Web Appendix for: "Competing on Good Politicians"

Vincenzo Galasso* Bocconi University, Dondena, IGIER & CEPR

Tommaso Nannicini‡ Bocconi University, IGIER & IZA

October 2010

Abstract

This Web Appendix provides the formal proofs of the propositions contained in the paper "Competing on Good Politicians," *American Political Science Review*.

Keywords: political competition, political selection, probabilistic voting.

*Bocconi University, Department of Economics, Via Rontgen 1, 20136 Milan, Italy. Email: vincenzo.galasso@unibocconi.it.

‡Bocconi University, Department of Economics, Via Rontgen 1, 20136 Milan, Italy. Email: tommaso.nannicini@unibocconi.it.

Web Appendix

Proof of Proposition 1

First, define $H\left(\Lambda_i^D\right) = G\left(\lambda_{II}^i\right) - G\left(\lambda_I^i\right)$ as the mass of districts in the interval Λ_i^D. Parties' objective in allocating their experts is to maximize the probability of winning the election, i.e., of winning more than 50% of the districts. Consider party D. Its probability of winning a district k is $\delta < d_k = \rho\left(V_C\left(y_D^k\right) - V_C\left(y_R^k\right)\right) - \lambda_k$. Thus, party D allocates experts to districts in order to modify $V_C\left(y_D^k\right)$ in the marginal districts. These are the district(s) such that, given the shock, winning the district(s) increases the probability of winning the election. For instance, if candidates are matched in every districts, we would have $V_C\left(y_D^k\right) = V_C\left(y_R^k\right)\ \forall k$, and party D probability of winning the election would be $\delta < -\lambda_0$, so that the interval of districts around λ_0 would be pivotal.

Case (I) Both parties have enough experts to span the interval between λ_w and λ_0, i.e., $\mu > \eta/2$. Consider an allocation Λ^L by party D which includes Λ_i^D s.t. $[\lambda_w, \lambda_\Xi] \subset \Lambda_i^D$. It is easy to see that an allocation Λ^R by party R which includes Λ_i^R s.t. $[\lambda_\varepsilon, \lambda_W] \subset \Lambda_i^R$ is a best response to Λ^D. In fact, given Λ^D, by sending its experts to the interval $[\lambda_\varepsilon, \lambda_W]$, party R restores its probability of winning the election to 1/2. In particular, party R wins the election for $\delta > 0$, and D for $\delta < 0$. Allocating additional experts may modify the share of seats won by party R, but not its probability of winning the election. The same reasoning shows that Λ^D with Λ_i^D s.t. $[\lambda_w, \lambda_\Xi] \subset \Lambda_i^D$ is a best response to Λ^R with Λ_i^R s.t. $[\lambda_\varepsilon, \lambda_W] \subset \Lambda_i^R$. Hence, a pair of allocations Λ^D and Λ^R that include (i) Λ_i^D s.t. $[\lambda_w, \lambda_\Xi] \subset \Lambda_i^D$ and $H\left(\Lambda^D\right) = \sum_i H\left(\Lambda_i^D\right) = \mu_D$, and (ii) Λ_i^R s.t. $[\lambda_\varepsilon, \lambda_W] \subset \Lambda_i^R$ and $H\left(\Lambda^R\right) = \sum_i H\left(\Lambda_i^R\right) = \mu_R$ is a Nash equilibrium of the allocation game.

To prove that any equilibrium allocation Λ^D has to include Λ_i^D s.t. $[\lambda_w, \lambda_\Xi] \subset \Lambda_i^D$, consider first an allocation $\widehat{\Lambda}^D$ with $\widehat{\Lambda}_i^D = \left[\lambda_I^i, \lambda_{II}^i\right]$ s.t. $0 > \lambda_I^i > \lambda_w$ and $\lambda_{II}^i > \lambda_\Xi$, and no other experts in $[\lambda_w, \lambda_I]$. Party-R best response is to allocate its experts in $[\lambda_w, \lambda_I] \cup [\lambda_0, \lambda_{II}]$. Following this strategy, party R wins the election with a probability greater than 1/2, since for $\delta = 0$ party R wins all districts with $\lambda > 0$ (and hence 50%), but also the districts in $[\lambda_w, \lambda_I]$. Hence, $\widehat{\Lambda}^D$ cannot be part of an equilibrium since simply matching the previous best response by party R would give party D 50% probability of winning the election. Finally, it is trivial to show that an equilibrium allocation Λ^D has to include the interval $[\lambda_\varepsilon, \lambda_\Xi]$. Consider $\widehat{\Lambda}^D$ with $\widehat{\Lambda}_i^D = \left[\lambda_I^i, \lambda_{II}^i\right] \in [\lambda_w, \lambda_\varepsilon]$ and $\widehat{\Lambda}_j^D = \left[\lambda_I^j, \lambda_{II}^j\right] \in [\lambda_\Xi, \lambda_W]$. Party-$R$ best response would be Λ^R such that $\Lambda_i^R = [\lambda_\varepsilon, \lambda_W]$, which yields party R a winning probability greater than 1/2. Hence, $\widehat{\Lambda}^D$ cannot be part of an equilibrium. It is easy to see that this holds also for $\mu = \eta/2$, in which case parties will allocate experts respectively to $[\lambda_w, \lambda_0]$ for party D and to $[\lambda_0, \lambda_W]$ for party R. Notice also that if $\mu_R = \eta/2$ and $\mu_D > \eta/2$ (or viceversa), the party having more experts will win the election with probability $\Pi_D = \frac{1}{2} + \psi \frac{1-\lambda^C}{\lambda^C} \frac{\eta}{4}$. This is because party D wins the election for population shocks such that $\delta < 0$, but it also ties the elections for $\delta \in [0, W]$, where $W = \frac{1-\lambda^C}{\lambda^C} \frac{\eta}{2}$.

Case (II) One party (say party D) has enough experts to span the crucial interval, but the other does not, i.e., $\mu_D > \eta/2 > \mu_R$. Suppose that party D allocates its experts to $[\lambda_a, \lambda_W]$, as displayed in Figure 1 in the text. What is party-R best response? Party-R does not have enough experts to match party-D experts and re-establish its probability of winning the election to 1/2, but it can reduce party-D probability of winning the election. To see how, consider the largest (positive) realization of the shock, δ, that still allows party-D to win the election, given that party-D has allocated experts as described above, and party-R has not allocated any. Party-R will have to targets with its experts those districts that are marginally in favor of party-D, for this level of the shock. This can be done by sending experts to $[\lambda_j, \lambda_m]$ with $\lambda_j = \max\{\lambda_a, \lambda_w\}$, since it never pays off to send experts outside the interval $[\lambda_w, \lambda_W]$. Moreover, it is trivial to see that for party-R sending experts to $[\lambda_j, \lambda_m]$ with $\lambda_j = \max\{\lambda_a, \lambda_w\}$, party-$D$ best response is to span the interval $[\lambda_a, \lambda_W]$. Hence this allocation constitutes an equilibrium. To see why under this allocation party D wins the election with probability $\Pi_D = \frac{1}{2} + \frac{1-\lambda^C}{2\lambda^C}\psi z$, consider Figure 1 in the text. Party D wins the election when more than 50% of the districts are in its favor; these districts are $\left[-\frac{1-\lambda^c}{2\lambda^c}, -\lambda_m + x\right] \cup [\lambda_m, \lambda_m + x]$, such that $\frac{\lambda^c}{1-\lambda^c}\left[-\lambda_m + x + \frac{1-\lambda^c}{2\lambda^c}\right] + \frac{\lambda^c}{1-\lambda^c}\left[\lambda_m + x - \lambda_m\right] = 1/2$. Hence, $x = \lambda_m/2$, where $\lambda_m = \frac{1-\lambda^c}{\lambda^c}z$, since $G(\lambda_W) - G(\lambda_a) = \mu_D = \mu_R + z$ and $G(\lambda_m) - G(\lambda_a) = \mu_R$. A simple inspection of Figure 6 in the Appendix of the paper shows that all these districts are won by party-D if $\delta < d_x = -x + \lambda_m = \frac{1-\lambda^c}{2\lambda^c}z$, which occurs with probability $\Pi_D = \frac{1}{2} + \frac{1-\lambda^C}{2\lambda^C}\psi z$.

Finally, to see that no other equilibrium allocation is possible, consider party-D allocating experts to $\Lambda^D = [\lambda_w, \lambda_s]$, such that $G(\lambda_s) - G(\lambda_w) = \mu_D$. Party-$R$ would have an incentive to allocate experts to $\Lambda^R = [\lambda_\varepsilon, \lambda_s]$, thereby winning the elections with a probability higher than 50%. But with this allocation by party-R, party-D best response would be to allocate its experts to $[\lambda_a, \lambda_W]$.

Case (III) Parties have equal shares of experts, and are unable to span the crucial districts, $\mu < \eta/2$. Suppose that party D allocates its experts to $[\lambda_0, \lambda_B]$. What is party-R best response? To re-establish its probability of winning the election to 1/2, party R can send its experts to $[\lambda_b, \lambda_0]$. As displayed in Figure 4 in the Appendix of the paper, party D wins the election for $\delta < \max[-\lambda_B, -\lambda_b - W]$, party R for $\delta > \min[-\lambda_b, -\lambda_B + W]$, and the election is tied for $\delta \in [-\lambda_B, -\lambda_b]$. Finally, notice that party R cannot increase its probability of winning the election above 1/2 by allocation experts in other districts (see Figure 7 in the Appendix of the paper). Hence, party-D allocation in $[\lambda_0, \lambda_B]$ and party-R allocation in $[\lambda_b, \lambda_0]$ is an equilibrium.

To prove that no other equilibrium allocation exists, first notice that allocating experts outside the interval $[\lambda_w, \lambda_W]$ is never part of an equilibrium, since it does not modify the probability of winning election, which can instead be achieved by allocating experts in this interval. Consider party-D allocation $\Lambda^D = [\lambda_b, \lambda_0]$. Party-$R$ best response would be to allocate experts to $[\lambda_w, \lambda_b]$, which would yield party R a winning probability above 1/2, since for $\delta = 0$ party R would win in districts with $\lambda > 0$ and in $[\lambda_w, \lambda_b]$. The same reasoning applies to any $\widehat{\Lambda}^D = [\lambda_I, \lambda_{II}]$ s.t. $\lambda_I \in [\lambda_w, \lambda_0)$, $\lambda_{II} \in [\lambda_w, \lambda_b)$ and $G(\lambda_{II}) - G(\lambda_I) = \mu$. And to $\widehat{\Lambda}^D = [\lambda_I, \lambda_W]$ and $G(\lambda_W) - G(\lambda_I) = \mu$.

Case (IV) Parties are unable to span the crucial districts, and have marginally different shares of experts. Suppose that party D, which has few more experts than party R (i.e., $z < (\frac{\eta}{2} - \mu_R)/2$), allocates its experts to $[\lambda_0, \lambda_B]$. What is party-R best response? Having fewer experts, party R is unable to match party-D action with a symmetric allocation (i.e., with $[\lambda_b, \lambda_0]$ as in Figure 8 in the Appendix of the paper) and to restore the probability to win the election to 50%. But it can reduce party-D probability of winning the election. To see how, consider the largest (positive) realization of the shock, δ, that still allows party-D to win the election, given that party-D has allocated experts as described above, and party-R has not allocated any. Party-R will have to targets with its experts those districts that are marginally in favor of party-D, for this level of the shock. This can be done by sending experts to $[\lambda_w, \lambda_g]$, such that $G(\lambda_g) - G(\lambda_w) = \mu_R$ (or alternatively to the right of λ_0). Moreover, notice that for this allocation by party R, party D best response is to allocate its experts to $[\lambda_0, \lambda_B]$ (or alternatively to $[\lambda_w, \lambda_g]$ and the remaining part to the right of λ_0). Hence, this allocation constitutes an equilibrium.

To see why under this allocation party-D wins the election with probability $\Pi_D = \frac{1}{2} + \frac{1-\lambda^C}{\lambda^C}\psi z$, consider again Figure 8. Party-D wins the election when more than 50% of the districts are in its favor; these districts are $\left[-\frac{1-\lambda^c}{2\lambda^c}, \lambda_w\right] \cup [\lambda_g, \lambda_w + x] \cup [\lambda_0, \lambda_B]$, such that $\frac{\lambda^c}{1-\lambda^c}\left[\lambda_w + \frac{1-\lambda^c}{2\lambda^c}\right] + \frac{\lambda^c}{1-\lambda^c}\left[\lambda_w + x - \lambda_g\right] + \frac{\lambda^c}{1-\lambda^c}\left[\lambda_B - \lambda_0\right] = 1/2$. Hence, $x = W - \lambda_B$, where $\lambda_B = \frac{1-\lambda^c}{\lambda^c}\mu_D$. A simple inspection of Figure 8 shows that all these districts are won by party-D if $\delta < d_x = -(x + \lambda_g) = \frac{1-\lambda^c}{\lambda^c} z$, which occurs with probability $\Pi_D = \frac{1}{2} + \frac{1-\lambda^C}{\lambda^C}\psi z$.

To prove that no other equilibrium allocation exists, notice that party D has no incentive to allocate experts anywhere in the interval $[\lambda_w, \lambda_0]$, since party R would best respond by sending experts to the subset of the interval $[\lambda_w, \lambda_0]$, where party D has instead sent loyalists, and would thus win the election with more than 50% probability. Party D sending experts to the interval $[\lambda_z, \lambda_W]$ is not part of an equilibrium either, since, regardless of party R response, party D could always do at least as well by sending them to $[\lambda_0, \lambda_B]$.

Case (V) Parties are unable to span the crucial districts, and have largely different shares of experts. Suppose that party D, which has more experts than party R (i.e., $z > (\frac{\eta}{2} - \mu_R)/2$), allocates its experts to $[\lambda_0, \lambda_B]$. What is party-R best response? In this case, having much fewer experts, party R can only try to reduce party-D probability of winning the election. To see how, consider the largest (positive) realization of the shock, δ, that still allows party-D to win the election, given that party-D has allocated experts as described above, and party-R has not allocated any. Party-R will have to targets with its experts those districts that are marginally in favor of party-D, for this level of the shock. This is easily done by sending its few experts to $[\lambda_0, \lambda_P]$, such that $G(\lambda_P) - G(\lambda_0) = \mu_R$. Moreover, notice that for this allocation by party R, party D is indifferent between allocating its experts to $[\lambda_0, \lambda_B]$ (or alternatively to the right of λ_P). Hence, this allocation constitutes an equilibrium.

To see why under this allocation party D wins the election with probability $\Pi_D = \frac{1}{2} + \frac{1-\lambda^C}{2\lambda^C}\psi(\frac{\eta}{2} - \mu_R)$, consider Figure 9 in the Appendix of the paper. Party-D wins the election when more than 50%

of the districts are in its favor; these districts are $\left[-\frac{1-\lambda^c}{2\lambda^c}, \lambda_w\right] \cup [\lambda_w, \lambda_w + x] \cup [\lambda_P, \lambda_w + x]$, such that $\frac{\lambda^c}{1-\lambda^c}\left[\lambda_w + \frac{1-\lambda^c}{2\lambda^c}\right] + \frac{\lambda^c}{1-\lambda^c}\left[\lambda_w + x - \lambda_w\right] + \frac{\lambda^c}{1-\lambda^c}\left[\lambda_w + x - \lambda_P\right] = 1/2$. Hence, $x = (W + \lambda_P)/2$, where $\lambda_P = \frac{1-\lambda^c}{\lambda^c}\mu_R$. A simple inspection of Figure 6 shows that all these districts are wons by party-D if $\delta < d_x = -(x + \lambda_w) = \frac{1-\lambda^c}{2\lambda^c}(\frac{\eta}{2} - \mu_R)$, which occurs with probability $\Pi_D = \frac{1}{2} + \frac{1-\lambda^C}{2\lambda^C}\psi(\frac{\eta}{2} - \mu_R)$.

To prove that no other equilibrium allocation exists, notice that party D has no incentive to allocate experts anywhere in the interval $[\lambda_w, \lambda_0]$, since party R would best respond by sending experts to the subset of the interval $[\lambda_w, \lambda_0]$, where party D has instead sent loyalists, and would thus win the election with more than 50% probability. Party D sending experts to the interval $[\lambda_z, \lambda_W]$ is not part of an equilibrium either, since, regardless of party R response, party D could always do at least as well by sending them to $[\lambda_0, \lambda_B]$. QED

Proof of Proposition 2

To see that for $\gamma < \frac{1-\lambda^C}{\lambda^C}\psi\Delta v$, the equilibrium share of experts chosen by both parties is $\mu_D = \mu_R = \frac{\eta}{2} + \varepsilon$, let consider each party best response to the other party selection (and allocation) decision.

Suppose that $\mu_R < \frac{\eta}{2}$. What is party D best response? Party D could select and allocate the same share of experts as party R, $\mu_D = \mu_R < \frac{\eta}{2}$. This would lead to a 50% probability of winning the election and, according to equation (4), to a utility $U_L = \Delta v/2 - \gamma\mu_R$. Alternatively, party D could try to increase its probability of winning the election by selecting more experts. For $\mu_R < \mu_D = \mu_R + z < \frac{\eta}{2}$, and $z < (\eta/2 - \mu_R)/2$, the probability of winning the election would become $\Pi_D = \frac{1}{2} + \frac{1-\lambda^C}{\lambda^C}\psi z$ (see case IV in Proposition 1 in the Appendix); and the corresponding utility would be $U_L = \frac{\Delta v}{2} + \frac{\Delta v\left(1-\lambda^C\right)\psi z}{\lambda^C} - \gamma\left(\mu_R + z\right)$. It is straightforward to see that for $\gamma < \frac{1-\lambda^C}{\lambda^C}\psi\Delta v$, party D would prefer to select $\mu_D\left(\mu_R\right) = \mu_R + z = \mu_R + (\eta/2 - \mu_R)/2$ experts to any share of experts $\mu_D \in [0, \eta/2]$. This is because for $\mu_D < \mu_D\left(\mu_R\right)$, the reduction in party-D probability of winning the election outweights the reduction in the cost of the experts; and for $\mu_D > \mu_D\left(\mu_R\right)$, the increase in the cost of the experts is not matched by a corresponding increase in the probability of winning the election (see case V in Proposition 1 in the Appendix). Party D could however choose to select an even larger share of experts, $\mu_D = \mu_R + z > \frac{\eta}{2}$. The utility associated with this strategy is $u_D = \frac{\Delta v}{2} + \frac{\Delta v\left(1-\lambda^C\right)\psi z}{2\lambda^C} - \gamma\left(\mu_R + z\right)$, s.t. $\mu_D = \mu_R + z > \frac{\eta}{2}$. It is easy to see that party D will prefer to select $\mu_D = \mu_R + z = \eta$ rather than $\mu_D = \mu_R < \frac{\eta}{2}$, if $\gamma < \frac{1-\lambda^C}{2\lambda^C}\psi\Delta v$ (and $\mu_D = \mu_R < \frac{\eta}{2}$ otherwise). It thus remains to be seen whether, when $\mu_R < \eta/2$ (and $\gamma < \frac{1-\lambda^C}{2\lambda^C}\psi\Delta v$) party D best response is to select $\mu_D\left(\mu_R\right) = \mu_R + (\eta/2 - \mu_R)/2$ or $\mu_D = \eta$. Comparing the utility provided to party D by these two selection decisions, it is easy to see that party D best response is

$$\mu_D\left(\mu_R\right) = \begin{cases} (\eta/2 + \mu_R)/2 \text{ if } \mu_R < \overline{\mu} \text{ and } \mu_R < \eta/2 \\ \eta \text{ if } \eta/2 > \mu_R > \overline{\mu} \end{cases}$$

where $\overline{\mu} = \frac{\eta}{2}\left(3\gamma - \frac{1-\lambda^C}{2\lambda^C}\psi\Delta v\right)$.

Suppose now that $\mu_R > \frac{\eta}{2}$. What is party D best response? It is straightforward to see that for $\mu_D = \frac{\eta}{2} + \varepsilon$, party D secures a probability of winning the election equal to 50%. Any larger share of experts would thus bring higher costs at no benefit.

Finally, what happens for $\mu_R = \frac{\eta}{2}$. Party D best response will be to select a share of experts equal to $\mu_D = \frac{\eta}{2} + \varepsilon$, which allows party D to win the election with probability $\Pi_D = \frac{1}{2} + \psi \frac{1-\lambda^C}{\lambda^C} \frac{\eta}{4}$.

To summarize, party D best response to party R selection of experts is

$$\mu_D(\mu_R) = \begin{cases} (\eta/2 + \mu_R)/2 \text{ if } \mu_R < \overline{\mu} < \eta/2 \\ \eta \text{ if } \eta/2 > \mu_R > \overline{\mu} \\ \eta/2 + \varepsilon \text{ if } \mu_R = \eta/2 \\ \eta/2 + \varepsilon \text{ if } \mu_R > \eta/2 \end{cases}$$

Due to symmetry, it is straightforward to see that a symmetric reaction function $\mu_R(\mu_D)$ applies to party R. Hence, the only equilibrium selection is $\mu_R = \mu_D = \eta/2 + \varepsilon$. The associated allocation of experts is $[\lambda_w, \lambda_\Xi]$ for party D, and $[\lambda_\varepsilon, \lambda_W]$ for party R, and both parties win the election with equal probabilities, $\Pi_D = \Pi_R = 1/2$ (as in case I at Proposition 1 in the Appendix). Moreover, it is straightforward to see that, since both parties send experts in $[\lambda_\varepsilon, \lambda_\Xi]$, only experts will be elected in these districts. QED

Proof of Proposition 3

Recall that $\eta = \frac{2\lambda^C}{1-\lambda^C} W$. Hence, $\frac{\partial \eta}{\partial \lambda^C} > 0$. And since $\mu_D = \mu_R = \frac{\eta}{2} + \varepsilon$, the share of experts increases with λ^C: $\frac{\partial \mu_D}{\partial \lambda^C} = \frac{\partial \mu_R}{\partial \lambda^C} > 0$. Notice however that an equilibrium with $\mu_D = \mu_R = \frac{\eta}{2} + \varepsilon$ requires the cost of experts to be low: $\gamma < \frac{1-\lambda^C}{\lambda^C} \psi \Delta v$. The right hand side of this inequality is clearly decreasing in λ^C, since the increase in the probability of being elected due to selecting and allocating additional experts is decreasing in λ^C. Thus, as long as $\gamma < \frac{1-\lambda^C}{\lambda^C} \psi \Delta v$ holds, we have $\frac{\partial \mu_R}{\partial \lambda^C} > 0$. But if the increase in λ^C leads to a reversal of the previous inequality, then we would have $\mu_D = \mu_R = 0$. QED