DO ELECTIONS ALWAYS MOTIVATE INCUMBENTS?

Eric le Borgne
And
Ben Lockwood

No 580

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

Do Elections Always Motivate Incumbents?

Eric Le Borgne and Ben Lockwood[*]
University of Warwick

First version:[†] April 2000
This version: November 2000

Abstract

This paper studies a principal-agent model of the relationship between office-holders and the electorate, where the office-holder is initially uninformed about her ability (following Holmström, 1999). If office-holder effort and ability interact in the "production function" that determines performance in office, then an office-holder has an incentive to experiment, i.e. raise effort so that performance becomes a more accurate signal of her ability. Elections reduce the experimentation effect, and the reduction in this effect may more than offset the positive "career concerns" effect of elections on effort. Moreover, when this occurs, appointment of officials (random selection from the citizenry and tenure) may Pareto-dominate elections.

Keywords: Career Concerns, Elections, Citizen-Candidate, Experimentation, Tournaments, Political Business Cycles.

JEL Classification Numbers: D72, D78, H41, J44, J45.

---

[*]Address for correspondance: Department of Economics, University of Warwick, Coventry CV4 7AL, United Kingdom. E-mails: eric.le-borgne@warwick.ac.uk, and B.Lockwood@warwick.ac.uk.

[†]With a different title: "The Career Concerns of Politicians: Efficiency in a Representative Democracy?".

1. Introduction

In recent years, economists and political scientists have applied principal-agent theory to study the relationship between voters and elected officials. The literature starts from the idea that there is a moral hazard problem between the elected official and the electorate: left to his own devices, the official will pursue his own interests, rather than those of the voters. This is modelled formally by supposing that the official can supply unobservable effort (Ferejohn, 1986; Austen-Smith and Banks, 1989; Banks and Sundaram, 1993, 1998) or has the opportunity to "steal" rent from tax revenue (Barro, 1973; Persson and Tabellini, 2000). However, this literature modifies standard principal-agent theory in two crucial ways. First, unlike employees,[1] elected officials cannot typically be offered monetary rewards for their performance on the job: the salaries of political office are usually independent of short-term performance. Second, dismissal (losing elections) is costly.

Under these two conditions, officials can only be motivated (to supply additional effort, to steal less rent) by "career concerns",[2] i.e. the fear of losing elections. The recent literature in this area has modelled this process formally, starting with the seminal work of Barro (1973) and Ferejohn (1986). This literature now comprises a variety of models (discussed in more detail in the Conclusion) but with two apparently very robust conclusions; (i) in (sequential) equilibrium, voters follow a cutoff rule, i.e. will only re-elect the incumbent if his observed performance is above a certain critical level; (ii) the cutoff rule always motivates the office-holder (to supply more effort, or extract less rent).

This paper argues that conclusion (ii) is in fact not robust. We present a simple two-period model of the agency problem between the electorate and the voters, drawing on the work of Holmström (1982, 1999) and Dewatripont, Jewitt, and Tirole (1999), and show that in sequential equilibrium, elections may demotivate: that is, the incumbent will supply less effort than without the "discipline" of an election. The intuition is simple. When the ability and the effort of the office-holder interact positively, the office-holder can learn more about his ability by supplying more effort. We call this the experimentation motive for supplying effort. However, if he is exposed to the possible future loss of office, his motive to experiment will be reduced. This diminution in the experimentation motive may more than offset the increase in effort induced by the desire to signal competence to the electorate (the career concerns effect). One way of interpreting the diminution is as short-termism; the incumbent underinvests, anticipating he will lose power (see also Besley and Coate, 1998, for examples of this type).

[1] Within a firm, various incentive mechanisms ranging from promotion and demotion, wage changes, performance contracts (e.g. stock options), are widely used (see Prendergast, 1999, and Gibbons and Waldam, 1999, for recent surveys).

[2] Career concerns refer to the fact that an agent's current actions (e.g. labour supply, effort on the job) are in part determined by taking into account the effect that these actions have on the agent's future career prospects even though no explicit incentives (e.g. performance contracts) links the two.

The existence of, and implications of, experimentation in a career-concerns setting is (as far as we know) a new …nding. This is because the existing literature assumes either (i) that potential o¢ce-holders are already fully (privately) informed about their ability, as in Banks and Sundaram (1993, 1998); (ii) an additive technology, where information has no value (Holmström, 1999); (iii) one period only, in which case information acquired currently cannot be used in the future (Dewatripont et al., 1999); (iv) there is no noise in the production function so that incumbents can perfectly observe their ability from performance at the end of the …rst period of o¢ce (Persson and Tabellini, 2000).

Our model is the following.[3] The economy is populated by a number of citizens, who may vary in competence if in political o¢ce. Their performance in o¢ce is described by a production function that maps competence, e¤ort, and a random shock into a scalar variable, the "public good". Following Holmström (1982, 1999), Dewatripont, Jewitt, and Tirole (1999), we assume that citizens do not know their own competence, but can only infer it from their performance in o¢ce.[4] Initially, we compare two institutions in this setting. The …rst is appointment, where the o¢ce-holder is randomly selected from the population, and is in o¢ce for two periods. The second is democracy, which di¤ers from the appointment in that at the beginning of the second period, there is an election, contested by the …rst-period incumbent and an opponent, the latter randomly selected from the population, with the winner holding o¢ce for the second period.

Our …rst …nding is that in this setting, even with appointment, e¤ort may vary over time, due to experimentation, which occurs when the incumbent deviates from the myopically optimal action that just maximises the current payo¤ in order to improve the information content of his signal about his own ability, namely the output of the public good.[5] We show that experimentation will occur in the …rst period of our two-period setting if and only if e¤ort and ability interact in the production of the public good, and that when it occurs, it unambiguously induces the o¢ce-holder to put in more e¤ort than

[3] The closest model to ours is the career concerns model of Chapter 4.5 of Persson and Tabellini (2000), which also builds on Holmstrom's work, and that we saw after the …rst draft of this paper was completed. We show in Section 7 that (subject to some inessential quali…cations) their model can be considered as a special case of ours where there is no randomness in the production function. As a consequence, in their model, the incumbent can perfectly observe his competence at the end of the …rst period of o¢ce, and so there is no experimentation e¤ect, which is the main topic of this paper. Our reading of their model is that it is intentionally kept very simple to permit an easy analysis of the way career concerns are a¤ected by electoral rules.

[4] This is in contrast to the more usual assumption in the principal-agent literature, which assumes that agents are privately informed. Empirically, there is support for both assumptions. One proxy for the self-knowledge of an elected o¢cial is the amount of time spent doing the job. In the US, all presidents except F.D.Roosevelt have served either one or two terms. On the other hand, there is considerable variation in o¢ce tenure in the UK, ranging from less than one year for Alec Douglas-Hume to 11 years for Margaret Thatcher.

[5] The experimentation literature initially studied the problem of a monopolist facing an unknown demand curve (Prescott (1972) and Grossman, Kihlstrom, and Mirman (1977) are early contributions). Mirman, Samuelson, and Urbano (1993) develop a tractable two-period monopolist game and establish conditions under which experimentation occurs. We make use of their results below. Keller and Rady (1999) surveys the literature.

the myopic level.

Our key observation is that when we move from appointment to democracy, the incentive to experiment unambiguously falls, for the reason described above. Of course, in our model, as in others in the literature, elections also have a positive e¤ect on equilibrium e¤ort via career concerns e¤ect;[6] the better the observable performance while in o¢ce, the higher the probability of being re-elected and therefore the higher the expected payo¤ in the future.

As we show, it is possible that the loss of the incentive to experiment may more than o¤set the career concerns e¤ect, so that equilibrium e¤ort may be lower in democracy than with appointment. More generally, we can say that career concerns and experimentation, while both inducing the incumbent to increase e¤ort, are substitutes under symmetric incomplete information: that is, democracy introduces career concerns, but also necessarily reduces the incentive to experiment.

In Section 5, we show that there is also an important relationship between the "e¢ciency" of equilibrium with democracy and the presence of an experimentation motive. Consider a constrained social planner who only knows the distribution of the competency variable initially (so he is only as well-informed as the citizens), and has the same powers as citizens, i.e. can "…re" the incumbent if performance falls below some cuto¤ value. We say that democracy (with or without endogenous entry of candidates) is constrained e¢cient[7] if a constrained social planner cannot make every citizen better o¤. It turns out that (subject to a uniqueness condition holding) when technology is additive (so there is no experimentation motive), the equilibrium with democracy is constrained e¢cient, but that this need not be the case with an experimentation motive.

A subsidiary objective of this paper is to address two other, related, weaknesses of the existing literature on the principal-agent relationship between voters and o¢ce-holders, namely (i) that the o¢ce-holders (the incumbent and challenger) are assumed to be randomly drawn from some population; and (ii) have di¤erent preferences than the voters.[8] Our model has already addressed the second problem, by having the incumbent and challenger randomly selected from the same population as the electorate. In Section 4 of the paper, we extend our model of democracy to deal with (i).

Speci…cally, we study democracy with endogenous (candidate) entry, where at the beginning of each of the two periods, any citizen can stand for election, and candidates

[6] In our model, the career concerns e¤ect can also be thought of as a "tournament" between incumbent and challenger (Lazear and Rosen, 1981; Green and Stokey, 1983). Whoever wins o¢ce gets "…rst prize" ; and whoever loses, "second prize".

[7] Of course, due to the underlying agency problem, the equilibrium outcome with democracy will never be …rst-best e¢cient, so that the latter is not a very interesting benchmark.

[8] In Banks and Sundaram (1993, 1998), the principal (voters) care about the output of the agent, but the agent's payo¤ is independent of this output. The same is true of Persson and Tabellini (2000), where voters care about the output of the public good, but the o¢ce-holder cares only about an exogenous ego-rent and the rents that he can extract from tax revenue.

are voted on by plurality rule, with the winner taking office for one period (becoming the office-holder). So, this approach combines the citizen-candidate modelling of selection of office-holders (Besley and Coate, 1997) with the principal-agent relationship between office-holder and voters.[9] It turns out that, given the information structure assumed, our main results do not change qualitatively. In particular, as the candidate entry stage cannot reveal any information to voters about their competence in office, there will still be experimentation in office. Candidate entry and voting (for all candidates in the first period, and for the challenger in the second) will be determined by other characteristics of the candidates.[10]

The rest of the paper is structured as follows. Section 2 describes the model. Section 3 presents the basic results, and Section 4 extends them to the case of democracy with endogenous entry. Section 5 is devoted to normative analysis. Section 6 discusses some extensions. Finally, Section 7 concludes and discusses related literature.

2. The Model

2.1. Technology

The economy is populated by a set N of citizens with $\#N = n \geq 3$ and evolves over two time periods, $t = 0, 1$. There is a political office that can only be occupied by one citizen, the "office-holder". The performance of the office-holder while in office is measured by a scalar variable $g \in \Re$ which we call the "public good".

The ability of an office-holder $i \in N$ is measured by θ_i, and his effort level in period t is $a_{i,t} \in [0, 1)$. Following Dewatripont, Jewitt, and Tirole (1999), this office-holder produces g_t units of the public good, where :

$$g_t = \mu(\theta_i + a_{i,t}) + (1 - \mu)\theta_i a_{i,t} + \varepsilon_t, \quad t = 0, 1 \tag{2.1}$$

where $\mu \in [0, 1]$. Also, $\varepsilon_0, \varepsilon_1$ are independently distributed random shocks. In either period, the office-holder has to decide on a level of effort before observing ε_t.

Note that the general production function (2.1) encompasses two important special cases. The first is where $\mu = 1$, in which case the technology is purely additive (as in Holmström, 1999). The second is where $\mu = 0$, in which case the technology is purely multiplicative (in the sense of Dewatripont, Jewitt and Tirole, 1999).

We assume that each θ_i is a random draw from a distribution that can take two values: $\theta_H > \theta_L > 0$ with probabilities $\pi_0, 1 - \pi_0$ respectively. This draw takes place at the beginning of period zero. So, the θ_i are uncorrelated across citizens. We refer to H, L as the types of the citizens.

[9] This is explored in more detail in a companion paper, Le Borgne and Lockwood (2000).

[10] Following Rogoff and Sibert (1988), Rogoff (1990), we allow voters to differ in "looks", i.e. characteristics that voters value but are unrelated to competence in office and therefore economic issues.

We assume that ε has a continuous distribution with probability density function f, cumulative distribution function F, and has full support on $\Re$. We assume that f satisfies the Monotone Likelihood Ratio Condition (MLRC) that $f'(\varepsilon)/f(\varepsilon)$ is a continuous and decreasing function.[11] We also assume that

A0. For any $a > 0$, there exists $\varepsilon', \varepsilon'', \ \varepsilon'' > \varepsilon'$, such that $\frac{f(\varepsilon' - a)}{f(\varepsilon')} < 1 < \frac{f(\varepsilon'' - a)}{f(\varepsilon'')}$.

It is well-known that a large number of distributions satisfy the MLRC (Milgrom, 1981), including the Normal, and it is easy to check that if ε is Normally distributed, A0 is also satisfied.

Our production function, plus the assumption that $\varepsilon_t \in \Re$, of course implies that g can be negative, and so cannot be literally interpreted as a public good in a public finance model. The reason for allowing shocks ε_t to be negative is that if we constrained ε_t to be positive, i.e. by assuming the lower bound of the support of ε_t to be zero, then if the incumbent observed $g_t < \gamma\theta_H$, he could be sure his type was low. This problem of "perfect inference" would complicate the analysis considerably. The simplest way to model non-negativity for g_t is to suppose that the random shock is multiplicative, i.e.

$$g_t = [\gamma(\theta_i + a_{i,t}) + (1 - \gamma)\theta_i a_{i,t}]\,\varepsilon_t \tag{2.2}$$

and has support $[0, \infty)$. The qualitative features of the analysis of this paper would be unchanged if we worked with (2.2).

2.2. Preferences

If $i \in N$ is an office-holder in period t, and produces g_t, then $j \neq i$ only cares about the level of performance of the office-holder, i.e. $u_{j,t} = g_t$. If an agent $i \in N$ is an office-holder in period t, she has payoff $u_{i,t} = g_t + R + rg_t - c(a_{i,t})$, where g_t is the net utility from the public good, as for $j \neq i$; $R + rg_t$ is an "ego-rent" from being in office (as in Rogoff and Sibert, 1988), deriving from the prestige in managing public affairs, and finally $c(a_{i,t})$ is the cost of effort. If $r > 0$, the ego-rent interacts positively with the amount of public good provided.[12] Following Rogoff and Sibert (1988), we assume for the moment that $r = 0$ (the case of $r > 0$ is discussed in Section 6.1 below). Also, we assume that $c(.)$ is strictly increasing and strictly convex, and[13] $c(0) = 0$, $c'(0) < 1$.

---

[11] The MLRC says that, for a given competency type, a high effort increases the probability of obtaining a high visible performance at least as much as it increases the probability of obtaining a low visible performance variable.

[12] Of course, $r > 0$ could also model a public duty/altruistic motive for the office-holder, capturing the fact that holders of public office may feel some obligation towards the citizens they represent, quite independently from the discipline that elections impose.

[13] The last condition $c'(0) < 1$ ensures that myopic effort is positive.

2.3. Institutions

The agent whose task it is to produce the public good (the office-holder) is selected in one of two ways. We allow for a third institution in Section 4.

1. Appointment

At the beginning of period $t = 0$, the office-holder is selected by random draw from the set of citizens, and is in place for both periods.

2. Democracy

At the beginning of period $t = 0$; an office-holder (the incumbent) is selected by random draw from the set of citizens. This office-holder is in place during period $t = 0$ but faces an election at the beginning of period 1. At this stage, an opponent is selected by random draw from the set of remaining citizens. The citizens then vote on the opponent versus the incumbent, and the winner is the office-holder in period $t = 1$.

Our modelling of democracy abstracts from the entry decisions of candidates (dealt with in Section 4 below) while allowing the electorate to "fire" bad office-holders. It also is quite close to the modelling of the electoral process in Rogoff and Sibert (1988), and Rogoff (1990).

In all cases, for consistency, we will impose the individual rationality condition that the office-holder prefers to be in office than not.

2.4. Information Structure

Following Holmström (1999), and Dewatripont, Jewitt and Tirole (1999), we assume that citizens do not know $\theta = (\theta_1, \ldots, \theta_n)$, but all know the joint distribution of θ (symmetric incomplete information). It is also assumed that the action a is only observable by the incumbent. Because of this, the office-holder cannot be rewarded on the basis of a: If she receives a salary, this is modelled as a component of R; the "ego-rent". It is also assumed that g is not verifiable, so the office-holder cannot be rewarded on the basis of g:

2.5. Myopic Choice of Effort

Consider the choice of effort by an office-holder who is in power for one period only, and believes he is high-ability with probability π: This office-holder solves the problem

$$v_0(\pi) = \max_a \left\{ \begin{array}{c} \pi[\mu(\theta_H + a) + (1-\mu)\theta_H a] \\ +(1-\pi)[\mu(\theta_L + a) + (1-\mu)\theta_L a] - c(a) + R \end{array} \right\} \quad (2.3)$$

The first-order condition is

$$\mu + (1-\mu)[\pi\theta_H + (1-\pi)\theta_L] - c'(a) = 0 \quad (2.4)$$

This solves to give $a^*(\pi)$; which we call the myopic optimal action by the office-holder, given a belief that he is competent with probability π: If $\mu = 1$; $a^*(\pi) \equiv a^*$, for all π.

Finally, we can define the utility of the non-office-holding citizen when both the citizen and the office-holder believe the office-holder to be competent with probability π;

$$v_c(\pi) = \pi[\beta(\theta_H + a^*(\pi)) + (1-\beta)\theta_H a^*(\pi)] + (1-\pi)[\beta(\theta_L + a^*(\pi)) + (1-\beta)\theta_L a^*(\pi)] \quad (2.5)$$

Some useful properties of a^* and the associated value functions v_o, v_c are the following. First, it is clear from the first-order condition (2.4) that

$$\frac{\partial a^*}{\partial \pi} = \frac{(1-\beta)(\theta_H - \theta_L)}{c''(a^*)} \quad (2.6)$$

So, a^* is independent of π if the technology is purely additive and strictly increasing in π otherwise.

Second, by direct application of the envelope theorem to (2.3), we have

$$v_o'(\pi) = \beta(\theta_H - \theta_L) + (1-\beta)(\theta_H - \theta_L)a^*(\pi) \quad (2.7)$$

so v_o is strictly increasing in π. By inspection of (2.6), v_c is also strictly increasing in π. Moreover, as $R > 0$, and by the properties of c, $v_o(\pi) > 0$, and by inspection, $v_c(\pi) > 0$.

3. Positive Analysis

3.1. Appointment

We solve the appointee's decision problem with the usual dynamic programming approach. In the second period, the appointee faces a myopic problem, so chooses $a_1 = a^*(\pi_1)$ where π_1 is the appointee's posterior belief that he is a high-ability type. The individual rationality condition for the appointee is that $v_o(\pi_1) \geq 0$ which is always satisfied.

Now, note from (2.7) above that as long as the technology has a multiplicative component, i.e. $\beta < 1$, his second-period payoff is strictly convex in π_1;

$$v_o''(\pi_1) = (1-\beta)(\theta_H - \theta_L)\frac{\partial a^*}{\partial \pi_1} > 0 \quad (3.1)$$

This means that information about θ obtained by Bayesian updating is strictly valuable. Now when updating, the appointee can observe both his own output of the public good in the first period, g_0, and his action in the first period, a_0. So, from Bayes' rule, the appointee's posterior belief that he is a high-type is

$$\pi_1(a_0; g_0) = \Pr(\theta = \theta_H \mid a_0; g_0) = \frac{\pi_0}{\pi_0 + (1-\pi_0)\, f_L(g_0; a_0)/f_H(g_0; a_0)} \quad (3.2)$$

where

$$f_k(g_0; a_0) \equiv f\left(g_0 - (1-\beta)\theta_k a_0 - \beta(\theta_k + a_0)\right);\ k = H; L \quad (3.3)$$

Note from (3.2) that changes in actions are informative, i.e. a change in a_0 affects the posterior probability that the office-holder is competent, given output ($\partial\pi_1(g_0, a_0)/\partial a_0 \neq 0$). So, the two well-known[14] conditions for optimal experimentation are satisfied in our model, i.e. the appointee has an incentive to deviate from the myopic effort level in the first period.

Now we go to the first-period problem for the appointee. Note that for a given value of a_0, g_0 is a random variable with distribution function

$$\begin{aligned} H(g_0, a_0) &= \pi_0 F(g_0 - (1-\mu)\theta_H a_0 - \mu(\theta_H + a_0)) + \\ &\quad (1-\pi_0) F((g_0 - (1-\mu)\theta_L a_0 - \mu(\theta_L + a_0)) \end{aligned} \tag{3.4}$$

Consequently, $\pi_1(g_0, a_0)$ is also a random variable, conditional on a_0, implying an expected optimized second-period payoff of $E_{g_0}[v_o(\pi_1(a_0, g_0))]$. So, the problem for the appointee in the first period is

$$\max_{a_0} \left\{ \begin{array}{c} \pi_0[\mu(\theta_H + a_0) + (1-\mu)\theta_H a_0] + (1-\pi_0)[\mu(\theta_L + a_0) + (1-\mu)\theta_L a_0] \\ - c(a_0) + R + E_{g_0}[v_o(\pi_1(a_0, g_0))] \end{array} \right\} \tag{3.5}$$

The first-order condition can be written

$$[\mu + (1-\mu)(\pi_0\theta_H + (1-\pi_0)\theta_L)] + \frac{\partial E_{g_0}[v_o(\pi_1(a_0, g_0))]}{\partial a_0} = c'(a_0) \tag{3.6}$$

The first term in the square brackets on the left-hand side is the first-period (myopic) gain from a small increase in effort. The second term on the left-hand side is the marginal experimentation benefit or cost from changing a_0 from its myopic level $a^*(\pi_0)$. Let the value of a_0 that solves (3.6) be a_0^A.

The question is now: what sign is the marginal experimentation term? Following the proof of Lemma 2 of Mirman, Samuelson and Urbano (1993) it is possible to show (derivation in Appendix B) that

$$\frac{\partial E_{g_0}[v_o(\pi_1(a_0, g_0))]}{\partial a_0} = \pi_0(1-\mu)(\theta_H - \theta_L)\int_{-\infty}^{+\infty} v_o''(1-\pi_1)\frac{d\pi_1}{dg_0} f_H(g_0, a_0)\, dg_0 \tag{3.7}$$

Now, from (3.1), $v_o'' > 0$ as long as $\mu < 1$, and

$$\frac{d\pi_1}{dg_0} = \frac{\pi_0(1-\pi_0)}{[\pi_0 f_H + (1-\pi_0) f_L]^2}(f_L f_H' - f_L' f_H) > 0 \tag{3.8}$$

from the MLRC. So, we see that

$$\frac{\partial E_{g_0}[v_o(\pi_1(a_0, g_0))]}{\partial a_0} > 0 \text{ iff } \mu < 1$$

[14] See, for instance, Proposition 1 of Mirman, Samuelson and Urbano (1993).

i.e. that the experimentation term is strictly positive iff the technology is partly multiplicative. So, the following result is immediate from the previous discussion and the strict concavity of c:

Proposition 1. *In the second period, the appointee chooses the myopic level of effort $a^*(\pi_1)$, conditional on her posterior belief. In the first period, the appointee will choose to experiment by choosing a higher effort than the myopic one, $a_0^A > a^*(\pi_0)$, unless the technology is purely additive ($\mu = 1$), in which case $a_0^A = a^*(\pi_0)$.*

3.2. Democracy

This case is more complex, as we have a game of incomplete information, where there is both experimentation (unless the technology is additive) and a career concerns effect. We characterise the perfect Bayesian equilibria (PBE) of this game, which turn out to be unique[15] except that (possibly) the incumbent may choose multiple actions in period 0. Suppose first that the challenger to the incumbent, $j \in N$, is elected. His choice of action is $a_{j,1} = a^*(\pi_0)$, because he has no additional information about his own competence. So, the expected utility to any member $i \neq j$ of the electorate from the opponent is $v_c(\pi_0)$.

Now, at the time the electorate votes, every citizen has had the chance to observe g_0, first-period public good provision. Let π_1 be the updated belief on the part of the electorate, having observed g_0, that the incumbent is a high-type. Now, when forming the posterior π_1, citizens rationally deduce that in the first period, the incumbent has taken equilibrium action a_0^*. So, their posterior probability that the incumbent is competent is

$$\pi_1^c(g_0) = \frac{\pi_0}{\pi_0 + (1 - \pi_0)\left[f_L(g_0; a_0^*)/f_H(g_0; a_0^*)\right]} \tag{3.9}$$

Note that we superscript $\pi_1^c(g_0)$ to distinguish it from the incumbent's own posterior, which is defined in (3.2). However, note that in equilibrium, $\pi_1^c(g_0) = \pi_1(g_0; a_0^*)$.

Then the expected utility that citizens can expect from the incumbent is $v_c(\pi_1^c(g_0))$. So, given the tie-breaking rule, all the citizens (apart possibly from the opponent), will vote for the incumbent when $v_c(\pi_1^c(g_0)) \geq v_c(\pi_0)$. As v_c is strictly increasing in its argument, this is equivalent to $\pi_1^c(g_0) \geq \pi_0$. From (3.8), (3.9), $\pi_1^c(g_0)$ is strictly increasing in g_0. Moreover, from this fact and assumption A0, there exists a unique critical value $\bar{g}_0$ such that $\pi_1^c(\bar{g}_0) = \pi_0$, with $\pi_1 > \pi_0$ for $g_0 > \bar{g}_0$, and $\pi_1 < \pi_0$ for $g_0 < \bar{g}_0$. The conclusion is that all voters (except the incumbent) follow the following cutoff rule: vote for the incumbent iff $g_0 \geq \bar{g}_0$, and for the opponent if $g_0 < \bar{g}_0$. As there are at least three voters by assumption, this cutoff rule determines the outcome of the election, i.e. how the incumbent votes is irrelevant.

[15] Sufficient conditions for uniqueness are derived below.

It remains to check that it is individually rational for both the incumbent and opponent to stand for election, given this cutoff rule. The net gain to winning the election for the incumbent is

$$\phi(\pi_1^c(g_0)) = v_o(\pi_1^c(g_0)) - v_c(\pi_0) \tag{3.10}$$

Now, the individual rationality condition requires that $\phi(\pi_1^c(g_0)) \geq 0;\ \pi_1^c(g_0) \geq \pi_0$. But from (3.10), $\phi'(\pi_1^c(g_0)) = v_o'(\pi_1^c(g_0)) > 0$ from Section 2.5. So, we only need that $\phi(\pi_0) \geq 0$. But by definition, $\phi(\pi_0) = R - c(a^*(\pi_0))$; So, we will assume:[16]

A1. $R > c(a^*(\pi_0))$

This simply says that the "net" ego-rent from holding office is positive given prior π_0. Given A1, a similar argument implies that the opponent also wishes to hold office.

So, in view of the preceding discussion, we can write the second-period equilibrium continuation payoff of the incumbent conditional on $g_0; a_0$ as:

$$w(g_0;a_0) = \begin{cases} v_o(\pi_1(g_0;a_0)), & \text{if } g_0 \geq \hat{g}_0 \\ v_c(\pi_0), & \text{if } g_0 < \hat{g}_0 \end{cases} \tag{3.11}$$

So, the expected second-period continuation payoff of the incumbent, conditional on first-period effort only, is

$$\begin{aligned} E_{g_0}\left[w\left(\pi_1\left(g_0;a_0\right)\right)\right] &= v_c(\pi_0)H\left(\hat{g}_0;a_0\right) \\ &+ \int_{\hat{g}_0}^{+1} v_o(\pi_1(g_0;a_0))h\left(g_0;a_0\right)dg_0 \end{aligned} \tag{3.12}$$

where $h\left(g_0;a_0\right) = \pi_0 f_H + (1-\pi_0)\, f_L$ is the density of H from (3.4).

Now consider the choice of first-period action for the incumbent, given his continuation payoff (3.12). This must solve:

$$u_0 = \max_{a_0} E_{g_0} \left\{ \begin{array}{c} \pi_0[\mu(\theta_H + a_0) + (1-\mu)\theta_H a_0] + (1-\pi_0)\left[\mu(\theta_L + a_0) + (1-\mu)\theta_L a_0\right] \\ -c\left(a_0\right) + R + E_{g_0}[w(\pi_1(a_0;g_0))] \end{array} \right\} \tag{3.13}$$

The first-order condition can be written as

$$\mu + (1-\mu)(\pi_0\theta_H + (1-\pi_0)\,\theta_L) + \frac{\partial E_{g_0}[w(\pi_1(a_0;g_0))]}{\partial a_0} = c'\left(a_0\right) \tag{3.14}$$

After some manipulation, the third term on the left-hand side, evaluated at a_0; is given by (derivation in Appendix B)

$$\left.\frac{\partial E_{g_0}[w(\pi_1(a_0;g_0))]}{\partial a_0}\right|_{a_0^*} = \pi_0\left(1-\mu\right)\left(\theta_H - \theta_L\right)\int_{\hat{g}_0}^{+1} v_0''\frac{d\pi_1}{dg_0}\left(1-\pi_1\right) f_H\left(g_0;a_0\right)dg_0$$

[16] The strict inequality in A1 rules out several troublesome borderline cases in the model of Section 4 below with endogenous entry.

$$+\pi_0(1-\pi_0)(1-\mu)v_0'(\pi_0)f_H(\tilde{g}_0;a_0) \qquad (3.15)$$
$$+[R-c(a^*(\pi_0))]\left(-\frac{\partial H(\tilde{g}_0;a_0)}{\partial a_0}\right)$$

where

$$-\frac{\partial H(\tilde{g}_0;a_0)}{\partial a_0} = \pi_0[\mu+(1-\mu)\theta_H]f_H(\tilde{g}_0;a_0^*) + (1-\pi_0)[\mu+(1-\mu)\theta_L]f_L(\tilde{g}_0;a_0^*) > 0 \qquad (3.16)$$

The first term on the right-hand side represents the "experimentation" effect that we encountered in the appointment case. However, in this case it is clear by inspection that this term is unambiguously smaller than in the appointment case. The intuition is that the democratically elected office-holder only reaps the benefits of experimentation in the event that she is re-elected, which occurs with probability less than one. The second term $\pi_0(1-\pi_0)(1-\mu)v_0'(\pi_0)f_H(\tilde{g}_0)$, which is positive, is an additional incentive to experiment.

More importantly, the last term in (3.15) is a new effect which we call the "career concerns" effect, and is the product of two terms. The first, $R-c(a^*(\pi_0))$ is the net gain, or "prize" to winning the election[17] when $\pi_1=\pi_0$. The second term, $-\frac{\partial H}{\partial a_0}$, is the increased probability of winning the "prize" due to a small increment in effort. So, this last term in (3.15) represents the marginal extra effort that the incumbent office-holder is willing to supply in order to win the election. Note that the last term is always strictly positive by A1.

Let any level of action that solves (3.14) be denoted a_0^D. As the career concerns effect is always positive, then $a_0^D > a^*(\pi_0)$. Then we can summarise:

Proposition 2. In the second period, the elected official chooses the myopic level of effort $a^*(\pi_1)$, conditional on her posterior belief. In the first period, the official will choose a higher effort than the myopic one, $a_0^D > a^*(\pi_0)$, even if the technology is purely additive ($\mu = 1$).

Because this is an equilibrium action in a game, we cannot be sure that it is unique. Indeed in their analysis of career concerns in the labour market for bureaucrats, Dewatripont, Jewitt and Tirole showed that in the Normal-quadratic version of the model (ε Normally distributed, c quadratic) if the technology is sufficiently multiplicative, there are multiple (two) equilibrium action levels, but if the technology is additive, the equilibrium action is unique.

[17] This can be related to the tournament literature (Lazear and Rosen, 1981). There, the motivation for effort is to gain the first prize instead of the second prize. Here, the first prize for the incumbent is taking office (with payoff $v_o(\pi_0)$) and second prize is losing the election in which case the opponent wins, giving the incumbent $v_c(\pi_0)$. Of course, $v_o(\pi_0) - v_c(\pi_0) = R - c(a^*(\pi_0))$. Therefore, as in the tournament literature, a policy maker's effort depends on the spread between winning and losing prizes.

In our model, in the additive case, from (3.15), we get

$$1+\left[\pi_0 f_H\left(\tilde{g}_0;a_0^D\right)+(1-\pi_0)\,f_L\left(\tilde{g}_0;a_0^D\right)\right](R-c(a^*))=c'(a_0^D) \tag{3.17}$$

where a^* is the myopic optimal action in period 1 (independent of π_1). So, as $c''>0$, and $c'(0)<1$, a sufficient condition for uniqueness is that left-hand side of (3.17), viewed as a function of a_0, is decreasing for all $a_0 \le a_0^D$. But for this, it is sufficient that $f_H'(\tilde{g}_0;a_0),\ f_L'\left(\tilde{g}_0;a_0^D\right)\ge 0,\ a_0\le a_0^D$, or, more explicitly

$$f'(x)\ge 0,\ x\le \tilde{g}_0-\theta_L-a_0^D \tag{3.18}$$

This condition will be useful in what follows. We are also able to show that in the Normal-quadratic case, if the technology is additive, the equilibrium action is unique [see Appendix C].

Moreover, simulations reported in Appendix C, show that for a range of parameter values, the equilibrium action is unique even when technology is almost completely multiplicative ($\mu\simeq 0$). So, when comparing institutions in Section 3.3, we will assume that a_0^D is unique.

3.3. Comparing Institutions

We can now turn to the main topic of the paper, the comparison of effort levels under appointment and democracy. In the final period, conditional on posterior belief about type, the same (inefficiently low) effort level occurs under both institutions. The interesting comparison is therefore in the first period. Here, it is instructive to compare the incentive to raise the effort level above the myopic optimum in the democratic case and the appointment case. The difference between this incentive in the democratic and appointment cases is

$$\begin{aligned}\Delta \;=\;& [R-c(a^*(\pi_0))]\left(-\frac{\partial H(\tilde{g}_0;a_0)}{\partial a_0}\right) \\ &+\left[\begin{array}{c}\pi_0(1-\pi_0)(1-\mu)\,v_0'(\pi_0)\,f_H(\tilde{g}_0;a_0^*) \\ -\pi_0(1-\mu)(\theta_H-\theta_L)\int_{-\infty}^{\tilde{g}_0} v_0''\frac{d\pi_1}{dg_0}(1-\pi_1)\,f_H(g_0;a_0^*)\,dg_0\end{array}\right]\end{aligned} \tag{3.19}$$

Again assuming uniqueness of a_0^D, by the convexity of $c(.)$, $a_0^D>a_0^A$ iff $\Delta>0$.

Now, the first term in Δ is the "career concerns" term, and is positive. The second term in square brackets is the additional incentive for experimentation in the democratic case. Although it is not analytically possible to sign it in general, it is clear that when the technology is (approximately) linear, i.e. $\mu\simeq 1$, the second term is zero, and so $\Delta>0$ overall, implying $a_0^D>a_0^A$ the conventional result that elections motivate. Illustrative calculations in row 1 of Table 1 show that when the variance of ε is high, the career concerns effect on effort may be large.

Our main focus of interest is to establish conditions under which elections may demotivate. Inspection of (3.19) indicates that this is likely to occur when the net ego-rent from office, $R - c(a^*(\pi_0))$ is close to zero. In this case, there is (approximately) no "career concerns" effect under democracy, so that as long as there is more incentive to experiment with appointment, we will have $\Delta < 0$ and hence $a_0^D < a_0^A$. For the Normal-quadratic case, simulation results reported in column 1 of Table 1 below show that this can easily happen, and the demotivating effect of elections is larger, the more multiplicative the technology is. A natural way to measure this is in terms of the increase relative to the myopic level of effort induced by either arrangement. When $\rho = 0.75$, $(a_0^D - a^*(\pi_0))/(a_0^A - a^*(\pi_0)) \approx 1$, but when $\rho = 0$, $(a_0^D - a^*(\pi_0))/(a_0^A - a^*(\pi_0)) \approx 5/8$. Table 1 also shows that it is possible that $\Delta < 0$ when $\theta_H - \theta_L$ is sufficiently large. In this case, information about θ is valuable, so the appointee's incentive to experiment is strong, and is much diminished by an electoral constraint.

Table 1: EQUILIBRIUM EFFORT LEVELS a_0^A, a_0^D IN THE N-Q CASE

ρ \ $R - c(a^*(\pi_0))$:	≈ 0	50	100	$a^*(\pi_0)$
ρ:				
1.00	1.00; 1.00	1.00; 1.20	1.00; 1.40	1.00
0.75	4.05; 4.03	4.05; 4.81	4.05; 5.58	4.00
0.50	7.69; 7.41	7.69; 8.61	7.69; 9.71	7.00
0.25	12.08; 11.22	12.08; 12.19	12.08; 13.00	10.00
0.00	15.00; 14.25	15.00; 14.69	15.00; 15.07	13.00
$\theta_H - \theta_L$:				
0.5	1.1250; 1.1253	1.1250; 1.35	1.1250; 1.57	1.1250
10	3.51; 3.52	3.51; 4.21	3.51; 4.91	3.50
20	6.34; 6.21	6.34; 7.33	6.34; 8.39	6.00
50	15.05; 14.43	15.05; 14.75	15.05; 15.03	13.5

When ρ is variable, other parameters are: $\pi_0 = 0.5$, $\theta_H = 25$, $\theta_L = 1$, $\sigma = 100$.
When $\theta_H - \theta_L$ is variable, other parameters are: $\pi_0 = 0.5$, $\rho = 0.5$, $\sigma = 100$.

Next, consider the expected probability (taken with respect to $\theta_1, \ldots, \theta_n$) that the office-holder is a high-type, under any institutional arrangement. The expected probability that the office-holder in either period is high-type under appointment is π_0, as is the probability that the office-holder is high-type in the first period, with democracy. In the second period, the expected probability that the office-holder is a high-type is

$$\Pr(\theta^i = \theta_H | g_0 \geq \bar{g}_0)(1 - H(\bar{g}_0; a_0^*)) + \pi_0 H(\bar{g}_0; a_0^*) = \bar{\pi}_1 > \pi_0$$

So, we can summarise the discussion as follows:

Proposition 3. The same (myopic) level of effort is chosen in all cases in the second period. In the first period, $a_0^D > a_0^A$; if the technology is linear (in which case there is no motive for experimentation), but it is possible that $a_0^D < a_0^A$ if the "prize" for office $R - c(a^*(\pi_0))$ is approximately zero and $\theta_H - \theta_L$ is sufficiently large.

With appointment, a high-ability type is selected in both periods with expected probability $\pi_0 < 1$. With democracy, a high-ability type is selected in the first period with probability π_0 and in the second with probability $\bar{\pi}_1 > \pi_0$.

This raises the possibility that democracy need not be more "efficient" than appointment, as the former, while undoubtedly raising the average quality of the second-period office-holder, may lower effort (relative to appointment). This is investigated further in Section 5 below.

4. Democracy with Endogenous Entry

Here, we consider the following institution, which (as explained in the introduction) allows for both candidate entry and voting to be modelled in a complete way without any ad hoc assumptions. There is an election at the beginning of each of the two time periods. The first stage of an election process is candidate entry. Any citizen can stand for election in either of the two periods, at a cost of $\delta > 0$. We restrict citizens to pure-strategy entry decisions.[18] The second stage is plurality voting over the set of candidates. That is, every citizen has one vote which he must cast for one of the candidates, (we rule out abstentions), and the candidate with the most votes wins.[19] We impose the restriction that voters vote sincerely, i.e. for their most favored candidate. The justification for this, and the consequences of relaxing it, are discussed in Section 6.2 below.

In the event of a tie (i.e. two or more candidates with equal numbers of votes) we adopt the standard tie-breaking rule that every candidate with the most votes is chosen with equal probability. In the event that nobody stands for election, a default option is selected by the constitution, which is that no public good is provided.

Finally, it is very convenient both for the statement and proof of our result (but not for the main idea) to ensure that voters $k \neq i, j$ cannot be indifferent between candidates i, j, even if he believes them to have the same (expected) competence. This is most easily done by assuming the following lexicographically secondary preference; If office-holders i, j both supply amount g_t, k strictly prefers[20] (g_t, j) to (g_t, i) iff $j > i$. The purpose of introducing these "looks" preferences is to break ties in preferences over candidates that

[18] Existence of equilibrium in this model is not a problem, and so we do not need to consider the extension to mixed strategies.

[19] Given the lexicographic preferences assumed (see below)) a voter is never indifferent between two or more candidates.

[20] This index could refer to any visible variable that belongs to a citizen but is unrelated to her economic performance. For instance, her "look" as in Rogoff and Sibert (1988) and Rogoff (1990). That beauty is an important determinant of a person's performance in the labour market is shown in Hamermesh and Biddle (1994).

lead to multiple voting equilibria.

Again, here we are interested in locating the PBE of this model. Here, as entry is endogenous, voters' off-the-equilibrium path beliefs about the types of citizens who do not enter in equilibrium are important. We will assume that at all information sets where k has entered, all $j \neq k$ believe that k is high-ability with probability π_0, except where k is the incumbent (i.e. was elected in the previous period). Given that every citizen who is not the incumbent believes himself to be high-ability with probability π_0; these seem the only reasonable off-the-equilibrium path beliefs.

Next, let a_0^{DEN} be the solution to (3.14) above, but where R is replaced by $R - \delta$. Again, we will assume a_0^{DEN} is unique, which as argued above, is a weak restriction in the Normal-quadratic case. The interpretation of a_0^{DEN} is that it is the first-period effort chosen by an incumbent with endogenous entry. Then it is clear that $a_0^{DEN} < a_0^{D}$, as in the endogenous entry case, the "prize" for winning the election is reduced by the amount of the entry cost.

To achieve a characterization of the PBE of this model, we need the following assumption, which ensures that candidate entry costs are low enough so that some agent will stand for election, and high enough to deter all agents from standing for election:

A2. $\frac{1}{n} v_o(\pi_0) < \delta < v_o(\pi_0)$

We then have the following result:

Proposition 4. Assume A0-A2. Then, there is a unique PBE with the following structure. In period 0, only $i = n$ stands for election and is elected. He chooses action a_0^{DEN}. In period 1, if $g_0 \geq \hat{g}_0$, only $i = n$ stands for election and is elected. He chooses action $a^*(\pi_1(g_0; a_0^*))$. If $g_0 < \hat{g}_0$, only $i = n - 1$ stands for election and is elected. He chooses action is $a^*(\pi_0)$.

The intuition for the entry decisions in equilibrium is that given that citizens do not know their own types, no citizen runs for election on the basis of her superior ability in the first period. Only the citizen with the best "look" stands for election and is elected: non-economic variables decide which citizen becomes candidate and office-holder in the first period. In the second period, the incumbent is re-elected if his track record is sufficiently good, and anticipating this, he stands. On the other hand, if his track record is weak, he does not bother to stand (rationally anticipating defeat if he does), thus allowing the remaining citizen with the best "look" to stand and win.

Note, however, that once in office in the first period, the incumbent's choice of effort is exactly the same (modulo the fact that δ reduces the ego-rent) as in the baseline case of democracy. So, our results are robust to the introduction of endogenous candidate entry.[21]

5. Normative Analysis

In this section, we address the question of whether the equilibrium outcome under our main institution of interest, democracy, is Pareto-efficient relative to some benchmark. So, we are following Wittman (1989), Besley and Coate (1998), in studying efficiency of democracy in the Pareto sense, rather than relative to some arbitrary social welfare function (e.g. Benthamite) for a social planner.[22]

An outcome here is defined as (i) a choice of office-holder in each period; (ii) a level of action by the office-holder in each period, conditional on his information about his type. With democracy, outcomes (i) and (ii) are described by Propositions 1, 2, and 4.

One widely used benchmark is what could be achieved by a social planner with complete information (i.e. knowing $\theta_1, \ldots, \theta_n$, and able to choose action levels) who can choose the identity and effort of the office-holder, and a full set of economic instruments (i.e. can make unrestricted transfers of some numéraire good between citizens).

Say that democracy with endogenous entry is unconstrained efficient if the social planner of this type (the unconstrained social planner) cannot choose a feasible outcome that makes every citizen better off. Assume for convenience that citizen utilities are linear in the numéraire good. As the social planner can make unrestricted transfers between agents, democracy (with or without endogenous entry) is unconstrained efficient if and only if it selects the same conditional actions in each period, and the same choice of office-holder, as does the social planner.

It is then clear that democracy cannot be unconstrained efficient. First, clearly, the social planner will always select a high-type office-holder; if the office holder i is a low type, all voters, except i are better off if a high-type is made office holder, and the gainers can clearly compensate i. Also, as there are n citizens, each of whom gets utility g from a level of the public good g, the social planner will choose a to maximize the expected value of ng minus $c(a)$; conditional on a high-type being in office, i.e. it solves

$$n(\mu + (1 - \mu)\theta_H) = c'(a) \tag{5.1}$$

Let the solution to (5.1) be a^{**}. Comparing these outcomes to the equilibrium ones in Propositions 1, 2, and 4, it is clear that equilibrium outcomes with democracy are never unconstrained efficient.

[21] As shown in a companion paper (Le Borgne and Lockwood, 2000) this need not be the case under an asymmetric information structure - as is commonly assumed in the literature.

[22] All citizens have identical preferences over outcomes in any period (a choice of office-holder and an effort level for this office-holder). However, as the "good" of office is indivisible, any outcome must be horizontally inequitable, and so the social planner faces the problem of preference aggregation.

The weakness of the unconstrained efficiency benchmark is of course that the social planner is given superior information and more economic instruments than the office-holder. Consider now a constrained social planner who has the same information as the citizens (i.e. only knows the distribution of θ initially), and has the same powers as citizens, i.e. can "fire" the incumbent if performance falls below some cutoff value (i.e. no ability to redistribute the numéraire good). Say that democracy (with or without endogenous entry) is constrained efficient if this social planner cannot choose a feasible outcome that makes every citizen better off. Constrained efficiency is a much weaker test for any institution.

It is easy to see that the only feasible actions for the constrained social planner are; (i) random selection of an office-holder in the first period; (ii) replacement of the initial office-holder by another citizen selected at random if the only publicly observable indicator of the incumbent's performance, g_0, falls into some "unacceptable" set U. From the assumption of the MLRC, the social planner can do no better than to set $U = \{g_0 | g_0 < g_0^*\}$, i.e. follow a cutoff rule. Obviously, if $g^* = -\infty$, this is simply appointment, and if $g_0^* = \hat{g}_0$, democracy.

Nevertheless, in the presence of an experimentation motive ($\mu < 1$), democracy may not even be constrained efficient. Indeed, we can state:

Proposition 5. Assume A1, A2. Democracy (with or without endogenous entry) is constrained efficient if the technology is additive ($\mu = 1$) and in addition (i) the sufficient condition (3.18) for uniqueness of a_0^D holds; (ii) $R > c(a^*) + \pi_0(\theta_H - \theta_L)$. However, with (partly) multiplicative technology ($\mu < 1$), there are parameter values for which "appointment", i.e. $g_0^* = -\infty$, may Pareto-dominate democracy, in which case democracy is not even constrained efficient.

The key idea is that with a linear technology, there is never unanimity about changing g_0^* from $\hat{g}_0$; the initial office-holder will always prefer $g_0^* = -\infty$, effectively making him an appointee, but all citizens who never hold office always prefer (ex ante) a g_0^* higher than $\hat{g}_0$, in order to motivate the initial office-holder to supply more effort. This argument breaks down when the technology becomes multiplicative, as then (due to the experimentation effect) the initial office-holder may be motivated[23] by lowering the cutoff g_0^*, as then he captures more of the gains from experimenting. So then, everybody may gain from a lowering of g_0^*.

Note finally that the condition $R > c(a^*) + \pi_0(\theta_H - \theta_L)$ is a strengthening of A1; in the linear case, A1 is of course $R > c(a^*)$.

[23] This also requires that the career concerns effect will be small, i.e. that the "prize" for winning the election ($R - c(a^*(\pi_0))$) is approximately zero.

6. Extensions

6.1. O¢ce-Holder Altruism ($r > 0$)

The assumptions of the model generate a very strong form of underprovision of e¤ort; as e¤ort is non-contractible, the o¢ce-holder only has 1=n of the correct incentive to provide e¤ort. Consequently, (at least for large n); the higher equilibrium e¤ort, the more e¢cient the e¤ort is. This strong result can be re…ned by the (admittedly, ad hoc) device of supposing that the position of o¢ce has some psychological impact on the o¢ce-holder, making him or her more altruistic.[24] If $r > 0$, the positive analysis of the paper is qualitatively unchanged, except that the total ego-rent from o¢ce is now R + r(n ¡ 1)g; i.e. the ego-rent depends on performance while in o¢ce.

6.2. Strategic Voting

Our analysis has assumed that voters vote sincerely (i.e. for their most preferred candidate) at each election, no matter what the candidate set is. However, it is well-known[25] that when there are three or more candidates, voting sincerely might not be the only Nash equilibrium strategy (see Besley and Coate, 1997; or Dhillon and Lockwood, 2000). For example, in our model, it is a Nash equilibrium for all voters to vote for the candidate with the lowest index (i.e. looks characteristic). This is because no single voter can change the outcome by deviating, and so it is a weak best response to vote this way. However, as looks are uncorrelated with competence, this would not change the equilibrium outcome described in Proposition 4 in any economically relevant way.

7. Related Literature and Conclusions

7.1. Related Literature

The papers[26] most closely related to this one are Ferejohn (1986), Austen-Smith and Banks (1989), Banks and Sundaram (1993, 1998), and Persson and Tabellini (2000). In all these models, there is a moral hazard problem between o¢ce-holder and voters, and periodic elections unambiguously induce incumbent o¢ce-holders to supply more e¤ort (or in the case of Persson and Tabellini (2000) extract less rent).

[24] Holmström and Milgrom (1991) have such a type of assumption in their multitask agency model: they assume that not all work is unpleasant for an agent so that even without explicit incentives, the agent will supply e¤ort on some tasks.

[25] There is also much empirical evidence that it occurs in single-seat elections by plurality rule (Cox, 1997).

[26] Barro (1973) was the …rst to explicitly model electoral control of politicians. However, in his model, the actions of o¢ce-holders were always observable, and so if o¢ce-holders are in…nitely lived, they can always be induced to take e¢cient actions, if discounting is su¢ciently low (by a simple folk theorem argument).

In a classic article, Ferejohn (1986) proposed a simple and elegant moral hazard model of electoral control of office-holders. In equilibrium, voters follow a cutoff rule by voting for the incumbent only if his observed performance does not fall below a certain level, and the candidate chooses effort so that performance remains just at the cutoff. So, office-holder effort is higher than it would be without elections (there is electoral control of the incumbent).

As Ferejohn himself recognized (see p10 of his paper) his analysis relies[27] on the assumption that official may stay in office for ever (no term limits). With term limits, incumbents can never be induced to supply more than their myopic level of effort in the final period, and an "unraveling" argument then shows that incumbents can then never be induced to supply more than their myopic level of effort in any period of office.[28]

More recently, Banks and Sundaram (1998) have shown that with finite term limits, there can be electoral control of the incumbent if there is also an adverse selection ingredient to the model, namely, some ability parameter of the potential office-holder that is initially unobservable to the electorate. In this case, it is no longer ex post optimal to "fire" the incumbent in his last term of office if he has revealed himself to be of high enough quality. Indeed, under some very weak regularity conditions, the threat of (electoral)[29] dismissal induces agents of all types to supply more effort than they would otherwise in their first term of office[30] (Proposition 3.3).

Persson and Tabellini (2000, Chapter 4.5), have a two-period model with both adverse selection and moral hazard, where, as in this paper, initially the incumbent does not know his type.[31] Given an incumbent with competence θ; the technology for supplying the public good is

$$g_t = \theta(\tau - r_t) \tag{7.1}$$

where g_t is output of the public good, τ is exogenous tax revenue, and r_t are rents misappropriated from tax revenues. So, incumbents transform tax revenues net of rents into public goods. Voters care only about the level of public good provision, and the office-holder in period t has payoff $R + r_t$, where R is an ego-rent, as in our model.

Although Persson and Tabellini model rents in monetary terms, one (formally very similar) way of interpreting rent is to assume that it is the degree to which the official

[27] With term limits, Ferejohn's model can only exhibit electoral control in equilibrium if voters can precommit to a cutoff rule, a rather unattractive assumption.

[28] For a formal statement of this result, see Banks and Sundaram (1998), Proposition 3.5.

[29] Banks and Sundaram have a general model where the principal can only control the agent by dismissing him. This has an electoral interpretation, amongst others.

[30] See Besley and Case (1995) for an empirical test of the effects of term limits on the behaviour of US State governors.

[31] Biglaiser and Mezzetti (1997) have a paper where in the first period, the incumbent chooses an observable discrete project, but where the value of the project depends on the incumbent's ability (initially unknown to everybody) and a random shock. The paper focuses on the issue of whether undertaking the project is a good or bad signal to the electorate about the incumbent's ability.

"slacks" from the first-best level of effort defined in (5.1), i.e. $r = a^{**} - a$. In that case, we can write our production function, assuming $\mu = 1$, as

$$g_t = \theta(a^{**} - r_t) + \varepsilon_t$$

which is of course formally identical to (7.1) except that we now have a random productivity shock.

Also, note that the payoffs to the office-holder in our model can be written $R + g_t - c(a^{**} - r_t)$. So, the payoffs in Persson and Tabellini correspond to the special case where c is linear and the incumbent does not care about the public good.[32] To conclude, the Persson and Tabellini career concerns model can be thought of as a "special case" of ours,[33] and moreover, one in which the experimentation effect is ruled out by construction.

Of course, the merit of their model is that it is very simple and easily analyzed, and so it very well-suited to an analysis of the way career concerns are affected by electoral rules (Persson and Tabellini (2000), chapter 9.1). This would be much more difficult with a model such as ours.

7.2. Conclusions

Under symmetric incomplete information, an important insight from our paper is that career concerns and experimentation, while both inducing the incumbent to increase effort, are substitutes: that is, democracy introduces career concerns, but also necessarily reduces the incentive to experiment leading to short-termism in office. This substitutability is not present in other career concerns models because of simplifying assumptions which prevent experimentation from occurring (e.g. static model or additive technology).

In our electoral model, a corollary of this substitutability is that (conditional on ability) first-period effort may be higher or lower with democracy than with appointment. Our result is however more general and applies to other labour markets: as long as the agent has some positive probability of being "fired" by the principal, that the model is dynamic and the technology the agent uses is at least partly multiplicative in talent and effort then both career concerns and experimentation will be present. The general message is that the selection and retention process of an agent are important elements of job design in agency relationships.

8.

[32] This last fact creates the modelling problem that in the final period, the incumbent will supply no effort, i.e. extract maximum rent, whatever his type, implying that voters do not care about the types of the elected officials. Persson and Tabellini deal with this in a relatively ad hoc way by imposing an upper bound on the amount of rent that can be extracted.

[33] Mathematically, it is not literally a special case, as in their model, θ is continuously distributed.

References

Austen-Smith, D. and J.S. Banks (1989), "Electoral Accountability and Incumbency ", in Models of Strategic Choice in Politics, (P. Ordershook, Ed.) University of Michigan Press, Ann Arbor.

Banks, J.S. and R.K. Sundaram (1998), "Optimal Retention in Agency Problems", Journal of Economic Theory, 82, 293-323.

Banks, J.S. and R.K. Sundaram (1993), "Adverse Selection and Moral hazard in a Repeated Elections Model", in Political Economy: Institutions, Information Competition, and Representation, (W. Barnett, et al., Eds.), Cambridge University Press, Cambridge UK.

Barro, R. (1973), "The Control of Politicians: An Economic Model ", Public Choice, 14, 19-42.

Besley, T. and A. Case (1995), "Does Electoral Accountability A¤ect Economic Policy Choices? Evidence From Gubernatorial Term Limits ", Quarterly Journal of Economics, 110, 769-98.

Besley, T. and S. Coate (1998), "Sources of Ine¢ciency in a Representative Democracy: A Dynamic Analysis", American Economic Review, 88, 139-56.

Besley, T. and S. Coate (1997), "An Economic Model of Representative Democracy ", Quarterly Journal of Economics, 112, 85-114.

Biglaiser, G. and C. Mezetti (1997), "Politicians' Decision Making With Re-election Concerns", Journal of Public Economics, 66, 425-47.

Cox, G.W. (1997), Making Votes Count, Cambridge University Press, Cambridge.

Dewatripont, M., I. Jewitt, and J. Tirole (1999), "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies", Review of Economic Studies, 66, 199-217.

Dhillon, A., and B. Lockwood (2000), "When are Plurality Rule Voting Games Dominance-Solvable?", mimeo, Department of Economics, University of Warwick.

Ferejohn, J. (1986), "Incumbent Performance and Electoral Control" Public Choice, 50, 5-26.

Gibbons, R., and M. Waldman (1999), "Careers in Organizations: Theory and Evidence", in Handbook of Labor Economics (O. Ashenfelter and D. Card, Eds.), ch. 36, 2373-437.

Green, J., and N. Stokey (1983), "Tournaments and Contracts", Journal of Political Economy, 91, 349-64.

Grossman, S.J., R.E. Kihlstrom, and L.J. Mirman (1977), "A Bayesian Approach to the Production of Information and Learning by Doing", Review of Economic Studies, 44, 533-47.

Hamermesh, D.S., and J.E. Biddle (1994), "Beauty and the Labor Market", American Economic Review, 84, 1174-94.

Holmström, B. (1999), "Managerial Incentive Problems: A Dynamic Perspective", Review of Economic Studies, 66, 169-82. Originally (1982) in Essays in Economics and Management in Honor of Lars Wahlbeck (Helsinki: Swedish School of Economics).

Holmström, B., and P. Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", Journal of Law, Economic and Organizations, 7, 24-51.

Keller, G., and S. Rady (1999), "Optimal Experimentation in a Changing Environment", Review of Economic Studies, 66, 475-507.

Lazear, E.P., and S. Rosen (1981), "Rank-Order Tournaments as Optimum Labor Contracts", Journal of Political Economy, 89, 841-64.

Le Borgne, E., and B. Lockwood (2000), "Candidate Entry, Screening, and Political Business Cycles", mimeo, University of Warwick, Department of Economics.

Milgrom, P. (1981), "Good News and Bad News: Representation Theorems and Applications", Bell Journal of Economics, 12, 380-91.

Mirman, L.J, L. Samuelson, and A. Urbano (1993), "Monopoly Experimentation", International Economic Review, 34, 549-63.

Persson, T., and G. Tabellini (2000), Political Economics: Explaining Economic Policy, MIT Press, Cambridge, MA.

Prendergaste, C. (1999), "The Provision of Incentives in Firms", Journal of Economic Literature, 37, 7-63.

Prescott, E.C. (1972), "The Multiperiod Control Problem Under Uncertainty", Econometrica, 40, 1043-58.

Rogo¤, K. (1990), "Equilibrium Political Budget Cycles", American Economic Review, 80, 21-36.

Rogo¤, K., and A. Sibert (1988), "Elections and Macroeconomic Policy Cycles", Review of Economic Studies, 55, 1-16.

Wittman, D. (1989), "Why Democracies Produce E¢cient Results", Journal of Political Economy, 97, 1395-426.

A. Proofs of Propositions

Proof of Proposition 4. We show that the perfect Bayesian equilibrium described exists and is unique by backwards induction. First, it is clear that any $i \in N$ who is elected at period 1 chooses $a^*(\pi_0)$ if he was not a first-period office-holder, and chooses $a^*(\pi_1^c(g_0))$ if he was a first-period office-holder.

Next, consider the behaviour of the voters, given any candidate set C_1. The first case is where the incumbent (say i) is not in C_1. Then, the voters in $N \backslash C_1$ will vote for their most preferred candidate in C_1. By our assumption about beliefs, voters believe that any member of C_1 is a high-type with probability π_0. So, they prefer the one with the highest index, $m(C_1) = \max_{i \in C_1}$. Finally, by A1, every candidate will vote for herself.

The second case is where the incumbent (say i) is in C_1. Then, all voters know that i is high-ability with probability $\pi_1^c(g_0)$, and believe that any $j \in C_1 \backslash \{i\}$ is high-ability with probability π_0. So, if $g_0 > \bar{g}_0$, all voters in $N \backslash C_1$ will vote for i. Also, i will vote for herself by A1. The remaining voters, i.e. $C \backslash \{i\}$ will either vote for themselves or i, as $v_o(\pi_0)$ is greater or less than $v_c(\pi_1^c(g_0))$. If $g_0 < \bar{g}_0$, all voters in $N \backslash C_1$ will vote for $m'(C_1) = \max_{j \in C_1 \backslash \{i\}}$. Also, i will vote for herself or $m'(C_1)$ as $v_o(\pi_1^c(g_0))$ is greater or less than $v_c(\pi_0)$. The remaining voters, i.e. $C_1 \backslash \{i\}$ will vote for themselves by A1. If $g_0 = \bar{g}_0$, all voters in $N \backslash C_1$ will vote for $m(C_1) = \max_{j \in C_1}$. The remaining voters, i.e. C_1 will vote for themselves by A1.

Now consider the candidate entry decision in period 1.

Case 1. $C_1 \neq N$. If $g_0 > \bar{g}_0$, and the incumbent i enters the election, she will surely win, no matter who else stands. So, the incumbent will be the only entrant. Now let

$$l = \begin{cases} n & \text{if } i < n \\ n-1 & \text{if } i = n \end{cases} \tag{A1}$$

If $g_0 < \bar{g}_0$, if l enters the election, $l = m'(C_1)$, so she will surely win, no matter who else stands (including the incumbent). So, l will be the only entrant. Finally, if $g_0 = \bar{g}_0$, if n enters the election, $n = m(C_1)$, so she will surely win, no matter who else stands (including the incumbent). So, n will be the only entrant.

Case 2. $C_1 = N$. In this case, at the voting stage, every candidate votes for himself and is elected to be office-holder with probability $1/n$. The payoff to any agent who is not the incumbent from entry is thus is $\frac{1}{n} v_o(\pi_0) - \delta$, which is negative by A2. On the other hand, any $i \in N$ can guarantee herself a positive payoff by not entering. So, this case is impossible in equilibrium.

So we have demonstrated that given a first-period incumbent i, with output g_0, in the second period, the unique equilibrium candidate set is

$$C_1(i; g_0) = \begin{cases} \{i\}, & g_0 > \bar{g}_0 \\ \{l\}, & g_0 < \bar{g}_0 \\ \{n\}, & g_0 = \bar{g}_0 \end{cases} \tag{A2}$$

Now consider the first period. Clearly, if $i \in N$ is elected, she rationally anticipates that she will stand for election next period (and win) iff either (i) $i = n;\ g_0 \geq \hat{g}_0$; or (ii) $i < n;\ g_0 > \hat{g}_0$: In either case, given that ε is absolutely continuous, she chooses a_0 to solve problem (3.13) where R is replaced by $R - \delta$: Moving to the voting stage, by previous arguments, all voters in $N \backslash C_0$ will vote for $m(C_0) = \max_{i \in C_0}$; and all voters in C_0 will vote for themselves. So, again by previous arguments, $C_0 = \{n\}$: ■

Proof of Proposition 5. (a) We prove first that with additive technology, equilibrium with democracy is weakly efficient. To do this, it is sufficient to show that there does not exist a cutoff $g_0^* \neq \hat{g}_0$ where all citizens are better off than at the equilibrium cutoff.

Let γ be an arbitrary cutoff. Without loss of generality, we can assume that the social planner chooses citizen n to be the first-period office holder, and $n-1$ to replace him in the second period if his performance falls below γ: Let $a_0(\gamma)$ be the office-holder's first-period action given the cutoff. So, $a_0(\gamma)$ solves (3.17) with γ replacing $\hat{g}_0$: Totally differentiating (3.17), we get

$$\begin{aligned} a_0'(\gamma) &= \frac{A}{c''(a_0(\gamma)) + A}; \\ \text{where}\quad A &= [\pi_0 f_H'(\gamma; a_0(\gamma)) + (1-\pi_0) f_L'(\gamma; a_0(\gamma))](R - c(a^*)) \end{aligned} \tag{A.1}$$

Also, $c'' > 0$; and as (3.18) holds we have $A \geq 0;\ \gamma \leq \hat{g}_0$: So, from (A.1) we have

$$0 \leq a_0'(\gamma) < 1;\ \gamma \leq \hat{g}_0 \tag{A.2}$$

We can first write down expected present value payoff of $i = n$ conditional on this cutoff, given that the office-holder optimizes his actions in both periods;

$$\begin{aligned} v_n(\gamma) &= \bar{\theta} + a_0(\gamma) + R - c(a_0(\gamma)) + \int_{\gamma}^{\infty} v_c(\pi_1(a_0(\gamma); g_0))h(g_0; a_0(\gamma))dg_0 \\ &\quad + H(\gamma; a_0(\gamma))v_c(\pi_0) \end{aligned} \tag{A.3}$$

where $\bar{\theta} = \pi_0\theta_H + (1-\pi_0)\theta_L$: Note first that from (A.3) and the fact that $a_0(\gamma)$ maximises (3.17):

$$\begin{aligned} v_n'(\gamma) &= h(\gamma; a_0(\gamma))[v_c(\pi_0) - v_o(\pi_1(a_0(\gamma); \gamma))] \\ &< h(\gamma; a_0(\gamma))[v_c(\pi_0) - v_o(0)] \\ &< 0 \end{aligned} \tag{A.4}$$

where the second line follows from the properties of $v_o; v_c$ given in Section 2.5, and the third from the assumption that $R > c(a^*) + \pi_0(\theta_H - \theta_L)$; which is equivalent to $v_o(0) > v_c(\pi_0)$ when $\mu = 1$. So, from (A.4), n prefers the lowest possible $\gamma = -\infty$ (i.e. no election).

So, the social planner cannot make everybody better off by raising γ from $\hat{g}_0$. Thus, to prove that the equilibrium is weakly efficient, it suffices to prove that some $j \neq n$ most prefers a cutoff at or above $\hat{g}_0$. For then, the social planner cannot make everybody better off by lowering γ from $\hat{g}_0$, either. Note that for $j < n - 1$:

$$v_j(\gamma) = \bar{\mu} + a_0(\gamma) + \int_{\gamma}^{1} v_c(\pi_1(a_0(\gamma), g_0))h(g_0, a_0(\gamma))dg_0 + H(\gamma, a_0(\gamma))v_c(\pi_0) \qquad \text{(A.5)}$$

Now differentiating (A.5), we have;

$$v_j'(\gamma) = \frac{\partial v_j}{\partial a_0}a_0'(\gamma) + h(\gamma, a_0(\gamma))[v_c(\pi_0) - v_o(\pi_1(a_0(\gamma), \gamma))], \; j < n - 1 \qquad \text{(A.6)}$$

Now, note from (3.2) that with a linear technology, $\pi_1(a_0(\gamma), \gamma) \equiv \pi_1(\gamma - a_0(\gamma))$, with $\pi_1'(.) > 0$ by the MLRC. So, from this fact and the fact from (A.2) that $\gamma - a_0(\gamma)$ is increasing in γ, from (A.6), we have

$$v_o(\pi_1(a_0(\gamma), \gamma)) \leq v_o(\pi_1(a_0(\hat{g}_0), \hat{g}_0)), \; \gamma \leq \hat{g}_0 \qquad \text{(A.7)}$$

Also, by previous definitions and results:

$$\begin{aligned} v_c(\pi_0) &= v_o(\pi_1(a_0(\hat{g}_0), \hat{g}_0)) - (R - c(a^*)) \qquad \text{(A.8)} \\ &< v_o(\pi_1(a_0(\hat{g}_0), \hat{g}_0)) \\ &\leq v_o(\pi_1(a_0(\hat{g}_0), \hat{g}_0)), \; \gamma \leq \hat{g}_0 \end{aligned}$$

In the first line, we have used the definition of $\hat{g}_0$ that $\pi_0 = \pi_1(a_0(\hat{g}_0), \hat{g}_0)$, and the definitions of v_c, v_o. In the second, we have used $R > c(a^*)$ from A1 (note with linearity, the myopic action a^* does not depend on π_1). In the third, we have used (A.7). Therefore, from (A.6), (A.8), we see that

$$v_j'(\gamma) \geq \frac{\partial v_j}{\partial a_0}a_0'(\gamma), \; \gamma \leq \hat{g}_0, \; j < n - 1 \qquad \text{(A.9)}$$

Finally, it is obvious that $\partial v_j / \partial a_0 > 0$, as a_0 is chosen optimally by the office-holder, n, but $j \neq n$ does not bear the cost of the action. So, from this fact, (A.2) and (A.9), we conclude that $v_j'(\gamma) \geq 0, \; \gamma \leq \hat{g}_0$ so $j < n - 1$ most prefers a cutoff at least $\hat{g}_0$, as required.

(b) An example with non-additive technology where appointment Pareto-dominates democracy can be constructed as follows. W.l.o.g., assume that the incumbent is n and the challenger is $n - 1$. Equilibrium payoffs under democracy, allowing $\mu \neq 1$, are:

$$v_n^D = \mu(\bar{\mu} + a_0^D) + (1 - \mu)\bar{\mu}a_0^D + R - c(a_0^D) + \int_{\hat{g}_0}^{1} [v_o(\pi_1(a_0^D, g_0))h(a_0^D, g_0)dg_0 +$$

$$H(\hat g_0, a_0(\hat g_0))v_c(\pi_0)$$

$$v^D_{n-1} = \bar\theta + a_0(\hat g_0) + \int_{\hat g_0}^{\infty} v_c(\pi_1(a_0(\hat g_0), g_0))h(g_0, a_0(\hat g_0))dg_0 + H(\hat g_0, a_0(\hat g_0))v_o(\pi_0)$$

$$v^D_j = \bar\theta + a_0(\hat g_0) + \int_{\hat g_0}^{\infty} v_c(\pi_1(a_0(\hat g_0), g_0))h(g_0, a_0(\hat g_0))dg_0 + H(\hat g_0, a_0(\hat g_0))v_o(\pi_0), \; j < n-1$$

Also, under appointment of citizen i, the expected utilities are

$$v^A_i = \mu(\bar\theta + a^A_0) + (1-\mu)\bar\theta a^A_0 + R - c(a^A_0) + \int_{-\infty}^{\infty} v_o(\pi_1(a^A_0, g_0))h(a^A_0, g_0)dg_0$$

$$v^A_j = \mu(\bar\theta + a^A_0) + (1-\mu)\bar\theta a^A_0 + \int_{-\infty}^{\infty} v_o(\pi_1(a^A_0, g_0))h(a^A_0, g_0)dg_0, \; j \neq i$$

The example is the following. First, ε is Normal, with mean zero and $\sigma = 50$, and $c(a) = a^2/2$, and other parameters are: $\mu = 0.5$, $\pi_0 = 0.55$, $R = 39.9$; $\theta_H = 25$; $\theta_L = 1$; $\delta = 1$. In this case, equilibrium payoffs can be calculated using the above formulae as:

$$v^A_i = 156.3; \; v^A_j = 152.1; \; j \neq i$$

$$v^D_n = 148.4; \; v^D_{n-1} = 148.2; \; v^D_j = 143.6; \; j \neq n, n-1$$

So, we see that $\max_{i \in N} v^D_i < \min v^A_i$, and so we can be sure that appointment Pareto-dominates democracy with endogenous entry. $\square$

B. Derivations

Derivation of (3.7)

(Adapted from the proof of Proposition 2 of Mirman et al., 1993). Before turning to the derivation of equation (3.7) itself, the following results are useful. First:

$$\frac{d\pi_1(g_0, a_0)}{dg_0} = \frac{\pi_0(1-\pi_0)}{[\pi_0 f_H + (1-\pi_0) f_L]^2}\left(f_L f'_H - f_H f'_L\right) \geq 0 \qquad \text{(B1)}$$

where $f_H = f(g_0 - (1-\mu)\theta_H a_0 - \mu(\theta_H + a_0))$, and $f_L = f(g_0 - (1-\mu)\theta_L a_0 - \mu(\theta_L + a_0))$. $\left(f_L f'_H - f_H f'_L\right) \geq 0$ follows from the MLR property. Second:

$$\frac{d\pi_1(g_0, a_0)}{da_0} = -[\mu + (1-\mu)\theta_H]\frac{d\pi_1}{dg_0} - \frac{\pi_0(1-\pi_0)(1-\mu)(\theta_H - \theta_L)}{[\pi_0 f_H + (1-\pi_0) f_L]^2} f_H f'_L \qquad \text{(B2)}$$

$$= -[\mu + (1-\mu)\theta_L]\frac{d\pi_1}{dg_0} - \frac{\pi_0(1-\pi_0)(1-\mu)(\theta_H - \theta_L)}{[\pi_0 f_H + (1-\pi_0) f_L]^2} f_L f'_H \qquad \text{(B3)}$$

We can now evaluate (3.7). Notice that

$$E_{g_0} v_o\left[\pi_1(g_0;a_0)\right] = \int_{-\infty}^{+\infty} v_o\left[\pi_1(g_0;a_0)\right] h(g_0;a_0)\, dg_0$$

where $h(g_0;a_0) = \pi_0 f_H + (1-\pi_0) f_L$. Thus, denoting $dE_{g_0} v_o\left[\pi_1(g_0;a_0)\right]/da_0 = \Gamma$; we have:

$$\Gamma = \int v_o' \frac{d\pi_1}{da_0}\left[\pi_0 f_H + (1-\pi_0) f_L\right] dg_0 - \int v_o \left[\begin{array}{c} \pi_0 \left(\mu + (1-\mu)\theta_H\right) f_H' \\ +(1-\pi_0)\left(\mu + (1-\mu)\theta_H\right) f_L' \end{array}\right] dg_0 \quad \text{(B4)}$$

Integrating by parts the second term of (B4) and then rearranging with the first term gives

$$\begin{aligned} \Gamma = & \int v_o' \left(\frac{d\pi_1}{da_0} + \left(\mu + (1-\mu)\theta_H\right)\frac{d\pi_1}{dg_0}\right) \pi_0 f_H\, dg_0 \\ & + \int v_o' \left(\frac{d\pi_1}{da_0} + \left(\mu + (1-\mu)\theta_L\right)\frac{d\pi_1}{dg_0}\right) (1-\pi_0) f_L\, dg_0 \end{aligned} \quad \text{(B5)}$$

Using (B2) and (B3), expression (B5) becomes

$$\begin{aligned} \Gamma = & -\int v_o' \frac{\pi_0(1-\pi_0)(1-\mu)}{\left[\pi_0 f_H + (1-\pi_0) f_L\right]^2}\left[\theta^H - \theta^L\right] f_H f_L' \pi_0 f_H\, dg_0 \\ & -\int v_o' \frac{\pi_0(1-\pi_0)(1-\mu)}{\left[\pi_0 f_H + (1-\pi_0) f_L\right]^2}\left[\theta^H - \theta^L\right] f_H' f_L (1-\pi_0) f_L\, dg_0 \end{aligned} \quad \text{(B6)}$$

Because $\pi_1 = \pi_0 f_H / \left[\pi_0 f_H + (1-\pi_0) f_L\right]$ and $(1-\pi_1) = (1-\pi_0) f_L / \left[\pi_0 f_H + (1-\pi_0) f_L\right]$, equation (B6) becomes

$$\Gamma = -(\theta_H - \theta_L)\left\{\int v_o' \pi_1^2 (1-\pi_0)(1-\mu) f_L'\, dg_0 + \int v_o' (1-\pi_1)^2 \pi_0 (1-\mu) f_H'\, dg_0\right\} \quad \text{(B7)}$$

Using the fact that $(1-\pi_1)^2 = (1-\pi_1) - \pi_1(1-\pi_1)$, we can rewrite (B7) as

$$\Gamma = -(\theta_H - \theta_L)\left\{\begin{array}{c} \int v_o' (1-\mu)\pi_1\left[\pi_1(1-\pi_0) f_L' - (1-\pi_1)\pi_0 f_H'\right] dg_0 \\ + \int v_o' (1-\mu)(1-\pi_1)\pi_0 f_H'\, dg_0 \end{array}\right\} \quad \text{(B8)}$$

Rearranging the posterior belief π_1, we have $\pi_1\left[\pi_0 f_H + (1-\pi_0) f_L\right] = \pi_0 f_H$, which, after differentiating with respect to g_0 gives (after rearranging)

$$f_L' \pi_1 (1-\pi_0) - f_H' \pi_0 (1-\pi_1) = -\frac{d\pi_1}{dg_0}\left[\pi_0 f_H + (1-\pi_0) f_L\right] \quad \text{(B9)}$$

Inserting (B9) in (B8) yields,

$$\Gamma = (\theta_H - \theta_L)\left\{\begin{array}{c} \int v_o' (1-\mu)\pi_1 \frac{d\pi_1}{dg_0}\pi_0 f_H\, dg_0 + \int v_o' (1-\mu)\pi_1 \frac{d\pi_1}{dg_0}(1-\pi_0) f_L\, dg_0 \\ - \int v_o' (1-\mu)(1-\pi_1)\pi_0 f_H'\, dg_0 \end{array}\right\} \quad \text{(B10)}$$

From the π_1 expression, we have $f_L(1-\pi_0)\pi_1 = f_H\pi_0(1-\pi_1)$, so (B10) becomes

$$\Gamma = (\theta_H - \theta_L)\left\{\int v_o'(1-\mu)\frac{d\pi_1}{dg_0}\pi_0 f_H\, dg_0 - \int v_o'(1-\mu)(1-\pi_1)\pi_0 f_H'\, dg_0\right\} \quad \text{(B11)}$$

Now we integrate the second term in (B11) by parts. This yields,

$$\begin{aligned}(1-\mu)\pi_0\int v_o'(1-\pi_1)f_H'\, dg_0 &= -(1-\mu)\pi_0\int v_o''(1-\pi_1)\frac{d\pi_1}{dg_0}f_H\, dg_0 \\ &\quad + (1-\mu)\pi_0\int v_o'\frac{d\pi_1}{dg_0}f_H\, dg_0\end{aligned} \quad \text{(B12)}$$

Inserting (B12) in (B11) gives (3.7). $\blacksquare$

Derivation of Equation (3.15)

The derivation is similar to that of equation (3.7). First, note that $E_{g_0}[w(\pi_1(g_0;a_0))]$ can be written

$$\begin{aligned}E_{g_0}[w(\pi_1(g_0;a_0))] &= (v_c(\pi_0) - v_o(\pi_0))H(\underline{g}_0;a_0) + \int_{\underline{g}_0}^{+\infty}(v_o(\pi_1(g_0;a_0)) - v_o(\pi_0))h(g_0;a_0)\, dg_0 \\ &= [c(a^*(\pi_0)) - R]H(\underline{g}_0;a_0) + \int_{\underline{g}_0}^{+\infty}\phi((\pi_1(g_0;a_0))h(g_0;a_0)\, dg_0\end{aligned}$$

where $\phi(\pi_1(g_0;a_0)) = v_o(\pi_1(g_0;a_0)) - v_o(\pi_0)$; so $\phi(\pi_1(\underline{g}_0;a_0^*)) = 0$; $\phi' = v_o'$. So,

$$\begin{aligned}\Gamma \equiv \frac{dE_{g_0}[w^*(\pi_1(g_0;a_0))]}{da_0} &= (v_c(\pi_0) - v_o(\pi_0))\left(-\frac{\partial H(\underline{g}_0;a_0)}{\partial a_0}\right) \\ &\quad + \int_{\underline{g}_0}^{+\infty}\phi'\frac{d\pi_1}{da_0}[\pi_0 f_H + (1-\pi_0)f_L]\, dg_0 \\ &\quad - \int_{\underline{g}_0}^{+\infty}\phi\left[\pi_0(\mu + (1-\mu)\theta_H)f_H' + (1-\pi_0)(\mu + (1-\mu)\theta_L)f_L'\right]dg_0\end{aligned} \quad \text{(B13)}$$

Integrating by parts the third term of (B13) and then rearranging with the first two terms gives

$$\begin{aligned}\Gamma &= \int_{\underline{g}_0}^{+\infty}\phi'\left(\frac{d\pi_1}{da_0} + (\mu + (1-\mu)\theta_H)\frac{d\pi_1}{dg_0}\right)\pi_0 f_H\, dg_0 \\ &\quad + \int_{\underline{g}_0}^{+\infty}\phi'\left(\frac{d\pi_1}{da_0} + (\mu + (1-\mu)\theta_L)\frac{d\pi_1}{dg_0}\right)(1-\pi_0)f_L\, dg_0 \\ &\quad + [R - c(a^*(\pi_0))]\left(-\frac{\partial H(\underline{g}_0;a_0)}{\partial a_0}\right)\end{aligned} \quad \text{(B14)}$$

After using similar manipulations as for the derivation of equation (3.7), we obtain

$$
\begin{aligned}
- \;=\; & (\theta_H - \theta_L) \left\{ \int_{\tilde{g}_0}^{+\infty} \phi'(1-\mu) \frac{d\pi_1}{dg_0} \pi_0 f_H \, dg_0 - \int_{\tilde{g}_0}^{+\infty} \phi'(1-\mu)(1-\pi_1) \pi_0 f_H' \, dg_0 \right\} \\
& + [R - c(a^*(\pi_0))] \left(- \frac{\partial H(\tilde{g}_0; a_0)}{\partial a_0} \right) \qquad \text{(B15)}
\end{aligned}
$$

Now we integrate the second term in (B15) by parts. This yields,

$$
\begin{aligned}
(1-\mu)\pi_0 \int_{\tilde{g}_0}^{+\infty} \phi'(1-\pi_1) f_H' \, dg_0 \;=\; & -(1-\mu)\pi_0 \int_{\tilde{g}_0}^{+\infty} \phi'' \left(\frac{d\pi_1}{dg_0} \right) (1-\pi_1) f_H \, dg_0 \qquad \text{(B16)} \\
& + (1-\mu)\pi_0 \int_{\tilde{g}_0}^{+\infty} \phi' \frac{d\pi_1}{dg_0} f_H \, dg_0 \\
& - \frac{d\phi(\pi_1(\tilde{g}_0; a_0))}{d\pi_1} (1 - \pi_1(\tilde{g}_0; a_0)) \pi_0 (1-\mu) f_H(\tilde{g}_0)
\end{aligned}
$$

Inserting (B16) in (B15), and using $\phi' = v_0'$; and finally evaluating at $a_0 = a_0^*$ (and recalling $\pi_1(\tilde{g}_0; a_0) = \pi_0$) gives equation (3.15). $\square$

C. The Normal-Quadratic Case

The example we use follows the specification of Dewatripont, Jewitt and Tirole (1999). The cost of effort function is quadratic (specifically $c(a) = a^2/2 + da$; with $d \geq 0$; so that $c'(0) \geq 0$), and the error term ε is Normally distributed with mean zero and variance σ^2. We now analyse the different sections of the model under our specific assumptions.

We can now prove that when the technology is additive (i.e. $\mu = 1$), a unique equilibrium arises. In the appointment case, this is immediate as in this case there is no experimentation. For the democratic cases, when $\mu = 1$, it is possible to calculate (details on request) that

$$
\begin{aligned}
\left. \frac{\partial E_{g_0}[w(\pi_1)]}{\partial a_0} \right|_{\mu=1} \;=\; & \left[R - \delta - \frac{(a^*(\pi_0))^2}{2} - d a^*(\pi_0) \right] \\
& \times \left(\frac{\pi_0}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{8\sigma^2} (\theta_L - \theta_H)^2 \right] + \frac{(1-\pi_0)}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{8\sigma^2} (\theta_H - \theta_L)^2 \right] \right)
\end{aligned}
$$

which is decreasing in $a^*(\pi_0)$. On the other hand, the marginal cost of effort is upward sloping. Hence a unique equilibrium exists. The simulations reported in the paper are also based on this special case. Details are available on request.