INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Paying Politicians

Matthias Messner and Mattias Polborn

Working Paper n. 246

November 2003

IGIER – Università Bocconi, Via Salasco 5, 20136 Milano –Italy
http://www.igier.uni-bocconi.it

Paying politicians*

Matthias Messner
Bocconi University and IGIER

Mattias K. Polborn[†]
UWO and University of Illinois

September 2003

Abstract

We consider a society that has to elect an official who provides a public service for the citizens. Potential candidates differ in their competence and every potential candidate has private information about his opportunity cost to perform the task of the elected official. We develop a new citizen candidate model with a unique equilibrium to analyze citizens' candidature decisions.

Under some weak additional assumptions, bad candidates run with a higher probability than good ones, and for unattractive positions, good candidates free-ride on bad ones. We also analyze the comparative static effects of wage increases and cost of running on the potential candidates' entry decisions.

Keywords: Citizen-candidate model, political economy, private provision of public goods, wage for politicians.
JEL code: D7, H0.

*We would like to thank seminar audiences at the University of Toronto and the University of Western Ontario, as well as Antonio Cabrales, Jim Davies, Srihari Govindan, Ig Horstmann, Tilman Klumpp, Massimo Morelli, Mike Peters, Al Slivinski and Gopal Vasudev for helpful comments.

[†]Corresponding author: Mattias Polborn, Department of Economics, University of Western Ontario, London, Ontario, N6A 5C2, Canada; email: mpolborn@uwo.ca

1 Introduction

Potential candidates for political office will be influenced in their decision whether to enter the competition – as in any other profession – by financial considerations. The wage and/or other benefits of the office will have an effect on the set of candidates who are willing to run for office, and hence possibly on the person of the elected official and the policy outcome.[1]

However, most existing political economy models ignore this issue. They exogenously and sometimes implicitly assume that being a politician is an attractive job, because of perks associated with the political office (also known as ego rents or spoils of office). If elected, politicians in these models receive more than their reservation utility. While this may be true for some prestigious political positions, there are also many elected offices that are not particularly attractive for officials. An example of the latter type are positions in university committees; members of these committees usually provide a sometimes high labor input without receiving an explicit compensation for this task.

Moreover, how attractive a position is, is usually a decision of society through the choice of (explicit or implicit) remuneration, rather than an exogenous characteristic of the job. If every member of an appointments committee were paid \$ 10,000 for this work, there would certainly be many more candidates volunteering for the committee positions than now, and similar effects will be present for most political jobs that "nobody wants" today. However, for some reason, the respective society *chooses* not to pay such a high compensation for these positions.[2] Since the attraction of a political office is determined by a choice of society, it should ultimately be determined endogenously in politico-economic models. This paper makes a first step in this direction by providing a tractable model to analyze the effect of remuneration on strategic running decisions in a citizen candidate framework.

In our model, there are I potential candidates coming from a total of N citizens. Every potential candidate has a cost of being the office holder if he is elected; this cost is drawn from a known distribution, but the realization is a candidate's private information. Potential candidates differ in their competence only, that is in the value they can create in the office for every citizen. This value is common knowledge among

[1]For empirical investigations of how financial considerations influence politicians' candidature decision, see our literature review below.

[2]Note that to pay such a high compensation is not a priori excluded "because the money is not available". In the long run equilibrium, the base salary of faculty would be correspondingly lower, and the department as a whole would still pay the same amount of money for the same amount of work.

voters, so that our model is most relevant as a model of small and homogeneous electorates where voters have similar preferences over candidates.[3] In the first stage, every potential candidate decides whether he is willing to run for office. In the second stage, people vote on the self declared candidates; if no candidate is willing to run, a default outcome obtains.

The whole game (including the entry stage) has a unique equilibrium. Under some relatively mild sufficient condition, we show that bad candidates are more likely to run for office than good candidates, even though candidates are also consumers of the service provided and from this point of view, it is more attractive for good candidates to serve. However, this effect is counterbalanced by two other effects: First, good candidates are at least stochastically more likely to have higher (opportunity) costs for serving. The second point is more subtle: When the job is quite unattractive and every candidate would prefer that some other candidate – rather than he himself – does the job, candidates have an incentive to free-ride on each other. We show that, in equilibrium, good candidates can free-ride on bad candidates. The reason is that the decision to run for office is relatively less costly for a bad candidate (he will only be elected if there is no good candidate willing to run) than for a good candidate, who knows that he will have to serve in office, if he declares himself willing to run.

Also, we derive some interesting comparative static results with respect to the politician's wage. We show that the expected quality of running candidates might actually decrease as the remuneration of the official increases. This result is interesting since one of the arguments in discussions about politicians' remuneration is that higher wages will attract more competent candidates. The intuitive reason why this argument does not apply in all circumstances is that, while a higher remuneration has a direct effect that it makes the office more attractive for the elected official, there is also an indirect effect that all other candidates are now more willing to run; the higher probability that other candidates run makes it more attractive for a competent candidate (who would choose to run if there were no one else to fill the job) to try to freeride on them, and so an increase in the remuneration might induce some competent candidates *not* to run. However, for sufficiently high levels of remuneration, the job becomes more and more attractive and eventually the direct effect will dominate.

While most previous literature has neglected the influence of financial issues on candidates' decision to run, there have recently been some interesting new papers on this issue. Caselli and Morelli (2003) analyze a citizen candidate model with vertically

[3] Also, these electorates normally have to decide on filling positions that are often rather unattractive for the office holder, which is the most interesting case in our model.

differentiated candidates. They assume that bad candidates have lower opportunity costs than good candidates. In this framework, Caselli and Morelli show that low quality candidates are more likely to run for election than good ones, a result that is also found in our model.

The crucial difference between Caselli and Morelli (2003) and our work is that they have a continuum of political positions to be filled, so that each individual candidate does not influence the quality of service provided and consequently does not consider his potential influence on average quality when he decides whether to run. In our model, there is just one position to be filled, and so a potential candidate considers both his direct remuneration and the possible improvement of the service quality level (if he rather than a worse candidate serves) as the benefits of running for office, which he weighs against the cost of serving.

The service quality effect is in reality most relevant in small electorates. This is also the scenario in which the assumption that each potential candidate's quality is common knowledge in the electorate is most convincing. Caselli and Morelli's model probably applies best to the election of a large national assembly, where each potential candidate cannot significantly improve the quality of legislating (but where in principle even small improvements would be very beneficial as so many people consume the service).[4]

The assumption that potential candidates care about the quality level of the office holder generates an interesting endogenous "free riding" effect in our model that works in the same direction as the direct cost effect: Good candidates often free ride on bad ones. This effect is likely to be relevant for many rather unattractive positions.

The free riding effect distinguishes our paper also from Poutvaara and Takalo (2002). They too analyze a model based on the assumption that more able candidates have higher opportunity costs for serving in office. However, their framework is not a pure citizen candidate model in which candidates decide directly whether to run for the office. Rather, they introduce a further selection stage after the entry decision. From the set of candidates who are willing to run for election only two randomly selected ones stand against each other in the final vote.This mechanism creates the possibility of a crowding out effect on good candidates: If a large number of low quality candidates enters, the chance of a high-ability candidate to be admitted to the final election decreases and so does his incentive to enter the race and spend the campaign costs.

Carillo and Mariotti (2001) also analyze the problem of candidate quality. However, in contrast to our model, the main actors in their setup are the parties who have to

[4]Caselli and Morelli also study the effects of reputational externalities between politicians and the endogenous determination of the wage in a dynamic framework (by the elected politicians themselves).

choose whether to nominate experienced politicians (about whom the electorate and the parties know a lot) or a fresh candidate of highly uncertain quality (and to start to learn about the quality of this new candidate). They show that the electoral competition induces parties to behave too conservatively from a social point of view. That is, they stick too often to mediocre incumbents. Our model and theirs are complements in the sense that they are concerned with the parties' choice (among individuals who are willing to run), while we analyze the entry decision of individual potential candidates.

There is also some evidence of how financial considerations influence politicians' candidature decision. For example, Hall and Houweling (1995) analyze the consequences of a 1990 law that increased significantly the pension of congressmen who retired after 1992, and find that a significant number of congressmen who otherwise would have retired in 1990, decided to re-run for office in order to receive this financial windfall. Groseclose and Krehbiel (1994) analyze a different natural experiment in the U.S. House of Representatives and also find a significant influence of financial considerations. Diermeier, Keane and Merlo (2002) analyze the decision of congressmen to re-run for office, and the importance of monetary versus non-monetary components in this decision. While the empirical studies we know of focus on the U.S. Congress, we would expect that financial and opportunity cost considerations are relatively even more important in less prestigious offices that come with smaller "ego-rents".

Our paper also draws on the literature on the private supply of public goods, and on the citizen candidate model. If there is no remuneration, candidates provide a public good if they are willing to serve as politician and so free riding and underprovision results from this literature of course apply in our model.[5] The paper from this literature that is most related to our question is Bilodeau and Slivinski (1996). They study a finite horizon war of attrition game of volunteering for some public service which everybody prefers that somebody else performs, but would rather perform himself than live without the service. Their result is that the most able candidate volunteers immediately and is in stark contrast to our model where in a similar situation, society will end up with the least able politician. The fundamental reason for this difference is that in Bilodeau and Slivinski (1996), the best player has the weakest position since there is a time when all but the best player are unwilling to provide the service, and so the best action of the best player is to volunteer immediately. In contrast, in our model, better candidates have a relatively good outside option. Suppose there are just two candidates, one slightly better than the other. If the good candidate decides not to

[5]Some references concerning the private provision of public goods include Bergstrom et. al. (1986), Bliss and Nalebuff (1984), Palfrey and Rosenthal (1984).

run, he can still hope that the bad candidate is running and provides the public service. A symmetric argument does not hold for the bad candidate, because his decision to run is relevant only if the good candidate is not running,[6] and so the bad candidate should make his decision whether to run optimally conditional on the good candidate not running. The lower outside option makes a bad candidate in our model relatively more likely to run.

Our paper is also related to the literature on citizen candidates pioneered by Osborne and Slivinski (1996) and Besley and Coate (1997). With respect to voters' preferences, our setting with vertical differentiation is a special case of the Besley and Coate setup. Furthermore, we focus on the remuneration of the office as one of the important features which can make running for office attractive and derive the benefit of being in office endogenously (which is exogenous in both Osborne and Slivinski, and in Besley and Coate). The other main difference between our basic model and both Osborne and Slivinski and Besley and Coate is that they assume a (small) fixed cost of running for office, which, in their models, is used to decrease the number of equilibria. While our basic model does not have campaign costs, we analyze this as an extension in section 4.

The rest of this paper proceeds as follows. Section 2 describes the basic model. In section 3, we analyze candidates' decision to run, in particular how this decision depends on the candidate's level of competence. In section 4, we derive some comparative static results concerning the effect of a wage increase and a campaign cost increase on potential candidates' decision to run. Section 5 concludes.

2 The basic model

The polity we consider is a group of N people (called the *voters*) who have to elect a politician who provides a public good for all members of the group. We call the public good provider a "politician", but as mentioned in the introduction, this term should be interpreted widely as any person who is elected to his office.

There are I "potential candidates" for this job and $N - I$ "ordinary voters" (who cannot be candidates) among the N people. Candidates differ in their competence: Candidate i, $i = 1, ..., I$ can create a per capita value for every member of the society (including himself) of v_i. Without significant loss of generality, we assume that candi-

[6] If the good candidate runs, he will be elected independent of whether the bad candidate runs or not, so the bad candidates decision is irrelevant in this case.

dates are ordered with respect to their ability such that $v_i > v_j$ if $i > j$.[7] We assume that the v_i's of all candidates can be perfectly observed by every voter; this is in line with the literature on citizen candidate models that assumes that voters can observe all relevant characteristics of candidates. If no candidate is willing to serve, every citizen gets a default payoff of $v_0 < v_1$.

For each candidate i, there is a (gross) personal cost c_i of being elected and serving as the politician. This cost includes the opportunity cost of the time spent in office. However, the candidate could also actually enjoy to perform the function because of the "ego-rents" associated with the office, so we do not exclude the case that $c_i < 0$. In exchange for incurring the personal costs, the politician receives a remuneration denoted $r \geq 0$, which is financed by a head tax levied on all other members of the society. Hence, the utility of candidate i if he is elected politician is

$$v_i + r - c_i, \tag{1}$$

and the utility of any other member (ordinary voter or unsuccessful candidate) is

$$v_i - \frac{r}{N-1}. \tag{2}$$

We assume that the personal costs of candidate i, c_i, can only be observed by the candidate himself; however, the distribution G_i from which it is drawn is common knowledge. Admitting different cost distributions for different candidates allows us to study the effects of correlation between productivity in the political position and (opportunity) costs.

We assume that for every potential candidate, it is possible that this candidate would not want to run: $Prob(v_i + r - c_i < v_0) > 0$ for all i, or equivalently

Assumption 1. *For any r and any i, $G_i(v_I - v_0 + r) < 1$.*

This assumption is satisfied for any unbounded random variable, for example, if c has a normal, log-normal or exponential distribution. Note that the probability that a candidate would not want to run can be arbitrarily small, as long as it is positive. We will show below that, if Assumption 1 holds, the game has a unique equilibrium that we can find by iterated elimination of strictly dominated strategies.

The timing is as follows: At first, candidates decide simultaneously whether to run for office. If no potential candidate is willing to run, then a default policy will be

[7] This is the generic case if abilities are determined by a random draw from a continuous distribution.

implemented that yields a benefit of $v_0 < v_1$ for every voter.[8] If only one candidate runs for office, then he automatically becomes the office holder.

If there are two or more candidates, then an election is held. Each vote is counted with probability $1 - \varepsilon$, for ε a small, but positive number. In this setting, Messner and Polborn (2003) show that trembling in the voting process induces a unique voting equilibrium, in which the candidate with the highest v_i is elected.[9]

The main difference between previous citizen candidate models and the present one is in our modeling of the costs of politicians. Osborne and Slivinski (1996) and Besley and Coate (1997) assume that there is a fixed cost of entry (sometimes interpreted as campaigning costs) for all candidates who decide to run for office. In contrast, we assume in the basic model that (just) *running* for office is costless.[10] This appears reasonable in our context. Given the set of running candidates, there is a unique winner of the election subgame, and the losing candidate(s) should not spend a lot of money on campaigning, in particular since by assumption voters know the competence of the potential candidates anyway.

Another difference is that we model explicitly the (gross) personal cost c_i of candidate i if he is elected and serves as the politician. This cost is "gross" in the sense that it does not take account of the remuneration for the job. It includes the opportunity cost of the time spent in office, but since the candidate could also actually enjoy to perform the function, we do not exclude the case that $c_i < 0$. While c_i can only be observed by the candidate himself, the distribution G_i from which it is drawn is common knowledge. Previous citizen candidate models assume perfect information about all characteristics of potential candidates. Our assumption concerning the information distribution with respect to the opportunity costs of potential candidates is probably more realistic in a number of settings and also, as we will see below, serves to narrow down the set of possible equilibria.

[8] Alternatively, v_0 can be interpreted as the (net) utility of every voter, if candidate 0, who is known to be willing to run, is elected. In this interpretation, I is the set of potential candidates who could do a better job than candidate 0, since only their decision to run is relevant (candidates who are worse than candidate 0 will never be elected).

[9] The intuition for this result is straightforward. Since votes are counted only with a probability strictly smaller than 1, every candidate who receives votes from at least one member of the society has a positive probability to win the elections. Therefore, for any voting profile in which not all votes are concentrated on the best candidate, all voters who cast their vote for the worst candidate among the ones who receive a positive number of votes are not behaving optimally. By switching their vote to any of the higher ranked candidates they could reduce the probability that the worst candidate is elected and increase the probability that one of the better candidates wins.

[10] However, in section 4.2, we will analyze what happens if we introduce positive campaign costs.

3 Equilibrium

We will now start to derive the equilibrium behavior of candidates. For ease of exposition, let us start with the case that there are only two potential candidates, 1 and 2, where candidate 2 has the higher ability. If candidate 2 decides to run, he will be elected as the politician, independent of whether candidate 1 decides to run for office or not. Therefore, candidate 1's candidature decision is only payoff relevant if player 2 decides not to run, and equation (1) guarantees that this happens with positive probability. Candidate 1 should run if and only if his utility from serving as politician is greater than his utility from the default outcome, v_0, so the following strategy is the best strategy for candidate 1, irrespective of which strategy candidate 2 plays:

$$s_1(c_1) = \begin{cases} \text{run} & \text{if} \quad v_1 - v_0 + r > c_1 \\ \text{don't run} & \text{otherwise} \end{cases} . \tag{3}$$

From candidate 2's perspective, candidate 1 runs with probability $G_1(v_1 - v_0 + r)$. Hence, if player 2 decides not to run, his expected payoff is

$$G_1(v_1 - v_0 + r)[v_1 - \frac{r}{N-1}] + [1 - G_1(v_1 - v_0 + r)]v_0. \tag{4}$$

If he runs, player 2's utility is $v_2 + r - c_2$, and so player 2's optimal strategy is

$$s_2(c_2) = \begin{cases} \text{run} & \text{if } v_2 + r - c_2 > G_1(v_1 - v_0 + r)[v_1 - \frac{r}{N-1}] + [1 - G_1(v_1 - v_0 + r)]v_0 \\ \text{don't run} & \text{otherwise} \end{cases} \tag{5}$$

It follows that, in the case of two candidates, the unique Bayesian Nash equilibrium is given by (3) and (5). The arguments presented easily generalize: Once the behavior of all players j, $j < i$, is pinned down, we can show by a similar reasoning that player i has a strictly dominant strategy.

Let p_j denote the probability that player j will run as a candidate. Furthermore, let Λ_i be the conditional expected utility of an ordinary (inactive) voter, given that no candidate ranked strictly higher than i is running for office:

$$\Lambda_i = \sum_{k=1}^{i} p_k \Big[\prod_{m=k+1}^{i} (1 - p_m) \Big] \Big[v_k - \frac{r}{N-1} \Big] + \prod_{m=1}^{i} (1 - p_m) v_0. \tag{6}$$

For future reference it is useful to note that, conditioning on whether candidate i runs or not, yields the following relation between two subsequent Λ_i:

$$\Lambda_i = p_i \left(v_i - \frac{r}{N-1} \right) + (1 - p_i)\Lambda_{i-1}, \tag{7}$$

with $\Lambda_0 \equiv v_0$.

As in the two candidate example above, the optimal strategy of player 1 is given by (3), and in general, the optimal strategy of player i is given by

$$s_i(c_i) = \begin{cases} \text{run} & \text{if } v_i + r - c_i \geq \Lambda_{i-1} \\ \text{don't run} & \text{otherwise} \end{cases} \tag{8}$$

Equation (8) is well defined as long as all p_j, $j = 1 \ldots, i-1$ are known, and determines

$$p_i = G_i\big(v_i + r - \Lambda_{i-1}\big). \tag{9}$$

Consequently, the equilibrium can be calculated recursively, and we have shown the following:

Proposition 1. *The unique Bayesian Nash equilibrium of the game with I candidates is characterized by all candidates $i = 1, \ldots, I$ playing the strategy given in (8).*

At this point, it is helpful to discuss which of our assumptions are necessary to generate this equilibrium. The decisive point is that our model generates an "electoral ranking" of potential candidates such that a candidate will be elected if and only if no higher ranked candidate decides to run. If candidates are strictly differentiated by their competence, then the electoral ranking is obviously equal to the competence ranking. However, even if there are two (or more) candidates who have exactly the same level of v, there are voting equilibria that generate a strict electoral ranking: For example, suppose that $v_1 = v_2$, but whenever both candidate 1 and 2 run (and no higher ranked candidate runs), then candidate 2 is elected; since voters are indifferent between the two candidates, such a voting equilibrium clearly exists.[11] Given this electoral ranking, the candidates make their entry decisions exactly as they do in our model when $v_2 > v_1$: Candidate 1 enters whenever this gives him a higher utility than the default outcome, and candidate 2 chooses to enter given candidate 1's equilibrium entry strategy.

What would happen if candidates do not know their competence ranking at the time when they decide whether to run for office? This destroys the strict electoral ranking. Suppose, for example, that the electorate and candidates 1 and 2 learn about their abilities only *after* the candidates' entry decisions are made; if both enter, each has a probability of 1/2 of being the more productive candidate (and hence of being elected if both run). Therefore, Candidate 1's optimal entry decision depends on Candidate

[11] There are of course also voting equilibria in which voters randomize over candidates; such a randomization equilibrium would be more problematic in that it does not generate a strict electoral ranking of potential candidates. Thus, if some candidates have exactly the same ability, there exist voting equilibria such that our results go through essentially unchanged, but other voting equilibria appear for which there may be multiple equilibria on the entry stage.

2's equilibrium strategy and vice versa. In such a situation, there may be multiple pairs of strategies for candidates that are mutually best responses to each other. The assumption that candidates know their place in the competence ranking is therefore essential for the recursive equilibrium determination that yields a unique equilibrium in our model.

3.1 Competence and the decision to run

We will turn to a characterization of the equilibrium, in particular we will establish a relation between a candidate's competence and the likelihood that the candidate runs.

This is an important question, because it tells us whether a democracy will be successful in choosing the best available candidate for a public position. In practice, there may be two problems with candidate selection through elections. First, voters may not be perfect in recognizing who would be the best candidate from a set of available choices. Second, the set of choices may be endogenous, and may not contain the best possible candidates.

In our model, we consider a framework in which the first problem is absent, since voters can perfectly recognize each candidate's quality. This is clearly an abstraction from reality, but it allows us to focus entirely on the second problem.[12]

Intuitively, there are two reasons why potential candidates might choose to run for office. First, being elected to the office may be attractive for a candidate, if his opportunity cost (minus the ego rents he derives from the office) is small compared to the remuneration of the office. If opportunity costs tend to be higher for good candidates (at least in a stochastic way), then bad candidates will be more likely to run.

A second reason may be the more important reason to run for a number of positions which are not too attractive for the office holder: Potential candidates are also consumers of the public service, hence, if a candidate does not run, he runs the risk that the position is filled by a worse candidate (or not at all). Here, there are two potential effects, going in different directions. On the one hand, a good candidate can deliver a better service than a bad candidate could, which makes running more attractive for a good candidate. On the other hand, we will show below that good candidates have a better opportunity to free ride on other candidates. Intuitively, conditional on the decision about running being important at all (i.e., there is no better candidate who

[12] Another paper that analyzes the importance of the endogeneity of the candidate set (in a different context) is Dutta, Jackson and LeBreton (2001).

runs), the *expected* quality of the next best candidate who runs is higher for a high quality candidate.

The following Lemma 1 establishes formally that Λ is a nondecreasing function. This means that the endogenous outside option, if not running, is better for high competence candidates.

Lemma 1. *$\Lambda_i \geq \Lambda_j \iff i \geq j$, and $\Lambda_i > \Lambda_j$ if $i > j$ and there exists k such that $j < k \leq i$ and $p_k > 0$*

Proof. From (6), it follows that $\Lambda_i \leq v_i - \frac{r}{N-1}$, with equality only if $p_i = 1$. Using (7), it follows that $\Lambda_i - \Lambda_{i-1} = p_i(v_i - \frac{r}{N-1} - \Lambda_{i-1}) \geq p_i(v_i - v_{i-1}) \geq 0$. □

The following assumption relates opportunity costs to candidate quality.

Assumption 2. *For $i > j$ (remember that this implies $v_i > v_j$), assume that*

$$G_i(c + v_i - v_j) < G_j(c), \text{ for all } c.$$

This implies that the distribution of c_i stochastically dominates the distribution of c_j at least by a shift of $v_i - v_j$. Intuitively, this assumption is fairly weak, when we consider that the opportunity cost is mostly equal to the forgone earnings that a candidate could make in other jobs. Remember that $v_i - v_j$ is the difference in the productivity of candidates i and j *per capita* of the population; the difference between the social product of candidates i and j is much larger, $N(v_i - v_j)$. Hence, if the total product of a candidate in the public office and in the private sector are roughly similar, Assumption 2 is quite likely to be satisfied.

If Assumption 2 holds, bad candidates are more likely to run for office than good ones:

Proposition 2. *Suppose that Assumption 2 holds. In equilibrium, $p_i < p_j$ for all $i > j$.*

Proof. From (9), $p_i < p_j$ is equivalent to $G_i(v_i + r - \Lambda_{i-1}) < G_j(v_j + r - \Lambda_{j-1})$. By $\Lambda_{i-1} \geq \Lambda_{j-1}$ (from Lemma 1) and Assumption 2, the claim follows. □

It is interesting to compare the result of Proposition 2 with Caselli and Morelli (2003). They also obtain the result that low quality candidates are more likely to run for election than good ones in a citizen candidate model with vertically differentiated candidates. Both models share the assumption that the opportunity cost of serving in office is higher for highly productive candidates.[13] In Caselli and Morelli, this is the only reason why bad candidates run more often than good candidates.

[13] In Caselli and Morelli, this cost difference is deterministic while in our model, it holds stochastically.

In addition, we assume that candidates consume the public service supplied by the politician. This effect is not present in Caselli and Morelli, because they assume that there is a continuum of public positions and all citizens' utility depends on the average quality of politicians, which cannot be influenced significantly by a single candidate.[14] In our model, this additional effect can go in both directions. If all candidates' costs were drawn from the same distribution (i.e., without Assumption 2), then the direct net payoff from serving as politician (i.e., remuneration minus opportunity cost) is increasing in a candidate's type, and consequently, if a good type were the only potential candidate, his probability of running would be higher than the corresponding probability of a bad type.

The incentives to enter the race are reversed only by the strategic interaction between candidates at the entry stage. Good candidates have a higher endogenous outside option from the possibility of another candidate volunteering for the job. The possibility of highly qualified candidates to free ride on low quality types reduces their net benefit from entering, and hence reduces their willingness to run for election. Assumption 2 gives a condition for the direct effect of increased competence on the net payoff to be at most zero. This is sufficient (but by no means necessary) for the overall effect to go in the direction of the indirect, strategic effect, so that good candidates are more reluctant to enter the race than bad ones.

While proposition 2 shows that bad candidates run more often than good candidates, this may not be too bad, if there are sufficiently many potential candidates. After all, just *one* good candidate who is willing to serve is needed. However, the following example shows that the effect in Proposition 2 may be so extreme that voters for sure have no good candidate available.

Consider an unattractive job for the politician, with no remuneration ($r = 0$) and positive costs for all potential candidates ($G(0) = 0$). In this situation, every candidate prefers that some other candidate who is higher ranked gets the job rather than to do the job as politician himself. The only motivation for running as a candidate is then that the quality of service might be worse if you do not run, either because some worse candidate is elected, or because no one is running for office and everyone gets the default payoff.

[14]Consequently, Caselli and Morelli is best interpreted as describing the election to a large national parliament in which the competence of a single politician does not significantly influence the overall quality of the elected body, while our model applies best for much smaller groups, in which each potential citizen-candidate knows that he could make a difference for the service quality that he will consume.

Example 1. *Suppose* $v_0 = 0$, $v_1 = 1$, *and there is an arbitrary number of candidates with* $v_i \in (1; 1.4]$. *For every candidate, the costs are drawn according to the following cdf*

$$G(c) = \begin{cases} 0 & \textit{for } c \leq 1/2 \\ 2(c - 0.5) & \textit{for } 1/2 < c \leq 0.99 \\ 0.98 & \textit{for } 0.99 < c \leq 2 \\ 1 & \textit{for } c \geq 2 \end{cases}$$

With this cost distribution, every player has a probability of 2% to have a very high cost realization of 2, and with 98% probability, the cost is drawn from a uniform distribution on $[0.5; 0.99]$.[15]

In Example 1, only the worst candidate (player 1) ever runs with positive probability. To see this, note first that the dominant strategy for player 1 is to run if and only if his costs are lower than 1, hence with 98% probability. Given player 1's strategy, any other player's outside option (his utility if he decides not to run) is greater or equal to $0.98v_1 = 0.98$. On the other hand, running as a candidate and consequently being elected gives a payoff of $v_i - c_i \leq 1.4 - 0.5 = 0.9 < 0.98$. Therefore, no other candidate apart from player 1 will ever consider to run.

The following proposition generalizes the result of Example 1 and gives a necessary and sufficient condition for incompetence to reign with certainty:

Proposition 3. *Let* $c^i_{min} = \sup\{x|G_i(x) = 0\}$ *denote the minimal possible cost realization. Only the worst candidate, candidate 1, will run with positive probability, if and only if the following inequality holds for all* $i \geq 2$

$$c^i_{min} \geq v_i + r - G_1(v_1 - v_0 + r))[v_1 - \frac{r}{N-1}] - (1 - G_1(v_1 - v_0 + r))v_0 \tag{10}$$

This is independent of the number of other potential candidates.

Proof. The proof for sufficiency follows the same arguments as given in Example 1 above and is therefore omitted. For necessity, suppose that in a certain equilibrium only the worst candidate runs with positive probability, but that (10) does not hold. But then, candidate I would be strictly better off running if he has a low cost realization. □

It is interesting to compare Proposition 3 with the results of Bilodeau and Slivinski's (1996, henceforth BS) model of "toilet-cleaning and chairing departments". They study

[15] As before, the possibility of a very high cost realization is essentially to be seen as a refinement that achieves that every candidate may find it not optimal to run, even if he knew that he were the only candidate. The details (that there is an atom on 2) do not matter in this respect.

a finite horizon war of attrition game of volunteering for some public service which everybody prefers that somebody else performs, but would rather perform himself than live without the service. This setting is relatively similar to the one we use in Example 1 and Proposition 3, even though we have asymmetric information concerning the candidates' costs, and there is a slight chance that a candidate might prefer not to perform the service even if he were the only potential candidate.

BS's result is that (ceteris paribus) the most able candidate volunteers immediately. The reason is that, in their model, there is a time after which all but the best candidate have the dominant strategy not to volunteer, because the payoff time is too short in comparison to the benefit they can create. Hence, in the subgame starting at this time, there is a unique equilibrium in which the best candidate volunteers. Knowing this, the other candidates have no incentives to run immediately before this moment, and the best candidate might as well volunteer earlier. The whole game unravels this way, and the best candidate will volunteer immediately. This result is in stark contrast to our result above that in certain situations, the society will end up with the least able politician.[16]

The important difference between the two models is that in our model, there is a stage where candidates have to decide whether to run, and only the elected candidate is required to perform the service. Since bad candidates are only elected if they are the best running candidate, it is relatively less costly for them to run than it is for the best candidate who is always elected if he decides to run.

We see our model as complementary to BS's one. For some public services, it is certainly realistic to assume that there is no election, but the task is performed by the first volunteer (e.g., toilet cleaning). For other public services, the decision process is probably more structured, and it is more realistic to assume that there is a fixed time when candidates have to decide whether to run. If this moment is before the time when bad candidates have "do not run" as their dominant strategy, then BS's backward induction argument loses its power, and our model of simultaneous decisions becomes more relevant.

We should note that there is another procedure apart from voting that leads to the same candidate behavior. Suppose there is a search committee whose task it is

[16] Another paper that considers a war of attrition game of volunteering is Bliss and Nalebuff (1984). In contrast to BS, they assume that players are infinitely lived. Furthermore, cost types in their model are private information as they are in ours. Their main result is that the waiting time before an individual decides to volunteer is increasing in the individual's cost parameter. Hence, as in BS, the individual with the lowest cost parameter will be the one to provide the public service, even though this will typically not happen immediately.

to find a suitable candidate to fill an unpleasant position, for example the chair of an economics department. Suppose that the search committee identifies the suitable candidates, ranks them from best to worst and starts its job by trying to persuade the best candidate on their list. If the first candidate declines, they move on to the next best candidate and so on, until they have found a "victim" who agrees to serve. In this setting, the decision of potential candidates whether to be willing to serve is essentially the same as when they have to declare simultaneously whether they would be willing to run in an election.

The difference is that the committee need not start with the best available candidate, and as shown, it might just make sense to choose some other sequence in which potential candidates are asked. For example, consider the setting of Example 1. If we ask the best candidate last, we know that he will run with a probability of 98 percent, since his choice is either to provide the public service, or to get v_0. Knowing that the best candidate will be willing to serve with very high probability, all worse candidates will decline, if asked earlier. Hence, in this example, it would be optimal to ask the best candidate last. However, we should stress that this result depends on the parameters, and is not true in general.

We now turn to the case that the condition in Proposition 3 is not satisfied, so that better candidates run with positive probability. However, we still look at unattractive jobs in the sense that the remuneration is $r = 0$, while every candidate has a positive opportunity cost.

If we increase the number of candidates, does it eventually become (almost) certain that a good candidate runs? If the answer to this question were affirmative, then it might make sense for a society with many potential candidates to choose a relatively low remuneration, knowing that they need just one good candidate to fill the job.

In order to analyze this case, it is helpful to have a measure of welfare in this society. Conditioning on whether the best candidate runs or not, and using the definition of Λ, the (unconditional) expected utility of an ordinary voter can be written

$$W = p_I\left(v_I - \frac{r}{N-1}\right) + (1 - p_I)\Lambda_{I-1}. \tag{11}$$

We have two reasons for being interested in this measure. From a normative point of view, if there are relatively few candidates in proportion to the population, it might be warranted to ignore the candidates' expected utility and to focus on W as a welfare measure. From a positive point of view, suppose that initially the society votes on the framework in which this public good provision game takes place (for example, how much r to pay to the politician, or whether to restrict the number of people who are allowed to run). In such a referendum on the wage or the institution, the number of ordinary voters will be larger than the number of potential candidates, and hence they will choose an institution in order to maximize W.

We can now show that there is a welfare loss even if there are many candidates and the condition in Proposition 3 is not satisfied, so that there is a positive probability that not just the worst potential candidate runs.

In particular, we consider a scenario where there are I candidates who are distributed equidistantly on the interval $[\underline{v};\bar{v}]$, with $v_1 = \underline{v}$ and $v_I = \bar{v}$ and whose costs are independent draws from the same distribution G on $[c_{min}; c_{max}]$, with $c_{min} > 0$. We consider the limit $I \to \infty$, and ask whether it is almost certain that a candidate close to $\bar{v}$ runs (and wins the election).

Suppose this were the case. But then, the outside option of a good candidate would be close to $\bar{v}$, and so he would be better off not running because running gives him (at most) $\bar{v} - c_{min}$, and the same conclusion holds for potential candidates with a slightly lower v, so this yields a contradiction.

Proposition 4. *Let $v_0 = 0$ and $r = 0$ and denote by $W(I)$ the welfare of an ordinary voter if there are I candidates who are distributed equidistantly on $v \in [\underline{v};\bar{v}]$, with $v_1 = \underline{v}$ and $v_I = \bar{v}$ and whose costs are independent draws from the same continuous and increasing distribution G on $[c_{min}; c_{max}]$, with $c_{min} > 0$. Then*

$$\lim_{I\to\infty} W(I) = \max\{G(v_1)v_1, \bar{v} - c_{min}\}$$

Proof. See appendix. □

This result is significant for a number of U.S. states, which do not pay their state legislators anything which comes close to a reasonable estimate of their opportunity cost. For example, New Hampshire pays its state legislators a total of \$ 200 for a two year term, and no living expenses during sessions.[17] Even if there are very many candidates who all would be willing to serve as legislator if they were the only candidate, it is rather improbable that a very good candidate will run for office in equilibrium, and a welfare loss is likely.

It is also interesting to relate the results of this subsection to the efficiency results for a representative democracy obtained by Besley and Coate (1997). They show that the equilibrium allocation in a representative democracy is Pareto efficient, provided that only one or two candidates are running in equilibrium and the costs of running are small.[18]

[17]See the Book of the States (2000/2001).

[18]For details, see Besley and Coate, Propositions 11 and 12.

This efficiency result seems at odds with the result of our Example 1 and Proposition 4, both of which have an air of inefficiency. However, the notion of efficiency in Besley and Coate (1997) is a very weak one (Pareto efficiency, including the politicians' utilities). If we apply it to our model, the equilibrium is "efficient" (in this weak sense) as well, as long as we only consider other allocations that have the same remuneration for the office holder. Consider example 1 in a setting with very many voters: The only way to make ordinary voters better off than in the equilibrium is to have a better candidate running for office. However, given that the remuneration is zero, this better candidate would be worse off than he is in equilibrium. A similar result holds in the setting of Proposition 4.

Note that the efficiency property of the equilibrium relies on the remuneration being fixed. If there are many voters, then it is evidently Pareto better to choose a remuneration which induces better candidates to run, and so the equilibrium with zero remuneration is not Pareto efficient in this class.[19]

4 Comparative statics and extensions

In this section, we will be looking at the comparative static effects of changing the conditions for candidates. Specifically, we will analyze the effect of a wage increase, of campaign costs, which every self-declared candidate needs to pay, and at the effect of excluding some potential candidates from running.

4.1 Wage

In this section, we want to look at the comparative static effect of a wage increase for the elected official. This is an important question also because one of the often heard arguments for increasing politicians' wages is that a higher wage will induce more competent candidates to run for office.

In order to consider this argument, we will look at the case that there are two potential candidates. The probability that the worse candidate runs is $G_1(v_1 - v_0 + r)$, which is clearly increasing in r. This is quite intuitive, since the worse candidate makes his decision to run conditional on being the only candidate; hence, for him the only effect is that a higher wage makes the job more attractive.

Consider now the decision of the better candidate. His optimal choice is described by (5), and consequently he runs with probability $G_2(v_2 + r - \Lambda_1)$. Differentiating with

[19] Of course, if a higher remuneration is Pareto better than zero remuneration, it would be chosen if the remuneration is determined endogenously in an election.

respect to r yields

$$\frac{dG_2}{dr} = g_2(\cdot)\left\{1 - g_1(v_1 - v_0 + r)\left[v_1 - v_0 - \frac{r}{N-1}\right] + \frac{G_1(v_1 - v_0 + r)}{N-1}\right\} \tag{12}$$

Hence, whether the probability that the good candidate runs increases or decreases with r depends on whether the term in curly brackets is positive or negative. A negative marginal effect of remuneration on the probability that the better candidate runs is more likely

- the higher is $g_1(v_1 - v_0 + r)$, because this measures how much the probability that candidate 1 runs, increases,
- the greater $v_1 - v_0$, as this measures how important the increased participation rate of candidate 1 is for candidate 2,[20]
- the smaller is r (the larger r is, the smaller the net benefit that candidate 2 has from candidate 1 running, as the wage of the politician must also be paid by candidate 2)

It is also interesting to note that candidate 2's probability of running will always decrease with a marginal wage increase if $v_1 - v_0 + r$ is an atom of candidate 1's cost distribution. We now turn to an example of these effects.

Example 2. *Let $v_0 = 0$, $v_1 = 1$ and $v_2 > v_1$. The cost of candidate 1 are normally distributed with mean μ and standard deviation σ. (No assumption about the cost distribution of candidate 2 is needed).*

In this example, candidate 2's outside option is

$$\begin{aligned}\Lambda_1 = \Phi\left(\frac{1+r-\mu}{\sigma}\right)\left[v_1 - \frac{r}{N-1}\right] + \left(1 - \Phi\left(\frac{1+r-\mu}{\sigma}\right)\right)v_0 = \\ v_0 + \Phi\left(\frac{1+r-\mu}{\sigma}\right)\left[v_1 - v_0 - \frac{r}{N-1}\right],\end{aligned} \tag{13}$$

where $\Phi(\cdot)$ is the cumulative distribution of the standard normal distribution. Differentiating with respect to r yields

$$\frac{d\Lambda_1}{dr} = \phi\left(\frac{1+r-\mu}{\sigma}\right)\frac{1}{\sigma}\left[v_1 - v_0 - \frac{r}{N-1}\right] - \frac{1}{N-1}\Phi\left(\frac{1+r-\mu}{\sigma}\right) \tag{14}$$

[20] We will always assume that $v_1 - \frac{r}{N-1} > v_0$, i.e., it is better for ordinary voters that candidate 1 takes the job than that no candidate is willing to serve.
If this assumption is not satisfied, we would need to specify what happens, if only candidate 1 runs. For example, if candidate 1 is then not elected, we can remove candidate 1 from the set of effective candidates and relabel the subscripts of the other candidates.

Assume now that the number of citizens N is large enough that we can neglect the terms containing N,[21] and substitute the values of v_0 and v_1:

$$\frac{d\Lambda_1}{dr} = \phi\left(\frac{1+r-\mu}{\sigma}\right)\frac{1}{\sigma} \tag{15}$$

Candidate 2's probability of running will be increasing in r if and only if $d\Lambda/dr < 1$. Using the definition of the density of the normal distribution, $d\Lambda/dr < 1$ is equivalent to

$$\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{1+r-\mu}{\sigma}\right)^2} < \sigma^2 \tag{16}$$

This is certainly satisfied if the remuneration is sufficiently attractive, i.e., if r is large relative to $\mu - 1$. However, if $\mu \geq 1$ and $\sigma < (2\pi)^{-1/4} \approx 0.63$, then there exist some levels of r (in a neighbourhood of $\mu - 1$) where good candidates run less often if r increases marginally.

4.2 Cost of running

Up to now, we have assumed that just running for office is costless. In our setting, this appears to be a reasonable assumption, since everyone knows the quality of each candidate and so campaigning is not very useful.

However, just as the remuneration of the office holder is in principle a matter of choice for society, society could also choose that there is an "entrance fee" to be paid by every candidate. For example, candidates who want to run in the 2003 leadership race of the Liberal party in Canada have to pay a fee of 75,000 dollar for the right to compete. It is therefore important to understand the effects of a cost δ that every candidate has to pay.

Again, consider a case with two potential candidates. The worse candidate's decision is relevant only if the better candidate does not run, hence with probability $(1-p_2)$; in this case, running and serving in office gives the worse candidate $v_1 + r - c_1$, rather than v_0, if he does not run. Candidate 1 should therefore run if and only if the expected utility gain from running (and this decision being relevant, because candidate 2 does not run) is larger than the cost of running that must be paid in any case, hence if,

$$(v_1 + r - v_0 - c_1)(1 - p_2) \geq \delta$$

[21]The assumption of a large number of voter is used only for convenience. It allows us to neglect the effect on the entry decision which comes through the increase of the contribution to the wage payment of a non-winning candidate. In this way we are able derive a simple closed form condition under which the better candidate's willingness to run is decreasing in the promised wage. It should be clear though that the effect which yields this result for large N is present also if the number of voters is small.

Therefore, candidate 1 runs with probability

$$p_1 = G_1\left(v_1 + r - v_0 - \frac{\delta}{1 - p_2}\right) \tag{17}$$

Note that, in the basic model, candidate 1 had only to consider the scenario that arises if the better candidate does not run for office, because only in this case, the decision of the worse candidate is relevant; hence, in the basic model, candidate 1 does not need to form an expectation about the probability of candidate 2 running. Here, however, p_2 enters in candidate 1's consideration, because δ must be paid whether or not candidate 2 enters the race. Consequently, if p_2 is high, candidate 1 might not find it attractive to spend the campaign cost δ, because it is very unlikely that he will get elected.

Candidate 2 will run if

$$v_2 + r - \Lambda_1 \geq \delta + c_2$$

where $\Lambda_1 = v_0 + p_1(v_1 - v_0 - \frac{r}{N-1})$. Hence, the probability that candidate 2 runs is

$$p_2 = G_2(v_2 + r - \Lambda_1 - \delta) \tag{18}$$

Two interesting observations can be made. First, the system defined by (17) and (18) may have more than just one solution. To illustrate this, consider the following example.

Example 3. *Let $v_0 = 0$, $v_1 = 1$, $v_2 = 1.2$, $r = 0$ and $\delta = 0.1$. Both candidates' costs are drawn from*

$$c = \begin{cases} 0.5 & \textit{with probability 0.9} \\ 2 & \textit{with probability 0.1} \end{cases}$$

There are two equilibria. In the first one, candidate 1 runs, if he has a low cost realization (i.e., $c_1 = 0.5$), and candidate 2 never runs. In the second equilibrium, the roles are reversed: Candidate 2 runs, if he has a low cost realization, and candidate 1 never runs.

Hence, with the introduction of campaign costs, it is possible that uniqueness of the equilibrium is lost. Note, however, that for sufficiently low δ only the equilibrium in which candidate 1 runs (for small cost realizations) survives, so our equilibrium in the basic model is locally robust to the introduction of small campaign costs.

Second, even if a unique equilibrium exists, an increase in δ may decrease or increase the probability with which the better candidate runs for office. The intuition is similar to the case when a wage increase reduced the probability that the good candidate runs: An increase in δ changes the outside option Λ_1, because it decreases the probability that candidate 1 will run.

To see this, differentiate (17) and (18) totally to yield

$$dp_1 = g_1(\cdot)\left[-\frac{1}{1-p_2}d\delta + \frac{\delta}{(1-p_2)^2}dp_2\right] \quad (19)$$

$$dp_2 = -g_2(\cdot)\left[d\delta + \left(v_1 - v_0 - \frac{r}{N-1}\right)dp_1\right] \quad (20)$$

Consider a marginal increase in δ, starting at $\delta = 0$. Solving the linear equation system yields

$$\frac{dp_2}{d\delta} = g_2(\cdot)\left[g_1(\cdot)\frac{v_1 - v_0 - \frac{r}{N-1}}{1-p_2} - 1\right] \quad (21)$$

Take the setting of Example 2 above, and assume $r = 0$ (or N is large). Then, a sufficient condition for an increase of δ to increase p_2 is that $\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(1-\mu)^2}{2\sigma^2}} > 1$. In particular, if $\mu = 1$, then if $\sigma < 1/\sqrt{2\pi}$ a small increase in δ, starting from $\delta = 0$, increases the probability that the better candidate will run.

4.3 Exclusion of candidates

Our results from the basic model indicate that sometimes the probability of a bad candidate running for office is much higher than the corresponding probability for a good candidate. This suggests that *reducing* the number of people who are "allowed" to volunteer for an unattractive job may be a reasonable strategy for ordinary voters.

This can be true even from an ex ante perspective: Suppose the society has to decide whether to exclude someone from running as a candidate before it knows the realizations of v. Assume, for example, that the v's are iid draws from the following distribution

$$V = \begin{cases} 1+\epsilon & \text{with probability } 1/2 \\ 1.4+\epsilon & \text{with probability } 1/2 \end{cases}, \quad (22)$$

where ϵ is an additional, very small continuous random variable (which makes sure that candidates' v's turn out to be equal only with probability 0).

Assume that $r = 0$, and that the probability of a cost realization that is between 0.3 and 1 is 99 percent, with the remaining probability placed on very high cost realizations. If society admits candidatures of both potential candidates, the ex ante expected value of W is $\frac{1}{4}0.99\cdot1.4+\frac{3}{4}0.99\cdot1 = 1.089$, as the only chance that a candidate with $v = 1.4$ will run is that both candidates have a high realization of v. If the society eliminates one of the candidates ex ante, the expected value of W is $\frac{1}{2}0.99\cdot1.4+\frac{1}{2}0.99\cdot1 = 1.188 > 1.089$.

Hence it might be optimal for a society to design rules that exclude certain possible candidates from running for office. While it is desirable to design rules that exclude candidates that are likely to be bad, this is not a necessary condition for the rule to improve welfare. After all, in the example above, one candidate is (ex ante) as good as the other candidate, and so in this case, an exclusion rule must be arbitrary (i.e., not correlated with expected quality).

Consider the following example from the former university of one of the authors. Every faculty has a promotion and tenure committee, and one of the members on this committee has to come from outside the faculty. Most people consider membership in this committee as not particularly attractive, given that there is no remuneration and the possibilities to trade favors with people from other faculties are very limited. The number of candidates who are willing to run is quite small. Nevertheless, the rules state that only tenured professors are eligible for positions on this committee.

Suppose that the value of the outside committee member rises with his experience, proxied by tenure. If the decision of tenured professors to run for this committee were fixed, admitting nontenured faculty as further candidates could not possibly decrease welfare. Voters would not vote for untenured candidates unless there were no tenured candidate running; but in this case, they would be strictly better off having an unexperienced candidate than having none at all. However, admitting untenured candidates might weaken the incentives of tenured candidates to run.

5 Conclusion

This paper set out to analyze a situation in which being a politician is not a priori an attractive job for every potential candidate, and so the remuneration and strategic interaction with other potential candidates determine which set of candidates is available for the public to choose from. We have analyzed this situation in a new citizen candidate model, in which the candidates differ in their private cost that they would incur as politicians; a candidate's cost parameter is his private information. For every level of remuneration, this game has a unique equilibrium, in which candidate's entry decisions depend on their private cost parameter and their ranking among the other potential candidates in the eyes of the voters.

Under quite general conditions, we have shown that bad candidates are more likely to run for office than good candidates. Even though society just needs *one* good candidate to fill the position, this problem may be acute especially for unattractive positions which every potential candidate would like to leave to another candidate (but would rather fill himself than leave the position open). In these cases, we have demonstrated by example that it is possible that only the least qualified candidate runs. Also, it is possible that a marginal increase in wage might in some circumstances decrease the quality of candidates who run.

Several questions remain for future research. Let us briefly address three issues.

First, we might look at what happens in a model in which not all voters agree on who is the most desirable candidate, for example a standard horizontal differentiation framework. If there are just two potential candidates, then the analysis of the present paper carries over almost completely. The strategic situation of the "extremist" candidate (i.e., the candidate who would lose if both candidates ran for office) is similar to the situation of the worse candidate in our model, as the extremist will only be elected if the moderate candidate does not run; hence, the extremist only needs to decide whether he would like the job given that the other candidate is not willing to run, in order to decide whether he is willing to run. This creates an equilibrium probability that the extremist candidate will run and thus generates the outside option for the more moderate candidate who behaves similar to the better candidate in our model.

If there are more than two potential candidates in a horizontal differentiation framework, the question arises about who would win the election if more than 2 candidates decided to run. Messner and Polborn (2003) consider "robust political equilibria" (a combination of trembling hand perfection and coalition proofness refinements) and show that under runoff rule, the "most moderate" candidate who runs is the unique equilibrium outcome, while there may be multiple equilibrium outcomes under plurality rule. If there are multiple equilibrium outcomes at the election stage, this would complicate considerably the decisions at the entry stage.

A second interesting question would be the endogenous determination of the wage. There are two possible approaches. First, one can look for the remuneration (and possibly campaign cost) scheme that would maximize the utility of an ordinary voter. This is the outcome we would expect if the electorate voted on the salary of the politician. Second, one can look at the salary level chosen if politicians can determine the salary level. Evidently, there need to be some restrictions imposed on politicians that prevent them from taking as much money as possible. One possibility is that politicians can only determine the salary paid in the next period. Caselli and Morelli (2003) have looked at this approach and derived some interesting dynamic results. In their model setup, bad politicians may choose a low salary level in order to prevent good candidates from running and hence being voted out of office.

A third variant of the model arises, if we change the assumption that citizens know the competence of all running candidates. If citizens have imperfect information, there is a scope for campaigning. Campaigning will be the fiercer, the more attractive the position is, and since campaign expenditures are socially wasteful, voters might choose a lower salary level than they would without this problem. Perhaps, a model along these lines can explain why we often see very low salaries for politicians which are difficult to explain as optimal in a perfect information environment.

References

[1] Bergstrom, T, L. Blume and H. Varian (1986), *On the Privat Provision of Public Goods*, Journal of Public Economics 29, 25-49.

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

[3] Bilodeau, M. and A. Slivinski, (1996), *Toilet Cleaning and Department Chairing: Volunteering a Public Service*, Journal of Public Economics, 59, 299-308.

[4] Bliss, C. and Nalebuff, B. (1984), *Dragon-slaying and Ballroom Dancing: The Private Supply of a Public Good*, Journal of Public Economics, 25, 1-12.

[5] Book of the States 2000/2001, vol. 33, Council of State Governments: Lexington, Kentucky.

[6] Caselli, F. and M. Morelli (2003), *Bad Politicians*, Journal of Public Economics, forthcoming.

[7] Carillo, J.D. and T. Mariotti (2003), *Electoral Competition and Politician Turnover*, European Economic Review 45, 1-25.

[8] Diermeier, D., M. Keane and A. Merlo (2002), *A Political Economy Model of Congressional Careers*, mimeo.

[9] Dutta, B., M.O. Jackson, and M. LeBreton (2001), *Voting by Successive Elimination and Strategic Candidacy*, Journal of Economic Theory, forthcoming.

[10] Groseclose, T. and K. Krehbiel (1994), *Golden Parachutes, Rubber Checks, and Strategic Retirements from the 102nd House*, American Journal of Political Science, 38, 75-99.

[11] Hall, R.L. and R.P. Van Houwelling (1995), *Avarice and Ambition in Congress: Representatives' Decisions to Run or Retire from the U.S. House*, American Political Science Review, 89, 121-136.

[12] Messner, M. and M. Polborn (2003), *Robust political equilibria under plurality and runoff rule*, mimeo.

[13] Osborne, M. and A. Slivinski (1996), *A Model of Political Competition with Citizen-Candidates*, Quarterly Journal of Economics, 111, 65-96.

[14] Palfrey, T., H. Rosenthal (1984), *Participation and the Discrete Provision of Public Goods*, Journal of Public Economics 24, 171-193.

[15] Poutvaara, P. and T. Takalo (2002), *Candidate Quality*, mimeo.

6 Appendix – Proof of Proposition 4

First, note that because $0 \leq W(I) \leq \bar{v}$ for all I, the sequence $\{W(I)\}$ must have a converging subsequence. In the following we will show that any converging subsequence must converge to the point $\max\{G(v_1)v_1, \bar{v} - c_{min}\}$.

If $G(v_1)v_1 \geq \bar{v} - c_{min}$, it follows from Proposition 3 that the worst candidate has the optimal strategy to run if $c_1 \leq v_1$, and that no other candidate will ever run. Hence, in this case, $W(I) = G(v_1)v_1$ for all I. For the remainder of the proof, consider the case that $G(v_1)v_1 < \bar{v} - c_{min}$. We show that $\lim_I W(I) = \bar{v} - c_{min}$, by first showing that $\lim_I W(I) \leq \bar{v} - c_{min}$, and then $\lim_I W(I) \geq \bar{v} - c_{min}$.

Step 1: Suppose first that, to the contrary, $\lim_I W(I) > \bar{v} - c_{min}$. Then, there must exist I_0 such that $W(I) > \bar{v} - c_{min}$ for all $I \geq I_0$.

Let $B(I)$ denote the best candidate who runs with positive probability, given that there are I candidates, let $v_B(I)$ denote his quality, and let $p_B(I)$ denote the probability that this candidate runs for office; when no confusion can arise, we will suppress the dependence on I. Conditioning on whether the best candidate runs or not, we have

$$W(I) = p_B v_B + (1 - p_B)\Lambda_{B-1}. \tag{23}$$

Since B is willing to run, we must have

$$v_B - c_{\min} \geq \Lambda_{B-1} \tag{24}$$

Combining (23) and (24), and the initial assumption that $\lim_I W(I) > \bar{v} - c_{min}$, we have

$$\bar{v} - c_{\min} < W(I) \leq v_B - (1 - p_B)c_{\min} \tag{25}$$

for all $I > I_0$. Consequently, there exists $\bar{p} > 0$ such that $p_B(I) \geq \bar{p}$ for all $I \geq I_0$.

Now observe that, for any candidate x, if $p_x = G(v_x - \Lambda_{x-1}) > 0$, then there exists an interval $[v_x - \delta, v_x]$ such that every candidate with a v in this interval runs with a probability greater than $G(v_x - \delta - \Lambda_{x-1}) > 0$, because all these candidates have a lower outside option than candidate x (because Λ is a nondecreasing function by Lemma 1).

As $I \to \infty$, there are infinitely many candidates in an arbitrarily small interval of the form $[v_B(I) - \epsilon, v_B(I))$. Hence, $\Lambda_{B-1} \approx v_B$, because for any $\epsilon > 0$, each of the candidates in this interval runs with a strictly positive probability, so that $\Lambda_{B-1} \geq v_B - \epsilon$. But then, candidate B should not run, because $v_B - c_B \leq v_B - c_{min} < \Lambda_{B-1}$, the desired contradiction that proves that we cannot have $\lim_{I \to \infty} W(I) > \bar{v} - c_{min}$.

Step 2: Now suppose that $\lim_{I \to \infty} W(I) < \bar{v} - c_{min}$. Then there exists $\delta > 0$ and I_2 such that $W(I) < \bar{v} - c_{min} - \delta$ for all $I \geq I_2$. Since $\Lambda_i \leq W(I)$ for all $i \leq I$, it follows that for all candidates in $(\bar{v}-\delta;\bar{v}]$, it must be that $\Lambda_{i-1} < v_i - c_{min}$. Hence, all these candidates will run with positive probability. As I grows, the number of candidates in this interval will grow too and consequently the probability that one of them wins will converge to 1. But this implies that there must be an $\bar{I}$ such that $W(I) > \bar{v} - c_{min} - \delta$ for all $I > \bar{I}$. Since δ can be chosen arbitrarily small, we have the desired contradiction, showing that $\lim_I W(I) \geq \bar{v} - c_{min}$.

Both steps combined show $\lim_{I\to\infty} W(I) = \bar{v} - c_{min}$, as claimed.