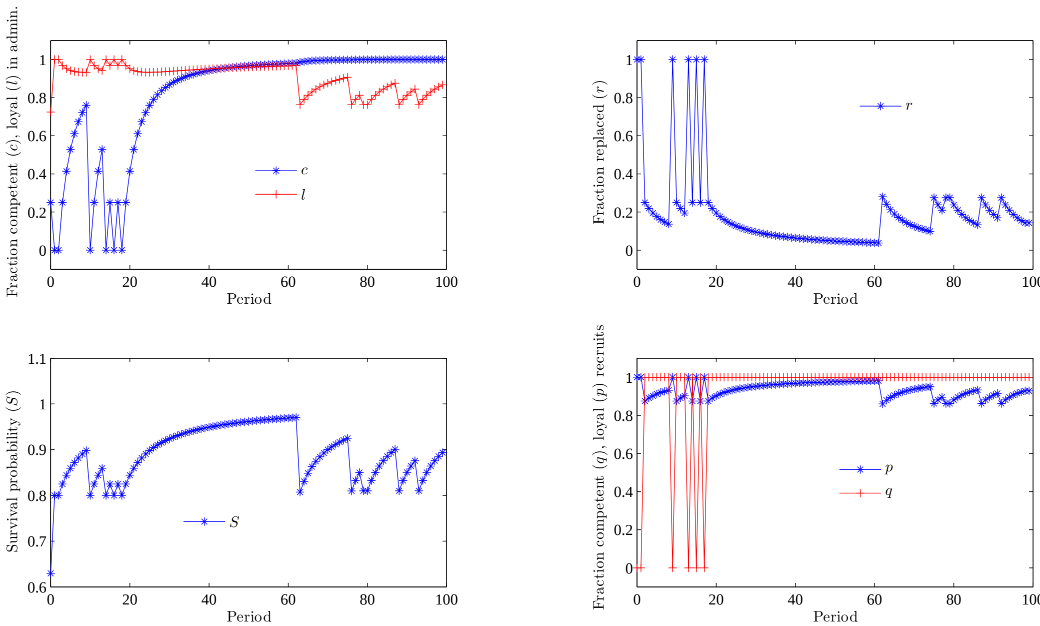

Figure 5: *Simulated time paths when dynasties survive with the endogenous probability* S*.*

decline in loyalty, mirroring the trends in mean levels. The decline in mean r and rise in mean S are also worth noting, reflecting rising levels of stability.

3.2.7 Sensitivity analysis

The Monte Carlo results in Figure 6 refer to the baseline set of parameter values given in Table 1. Figure 7 shows how the paths are altered when changing some of those.

Higher probability of external threat Consider first a higher risk of external threat, $z = 0.25$ (instead of $z = 0.2$), keeping other parameters and initial conditions at their baseline values. Most notable, the rise in mean competence comes earlier, which is not surprising, since z effectively measures the value of competence.

For later periods, a higher z generates slightly more stability (lower r and higher S) compared to the baseline case. For earlier periods, however, the relationship is the reverse. Intuitively, newly arriving dynasties are more likely to be ousted after their competence-eliminating purges, since they are helpless when the threat is external, an event which here occurs with higher probability.

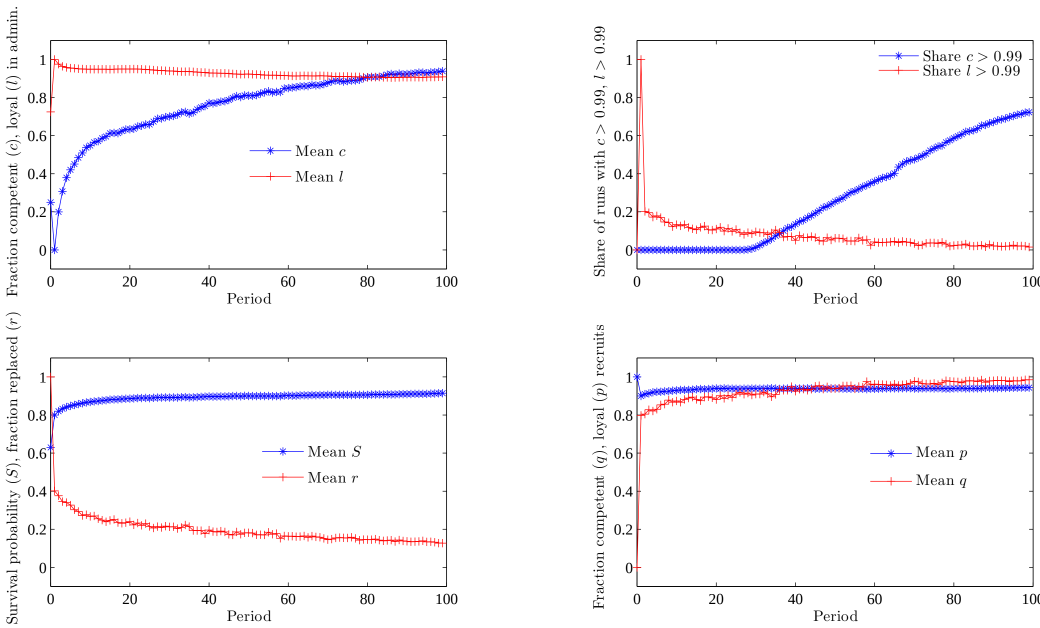

Figure 6: *Monte Carlo simulation. The lines are based on 1,000 runs like that shown in Figure 5.*

Relaxed competence-loyalty trade-off Consider next a lower level of ϕ. With $\phi = 0.45$ (instead of $\phi = 0.5$) the rise in mean competence and decline in mean loyalty in Figure 7 both come faster. Also, the increase in stability is quicker as reflected by the faster decline in r and a higher mean level of S throughout.

Lower survival probabilities Consider finally a different specification of the incumbent dynasty's survival probabilities, now defined as

$$S = \omega \left[zc + (1 - z)l \right]. \tag{32}$$

The baseline case thus corresponds to $\omega = 1$; see (10). When setting $\omega = 0.9$ there is no visible transition at all over the 100-period window shown. The reason is that much fewer dynasties survive long enough to accumulate sufficient competence to induce a new dynasty not to purge the inherited administration. In other words, exogenously increasing instability in this model endogenously generates more instability, in the form of more purges when there is a change in power.

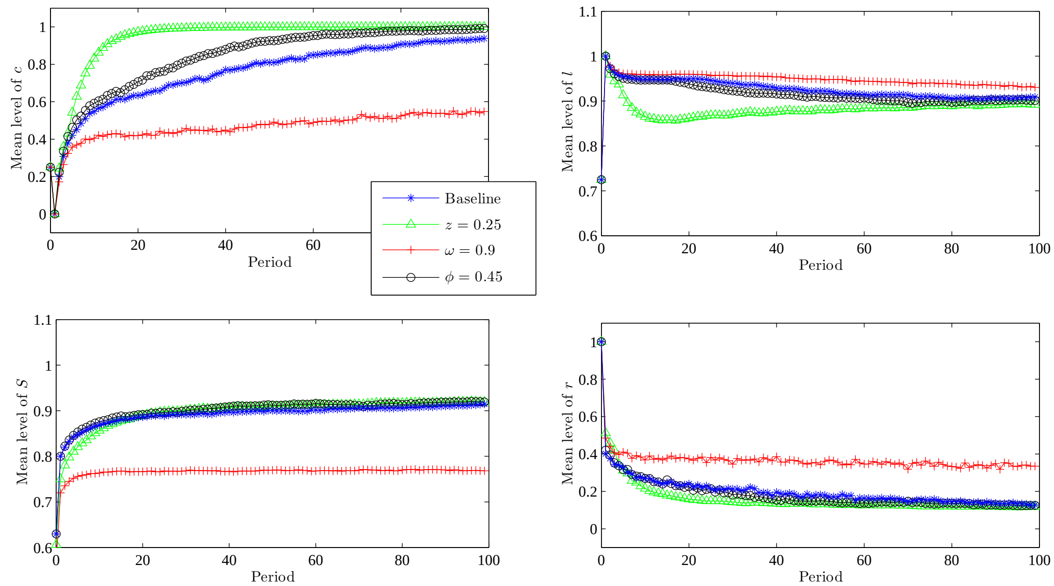

Figure 7: *Sensitivity analysis. The baseline case corresponds to the parameter values and initial conditions in Table 1 and* $\omega = 1$.

4 Conclusions

We have proposed a dynamic framework to study transitions from bad to better governance in a nondemocracy. Subsequent rulers replace one another, each seeking to influence his offspring's chances of staying in power by choosing how much and in what direction the composition of the administration is adjusted, in terms of its competence and loyalty.

The underlying idea is that loyalty is specific to a dynasty, whereas competence is not. By selecting administrators who are in some broad sense close to him, a ruler can insure that they are loyal to his ideological or biological offspring. However, to find competent administrators he needs to search over a longer ideological or biological distance. This generates a trade-off between competence and loyalty, as well as the size of the changes undertaken: big and rapid changes in the administration are costlier than gradual changes. Improving both loyalty and competence can be done only if changes are gradual, but destroying competence, by hiring only incompetent loyalists, can be done in one sweep.

When there is a change in power, the new ruler inherits an administration whose loyalty lies largely with the previous dynasty. He may therefore choose to purge the administration, filling the new positions with incompetent loyalists. He is less inclined to such a purge if more competence has been accumulated by the preceding dynasty. Intuitively, the higher is the inherited competence level, the costlier a purge is in terms of lost competence. This gives rise to path dependence. Once some dynasty has accumulated enough competence a transition takes place, away from repeated purges to a more stable no-purging path and a simultaneous rise in the levels of administrative competence.

One may interpret competence in our model broadly. A competent administrator may be less corrupt, more supportive of growth enhancing reforms, and provide better policy advice. Even more broadly, a larger fraction competent administrators may be interpreted as better institutions and political culture, or stronger state capacity (cf. Besley and Persson 2009, 2010). Such "capital" is indeed accumulated slowly, and can also (in some instances) be destroyed quickly, similarly to competence in our model. In that sense, our simple framework may offer insights into the links between political and economic development.

We imposed several strong but helpful assumptions when setting up this model. For example, we let the ruler care only about his immediate offspring's survival chances, and not members of the same dynasty ruling two periods ahead, which makes him more inclined to destroy competence capital.

We also assumed that a ruler can fire only a random sample of the administration. If he could instead primarily replace the disloyal and/or incompetent, he would be able to build up competence faster, which would speed up the transition to a no-purging path, although the transition would not be immediate, as long as the ruler faces a competence-loyalty trade-off in recruiting.

One could also let some competent administrators become incompetent between periods (capturing, e.g., aging or technical change, a form of depreciation of competence capital). This would slow accumulation of competence, thus delaying the transition to a no-purging path.

Still, the mechanisms that the model captures are intuitive and the model may serve as a starting point for thinking about other frameworks to study these issues.

APPENDIX

A Finding optimal r

A.1 Case (A): $c > \overline{c}$.

First note that $\widehat{S}(r; c, l) = S(r, 0; c, l)$ is linear in r for $r > \widehat{r}$. If $c > \overline{c}$, then $S(r, 0; c, l)$ is decreasing in r. In other words, conditional on recruiting only incompetent loyalists ($r > \widehat{r}$ and $q = 0$) the ruler wants to replace as few administrators as possible by setting r as low as possible. It follows that the maximum point cannot be greater than $\widehat{r}$, and must be on $[0, \widehat{r}]$. For $r \leq \widehat{r}$, we see from (16) that $\widehat{S}(r; c, l) = S(r, 1; c, l)$ is quadratic in r, and the first-order condition gives the (unconstrained) maximum point of $S(r, 1; c, l)$ as $r_{\max}$ defined in (19). Note that $r_{\max} > 0$ (for $c < 1$ and $l < 1$), so $\widehat{S}(r; c, l)$ cannot have a maximum point at $r = 0$. We must also examine if $r_{\max}$ falls below $\widehat{r}$. Suppose that $r_{\max} \geq \widehat{r}$. Comparing (15) and (19) we see that this is the case if, and only if, $c \leq \underline{c}$, where $\underline{c}$ is given by (18). But $c \leq \underline{c}$ cannot hold if $c > \overline{c}$. Thus, if $c > \overline{c}$, then the optimal choice of r is $r_{\max} < \widehat{r}$ and (14) implies that optimal q equals 1.

A.2 Case (B): $c < \underline{c}$.

From (15) and (19) we see that $c < \underline{c}$ implies $r_{\max} > \widehat{r}$. This means that $\widehat{S}(r; c, l) = S(r, 1; c, l)$ is increasing in r for $r < \widehat{r}$. In other words, conditional on recruiting only competent candidates ($r < \widehat{r}$ and $q = 1$) the ruler wants to replace as many administrators as possible by setting r as high as possible on $[0, \widehat{r}]$. Thus, the maximum point cannot be on $[0, \widehat{r})$. Moreover, for $c < \underline{c}$ (which implies $c < \overline{c}$) we can see from (16) that $\widehat{S}(r; c, l) = S(r, 0; c, l)$ is increasing in r for $r > \widehat{r}$. Thus, if $c < \underline{c}$, then the optimal choice of r is 1. Also, (14) then implies that optimal q equals 0.

A.3 Case (C): $c \in [\underline{c}, \overline{c}]$.

In this case $\widehat{S}(r; c, l)$ in (16) has two local maximum points, $r = 1$ and $r = r_{\max}$, where $r_{\max} \in [0, \widehat{r}]$. See Figure 2 for an illustration.

First note from (16) that

$$\widehat{S}(1; c, l) = S(1, 0; c, l) = 1 - z, \tag{A1}$$

which simply states that the offspring of a ruler who replaces the whole administration with incompetent loyalists (setting $r = 1$ and $q = 0$) survives with certainty if the threat is internal, which happens with probability $1 - z$; else the offspring is ousted with certainty.

Next, using (16) we see that $\widehat{S}(r_{\max}; c, l) = S(r_{\max}, 1; c, l) > (<, =)\widehat{S}(1; c, l) = 1 - z$ is equivalent to

$$zc+(1-z)l+r_{\max}[z(1-c)+(1-z)(1-l)]-(1-z)\phi\left(r_{\max}\right)^2 > (<,=)1-z, \quad \text{(A2)}$$

which can be written as

$$P(r_{\max}, D) > (<, =)0 \quad \text{(A3)}$$

where

$$P(r, D) = -\phi r^2 + \left(\frac{z}{1-z} + D\right) r - D, \quad \text{(A4)}$$

and

$$D = 1 - l - \left(\frac{z}{1-z}\right) c. \quad \text{(A5)}$$

Note from (17), (18) and (A5) that $c \in [\underline{c}, \overline{c}]$ implies that

$$D \in \left[0, \frac{z}{1-z}\right]. \quad \text{(A6)}$$

Note that $P(0, D) = -D \leq 0$, and $P(1, D) = [z/(1-z)] - \phi < 0$, since we have assumed that $z/(1-z) < \phi$ to ensure that $\widehat{r} < 1$; see (15). It can also be seen that $\partial P(r, D)/\partial r = 0$ at $r = r_{\max}$, and

$$\frac{\partial^2 P(r, D)}{\partial r^2} = -2\phi < 0. \quad \text{(A7)}$$

Thus, $P(r, D)$ is inversely U-shaped when graphed against r, reaching its maximum level, $P(r_{\max}, D)$, when $r = r_{\max}$. Next note from (A4) that an increase in D, at given $r < 1$, shifts $P(r, D)$ down:

$$\frac{\partial P(r, D)}{\partial D} < 0, \quad \text{(A8)}$$

for all $r < 1$. We can now define μ as the level of D for which the maximum level of $P(r, D)$ is exactly zero, i.e.,

$$P(r_{\max}, \mu) \equiv 0. \quad \text{(A9)}$$

That is, if c and l are such that $D = \mu$, then $\widehat{S}(r_{\max}; c, l) = \widehat{S}(1; c, l) = 1-z$; this is the case illustrated in Figure 2. More generally, from (A8) and (A9) and the definition of $P(r, D)$ the following chain of implications now follows:

$$D > (<, =)\mu \Leftrightarrow P(r_{\max}, D) < (>, =)0 \Leftrightarrow \widehat{S}(r_{\max}; c, l) < (>, =)\widehat{S}(1; c, l). \tag{A10}$$

In other words, if D lies above μ, then the ruler's optimal choice is $r = 1$ and $q = 0$ (a purge); if D falls below μ, then the ruler's optimal choice is $r = r_{\max}$, and $q = 1$. Recall from (A5) that D depends (linearly) on c and l, so if we can find an expression for μ in terms of exogenous parameters, then we have a (linear) condition determining whether $\widehat{S}(r_{\max}; c, l)$ or $\widehat{S}(1; c, l)$ is larger. Below we solve for μ in terms of the exogenous z and ϕ.

A.3.1 Finding an expression for μ

Using (19) and the definition of D in (A5) we can write $P(r_{\max}, D)$ in terms of D, z, and ϕ only. First note that

$$r_{\max} = \frac{z(1-c) + (1-z)(1-l)}{2\phi(1-z)} = \frac{\frac{z}{1-z} + D}{2\phi}. \tag{A11}$$

We can then use (A4) and (A11) to derive the following:

$$\begin{aligned}
&P(r_{\max}, D) \\
&= -\phi(r_{\max})^2 + \left(\tfrac{z}{1-z} + D\right) r_{\max} - D \\
&= -\phi(r_{\max})^2 + 2\phi(r_{\max})^2 - D \\
&= \phi(r_{\max})^2 - D \\
&= \phi\left(\frac{\frac{z}{1-z}+D}{2\phi}\right)^2 - D \\
&= \tfrac{1}{4\phi}\left(\tfrac{z}{1-z} + D\right)^2 - D \\
&= \tfrac{1}{4\phi}\left[\left(\tfrac{z}{1-z}\right)^2 + D^2 - 4\phi\left(1 - \tfrac{z}{2\phi(1-z)}\right) D\right].
\end{aligned} \tag{A12}$$

Evaluating (A12) at $D = \mu$ it follows that we can define

$$\begin{aligned}
\alpha &= \left(\tfrac{z}{1-z}\right)^2, \\
\gamma &= \phi\left[1 - \tfrac{z}{2\phi(1-z)}\right] > 0,
\end{aligned} \tag{A13}$$

and then write $P(r_{\max}, \mu) = 0$ as

$$\mu^2 - 4\gamma\mu + \alpha = 0, \tag{A14}$$

which has solutions

$$\mu = 2\gamma \pm \sqrt{4\gamma^2 - \alpha}. \tag{A15}$$

Recall from (A6) that we are here restricting attention to $D \leq z/(1-z)$. Thus, the threshold level of D that we are looking for, i.e. μ, must fall below $z/(1-z)$ too. Using (A13) some algebra then shows that we can rule out the larger root in (A15). We can then use (A13) to derive

$$4\gamma^2 - \alpha = 4\phi^2 \left[1 - \frac{z}{\phi(1-z)}\right], \tag{A16}$$

and

$$\mu = 2\gamma - \sqrt{4\gamma^2 - \alpha} = 2\phi \left[1 - \sqrt{1 - \frac{z}{\phi(1-z)}}\right] - \frac{z}{1-z}, \tag{A17}$$

which is an expression for the threshold level of D that we were seeking, in terms of the exogenous variables ϕ and z. Note from (A17) that

$$\left(\frac{1-z}{z}\right)\mu = \frac{2\phi(1-z)}{z}\left[1 - \sqrt{1 - \frac{z}{\phi(1-z)}}\right] = \frac{2}{\widehat{r}}\left[1 - \sqrt{1-\widehat{r}}\right] - 1 \equiv \Omega, \tag{A18}$$

where the second equality uses the notation in (15).

Thus, using (A5) we see that $D = 1 - l - [z/(1-z)]c = \mu$ is equivalent to

$$c = \left(\frac{1-z}{z}\right)(1 - \mu - l) = \left(\frac{1-z}{z}\right)(1-l) - \Omega \equiv \Psi(l), \tag{A19}$$

where the second equality uses (A18), and where $\Psi(l)$ is the same expression as in (20). More generally, it is now easily seen, using (A5) and (A10), that

$$c < (>,=)\Psi(l) \Leftrightarrow D > (<,=)\mu \Leftrightarrow \widehat{S}(r_{\max}; c, l) < (>,=)\widehat{S}(1; c, l). \tag{A20}$$

In other words, we have now shown that if $c < \Psi(l)$, then $\widehat{S}(r_{\max}; c, l) < \widehat{S}(1; c, l)$ and the ruler's optimal choice is $r = 1$ and $q = 0$; if $c > \Psi(l)$, then the ruler prefers $r = r_{\max}$ and $q = 1$. In the special case when $c = \Psi(l)$, the ruler is indifferent and we may assume that he chooses $r = r_{\max}$ and $q = 1$.

A.4 Cases (A) to (C) together

It can be seen that $\Omega \in [0, 1]$; see Section B of this appendix. From $\Omega \leq 1$ it follows that $\underline{c} \leq \Psi(l)$; see (A19) and (18). Thus, $c < \underline{c}$ implies $c < \Psi(l)$. Case (B) above, for which $r = 1$ and $q = 0$ in optimum, then becomes a special case of $c < \Psi(l)$.

Similarly, $\Omega \geq 0$, (A19) and (17) imply that $\overline{c} \geq \Psi(l)$. Thus Case (A) above, for which $r = r_{\max}$ and $q = 1$ in optimum, becomes a special cases of $c > \Psi(l)$.

B Showing that $\Omega \in [0, 1]$

Here we show that Ω in (A18) falls on the interval $[0, 1]$, which in turn can easily be seen to imply that $\Psi(l) \in [\underline{c}, \overline{c}]$. First note that $\Omega \geq 0$ is equivalent to the following inequalities:

$$\begin{aligned} & 1 \leq \tfrac{2}{\widehat{r}}\left[1-\sqrt{1-\widehat{r}}\right] \\ & \tfrac{\widehat{r}}{2} \leq 1-\sqrt{1-\widehat{r}} \\ & \sqrt{1-\widehat{r}} \leq 1-\tfrac{\widehat{r}}{2} > 0 \\ & 1-\widehat{r} \leq 1-\widehat{r}+\left(\tfrac{\widehat{r}}{2}\right)^2, \end{aligned} \tag{A21}$$

which always holds. Next note from (A18) that $\Omega \leq 1$ is equivalent to the following inequalities:

$$\begin{aligned} & 1 \geq \tfrac{2}{\widehat{r}}\left[1-\sqrt{1-\widehat{r}}\right]-1 \\ & 2 \geq \tfrac{2}{\widehat{r}}\left[1-\sqrt{1-\widehat{r}}\right] \\ & \widehat{r} \geq 1-\sqrt{1-\widehat{r}} \\ & \sqrt{1-\widehat{r}} \geq 1-\widehat{r} = \sqrt{1-\widehat{r}}\cdot\sqrt{1-\widehat{r}} \\ & 1 \geq \sqrt{1-\widehat{r}} \end{aligned} \tag{A22}$$

which always holds.

C Deriving l' when $c \geq \Psi(l)$

When $c \geq \Psi(l)$, and thus $r = r_{\max}$ and $q = 1$, we see from (12) that l' is given by

$$l' = l(1-r_{\max}) + r_{\max} - \phi(r_{\max})^2 = l + (1-l)r_{\max} - \phi(r_{\max})^2. \tag{A23}$$

First focus separately on the last two terms of (A23). Applying (19) we see that

$$(1-l)r_{\max} = \frac{z(1-c)(1-l)+(1-z)(1-l)^2}{2\phi(1-z)} = \frac{1}{4\phi}\left[\frac{2z(1-c)(1-l)}{1-z} + 2(1-l)^2\right], \tag{A24}$$

and

$$
\begin{aligned}
\phi(r_{\max})^2 &= \phi\left[\tfrac{z(1-c)+(1-z)(1-l)}{2\phi(1-z)}\right]^2 \\
&= \tfrac{z^2(1-c)^2+(1-z)^2(1-l)^2+2z(1-c)(1-z)(1-l)}{4\phi(1-z)^2} \\
&= \tfrac{1}{4\phi}\left[\left(\tfrac{z}{1-z}\right)^2(1-c)^2+(1-l)^2+\tfrac{2z(1-c)(1-l)}{1-z}\right].
\end{aligned}
\tag{A25}
$$

Now (A23), (A24) and (A25) show that

$$
\begin{aligned}
l' &= l+\tfrac{1}{4\phi}\left[\tfrac{2z(1-c)(1-l)}{1-z}+2(1-l)^2-\tfrac{z^2(1-c)^2}{(1-z)^2}-(1-l)^2-\tfrac{2z(1-c)(1-l)}{1-z}\right] \\
&= l+\tfrac{1}{4\phi}\left[(1-l)^2-\left(\tfrac{z}{1-z}\right)^2(1-c)^2\right].
\end{aligned}
\tag{A26}
$$

D Showing that the path cannot lead from Region II to I

In Region II it holds that $c \geq \Psi(l)$ and $c < \Gamma(l)$, implying from (25), (26) and (31) that $c' \geq c$ and $l' < l$. To show that c' and l' cannot lie in Region I we need to show that $c' \geq \Psi(l')$. From (26) we can write the fraction disloyal in the next period as

$$
1-l' = 1-l+\frac{1}{4\phi}\left[\left(\frac{z}{1-z}\right)^2(1-c)^2-(1-l)^2\right], \tag{A27}
$$

which together with (20) gives

$$
\begin{aligned}
\Psi(l') &= \left(\tfrac{1-z}{z}\right)(1-l')-\Omega \\
&= \left(\tfrac{1-z}{z}\right)(1-l)-\Omega+\tfrac{1-z}{4\phi z}\left[\left(\tfrac{z}{1-z}\right)^2(1-c)^2-(1-l)^2\right] \\
&= \Psi(l)+\tfrac{1-z}{4\phi z}\left[\left(\tfrac{z}{1-z}\right)^2(1-c)^2-(1-l)^2\right] \\
&= \Psi(l)+\tfrac{z(1-c)^2}{4\phi(1-z)}-\tfrac{(1-z)(1-l)^2}{4\phi z},
\end{aligned}
\tag{A28}
$$

where the third equality uses $\Psi(l) = \left(\frac{1-z}{z}\right)(1-l)-\Omega$ from (20). Next, (25) gives the fraction competent in the next period as

$$
\begin{aligned}
c' &= c+(1-c)\left[\tfrac{z(1-c)+(1-z)(1-l)}{2\phi(1-z)}\right] \\
&= c+\tfrac{2z(1-c)^2}{4\phi(1-z)}+\tfrac{(1-l)(1-c)}{2\phi}.
\end{aligned}
\tag{A29}
$$

Using (A28) and (A29) we see that $c' \geq \Psi(l')$ can be written as

$$c + \frac{2z(1-c)^2}{4\phi(1-z)} + \frac{(1-l)(1-c)}{2\phi} \geq \Psi(l) + \frac{z(1-c)^2}{4\phi(1-z)} - \frac{(1-z)(1-l)^2}{4\phi z}, \quad \text{(A30)}$$

or

$$c - \Psi(l) + \frac{z(1-c)^2}{4\phi(1-z)} + \frac{(1-l)(1-c)}{4\phi} + \frac{(1-z)(1-l)^2}{4\phi z} \geq 0, \quad \text{(A31)}$$

which always holds, since $c \geq \Psi(l)$, and the remaining three terms are all non-negative.

E How $\Psi(l)$ shifts when changing z and ϕ

The expression for $\Psi(l)$ in (20), using the expressions for Ω in (21) and $\widehat{r} = z/[\phi(1-z)]$ in (15), can be rewritten as

$$\begin{aligned} \Psi(l) &= \left(\tfrac{1-z}{z}\right)(1-l) + 1 - \tfrac{2}{\widehat{r}}\left[1 - \sqrt{1-\widehat{r}}\right] \\ &= \left(\tfrac{1-z}{z}\right)(1-l) + 1 - \tfrac{2\phi(1-z)}{z}\left[1 - \sqrt{1 - \tfrac{z}{\phi(1-z)}}\right] \\ &= \left(\tfrac{1-z}{z}\right)(1-l) + 1 - \tfrac{2\phi(1-z)}{z}\left[1 - \sqrt{1 - \tfrac{z}{\phi(1-z)}}\right] \\ &= \tfrac{1-l}{x} + 1 - \tfrac{2\phi}{x}\left[1 - \sqrt{1 - \tfrac{x}{\phi}}\right] \equiv \widetilde{\Psi}(l, x, \phi), \end{aligned} \quad \text{(A32)}$$

where we have defined $x = z/(1-z)$. Note that $\widehat{r} = z/[\phi(1-z)] = x/\phi < 1$ (by earlier assumption) and that $\partial x/\partial z > 0$. To show that $\Psi(l)$ shifts down when z (or x) increases, and up when ϕ increases, the task is to show that $\partial\widetilde{\Psi}(l, x, \phi)/\partial x < 0$ and $\partial\widetilde{\Psi}(l, x, \phi)/\partial\phi > 0$. First note that

$$\begin{aligned} \tfrac{\partial\widetilde{\Psi}(l,x,\phi)}{\partial x} &= -\tfrac{(1-l)}{x^2} - 2\phi\left[\frac{-\frac{1}{2}\left(1-\frac{x}{\phi}\right)^{-\frac{1}{2}}\left(-\frac{x}{\phi}\right) - \left[1-\left(1-\frac{x}{\phi}\right)^{\frac{1}{2}}\right]}{x^2}\right] \\ &= -\tfrac{(1-l)}{x^2} - \frac{2\phi}{x^2\left(1-\frac{x}{\phi}\right)^{\frac{1}{2}}}\left[\tfrac{x}{2\phi} - \left(1 - \tfrac{x}{\phi}\right)^{\frac{1}{2}} + \left(1 - \tfrac{x}{\phi}\right)\right] \\ &= -\tfrac{(1-l)}{x^2} - \frac{2\phi}{x^2\left(1-\frac{x}{\phi}\right)^{\frac{1}{2}}}\left[1 - \tfrac{x}{2\phi} - \left(1 - \tfrac{x}{\phi}\right)^{\frac{1}{2}}\right] < 0, \end{aligned} \quad \text{(A33)}$$

where it can be seen that the expression in square brackets following the last equality is positive. Then note that

$$
\begin{aligned}
\frac{\partial\widetilde{\Psi}(l,x,\phi)}{\partial\phi} &= -\frac{2}{x}\left[1-\left(1-\frac{x}{\phi}\right)^{\frac{1}{2}}\right] - \frac{2\phi}{x}\left[-\frac{1}{2}\left(1-\frac{x}{\phi}\right)^{-\frac{1}{2}}\left(-\frac{x}{\phi^2}\right)(-1)\right] \\
&= -\frac{2}{x}\left[1-\left(1-\frac{x}{\phi}\right)^{\frac{1}{2}}\right] + \frac{2}{x}\left[\frac{1}{2}\left(1-\frac{x}{\phi}\right)^{-\frac{1}{2}}\frac{x}{\phi}\right] \qquad \text{(A34)} \\
&= \frac{2}{x\left(1-\frac{x}{\phi}\right)^{\frac{1}{2}}}\left[1-\frac{x}{2\phi}-\left(1-\frac{x}{\phi}\right)^{\frac{1}{2}}\right] > 0,
\end{aligned}
$$

where (again) it can be seen that the expression in square brackets following the last equality is positive.

References

[1] Acemoglu, D. and J.A. Robinson, 2000, Why did the West extend the franchise? Democracy, inequality, and growth in historical perspective, Quarterly Journal of Economics 115, 1167-1199.

[2] Acemoglu, D. and J.A. Robinson, 2001, A theory of political transitions, American Economic Review 91, 938-963.

[3] Acemoglu, D. and J.A. Robinson, 2006, Economic Origins of Dictatorship and Democracy, Cambridge University Press, New York.

[4] Acemoglu, D., G. Egorov, and K. Sonin, 2008, Coalition formation in non-democracies, Review of Economic Studies 75, 987-1009.

[5] Acemoglu, D., G. Egorov, and K. Sonin, 2009, Do juntas lead personal rule?, American Economic Review Papers and Proceedings 99, 298-303.

[6] Acemoglu, D., G. Egorov, and K. Sonin, 2010, Political selection and persistence of bad governments, Quarterly Journal of Economics 125, 1511-1575.

[7] Acemoglu, D., S. Johnson, J.A. Robinson, and P. Yared, 2008, Income and democracy, American Economic Review 98, 808-842.

[8] Acemoglu, D., D. Ticchi and A. Vindigni, 2010, A theory of military dictatorships, American Economic Journal: Macroeconomics 2, 1–42.

[9] Acemoglu, D., D. Ticchi and A. Vindigni, 2011, Emergence and persistence of inefficient states, Journal of the European Economic Association, forthcoming.

[10] Alesina, A., and R. Perotti, 1996, Income distribution, political instability, and investment, European Economic Review 40, 1203-1228.

[11] Alesina, A., S. Özler, N. Roubini, and P. Swagel, 1996, Political instability and economic growth, Journal of Economic Growth 1, 189-211.

[12] Besley, T., and S. Coate, 1997, An economic model of representative democracy, The Quarterly Journal of Economics 112, 85-114.

[13] Besley, T., and T. Persson, 2008, Wars and state capacity, Journal of the European Economic Association 6, 522-530.

[14] Besley, T., and T. Persson, 2009, The origins of state capacity: Property rights, taxation, and policy, American Economic Review 99, 1218-1244.

[15] Besley, T., and T. Persson, 2010, State capacity, conflict and development, Econometrica 78, 1-34.

[16] Cervellati. M., P. Fortunato, and U. Sunde, 2008, Hobbes to Rousseau: inequality, institutions and development, Economic Journal 118, 1354-1384.

[17] Clague, C., P. Keefer, S. Knack, and M. Olson, 1996, Property and contract rights in autocracies and democracies, Journal of Economic Growth 1, 243-276.

[18] Downs, A, 1957, An Economic Theory of Democracy, New York: Harper Collins.

[19] Egorov, G., and K. Sonin, 2011, Dictators and their viziers: endogenizing the loyalty-competence trade-off, Journal of the European Economic Association, forthcoming.

[20] Galor, O., and D. Weil, 2000, Population, technology, and growth: from Malthusian stagnation to the demographic transition and beyond, American Economic Review 90, 806-829.

[21] Gradstein, M., 2007, Inequality, democracy and the protection of property rights, Economic Journal 117, 252-269.

[22] Gradstein, M., 2008, Institutional traps and economic growth, International Economic Review 49, 1043-1066.

[23] Grossman, H.I., 2000, The state: agent or proprietor?, Economics of Governance 1, 3-11.

[24] Grossman, H.I., and S.J. Noh, 1994, Proprietary public finance and economic welfare, Journal of Public Economics 53, 187-204.

[25] Hansen, G.D., and E.C. Prescott, 2002, Malthus to Solow, American Economic Review 92, 1205-1217.

[26] Lagerlöf, N.P., 2003, From Malthus to modern growth: can epidemics explain the three regimes?, International Economic Review 44, 755-777.

[27] Lagerlöf, N.P., 2010, From Malthusian war to Solovian peace, Review of Economic Dynamics 3, 616-636.

[28] Lipset, S., 1959, Some social prerequisites for democracy: Economic development and political legitimacy, American Political Science Review 53, 69-105.

[29] Marr, P., 2004, The Modern History of Iraq, Westview Press, Boulder, Colorado.

[30] Mauro, P., 1995, Corruption and growth, Quarterly Journal of Economics 110, 681-712.

[31] Montefiore, S.S., 2003, Stalin: The Court of the Red Tsar, Knopf, New York.

[32] Moore, B., 1966, The Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World, Beacon Press, Boston.

[33] Mueller, H., 2009, Patronage and meritocracy: political institutions and bureaucratic efficiency, mimeo, Institut d'Analisi Economica, CSIC.

[34] Mulligan, C.B., and K.K. Tsui, 2008, Political Entry, Public Policies, and the Economy, NBER Working Paper No. 13830.

[35] North, D., and B. Weingast, 1989, Constitutions and commitment: the evolution of institutions governing public choice in seventeenth century England, Journal of Economic History 49, 803-832.

[36] Oh, J.K.-C., 1968, Korea – Democracy on Trial, Cornell University Press, Ithaca, New York.

[37] Olson, M., 1993, Dictatorship, democracy, and development, American Political Science Review 87, 567-576.

[38] Osborne, M.J., and A. Slivinski, 1996, A model of political competition with citizen-candidates, The Quarterly Journal of Economics 111, 65-96.

[39] Packer, G., 2005, The Assassins' Gate: America in Iraq, Farrar, Straus and Giroux, New York.

[40] Persson, T., and G. Tabellini, 2000, Political Economics: Explaining Economic Policy, MIT Press, Cambridge, Massachusetts.

[41] Przeworski, A., M.E. Alvarez, J.A. Cheibub, and F. Limongi, 2000, Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990, Cambridge University Press, Cambridge, United Kingdom.

[42] Svensson, J., 2005, Eight questions about corruption, Journal of Economic Perspectives 19, 19-42.

[43] Tullock, G., 1987, Autocracy, Klumer Academic Publishers, Boston.

[44] Weitzman, M., 1970, Iterative multilevel planning with production targets, Econometrica 38, 50-65.