An Electoral Model of Political Dynasties *

Bingyong Zheng† Alberto Batinti‡ Sambuddha Ghosh§ Quan Wen¶

September 23, 2016

Abstract

In this paper we propose an electoral model that studies the interaction between the distribution of voters ideal policies and the success of dynastic politicians. Electoral success of politicians depends on their ability to implement better policies for the voters because campaign promises are empty. Good politicians are more likely to implement the policy that maximizes social welfare. Dynastic politicians are believed to be good or bad depending on the policy implemented by their predecessors while in office. When the mean and median voters’ preferences coincide, good dynastic politicians enjoy a competitive advantage, but still can lose in some situations. However, as the gap between average and median policy preferences widens good dynastic candidates lose their competitive edge, and bad dynasts recover part of their disadvantage. We conclude that democracies with larger political polarization, may show a larger share of bad to good dynastic politicians, but not necessarily a larger share of dynastic politicians overall. We finally extend this model by introducing candidates’ concerns about future generations, and an electorate with long-term memory.

Keywords: Political dynasties, political selection, ideological polarization.

*We thank Pedro Riera Sagrera, Qi Zhang, Jiahua Che, Zhao Chen, Xi Qu, Xing Jing, Fei Shi, Xiaoshu Xu, conference and seminar participants at the 2016 European Political Science Association Conference, Fudan University and Shanghai Jaotong University for comments.

†School of Economics, Shanghai University of Finance & Economics, Shanghai, China. Email: bingyongzheng@gmail.com

‡School of Public Economics and Administration, Shanghai University of Finance & Economics

§School of Economics, Shanghai University of Finance & Economics

¶Department of Economics, University of Washington
“The New York Times reckons that the son of a governor is 6,000 times more likely than the average American male baby-boomer to become a governor himself, and the son of a senator is 8,500 times more likely to become a senator. The concentration of power and wealth in a small elite raises questions about legitimacy. ”

— Dynasties; The power of families, The Economist

1 Introduction

Classical political theory predicts that democracy ought to dispense with political dynasties; and powerful families would fade as ordinary citizens get the vote. Yet, today power still concentrates in families in many countries, including some of the most advanced democracies. A body of empirical research (e.g., Dal Bó et al. 2009, Feinstein 2010) has suggested that a “name recognition” or “brand name advantage” plays an important role in the enduring family power in politics. This view is seldom formalized or developed, however, and we lack an understanding of the mechanism that lends an advantage to dynastic politicians. As a consequence, without an electoral model of political dynasties, we are often unable to determine if powerful families are here to stay, and whether they are a force for good.

In this paper we develop a simple framework to analyze the role played by family name in affecting the electoral success of dynastic politicians. Voters’ policy preferences are exogenous and single-peaked, and their ideal policy depends on the realized state at the time an election is held. Candidates are of two types, low and high, in terms of quality. They have private incentives for running, and care about the well-being of the citizens as well. A candidate learns about her type after the election and, in addition, exerts a costly effort to find out the true state and implement the optimal policy. Voters do not know the type of any of the candidates. Moreover, there is positive correlation of types between family members, so voters may have some information about a dynastic candidate whose ancestor was in office before.

Thus, the name recognition of a dynastic candidate comes from the information’s
revelation carried out with the family’s political history. Moreover, we obtain inter-
esting results when assuming that the probability of being high quality for a dynastic
politician whose ancestor implemented a suboptimal policy will be below the uncondi-
tional average in the population, and above this level if the ancestor implemented
the optimal policy. From this point of view, this paper is also close to models about
the selections of good and bad politicians as in Caselli and Morelli (2004). We apply
this idea of political selection in the realm of dynastic politics, providing a novel in-
vestigation of the underlying factors that promote the success of bad dynasts or good
ones.

The baseline and static versions of this model produce two sets of predictions de-
pending on the distribution of policy preferences. When voters’ preferences are sym-
metrically distributed a candidate whose ancestor has performed well in office, and
therefore deemed more likely to be a good politician, wins the election when opposed
to a non dynastic one. Vice versa, a dynastic candidate who is believed to be more
likely to be of low quality will lose elections. Second, and more important, when
voters’ policy preferences are not symmetrically distributed, a high quality dynastic
candidate will not win in all circumstances even if voters believe she is more likely to
implement the socially optimal policy. As a result, having a good family name does
not always help and can be a disadvantage for a dynastic candidate sometimes. This
happens when the median voter prefers the alternative policy to the socially optimal
policy. Similarly, a low quality dynastic candidate will still have some chances of win-
ning. Moreover, the competitive advantage and disadvantage in these two cases will
wear out as the distance between the median and average policy preferences widens.
As this gap reaches its maximum, the two probabilities of winning for the two types
of dynastic candidates will converge to $\frac{1}{2}$ from above and below respectively. In other
terms, we obtain a novel insight from this model.

We provide a theory explaining both the overall electoral success of dynastic politi-
cians and their composition in terms of their quality. We show that, depending on the
polarization of voters’ preferences, electoral chances for high or low quality dynas-
tic candidates tend to be similar. Intuitively, only the first part of this conclusion,
regarding the high quality dynastic politician, holds when additional assumptions about entry costs are added so that running for a political office is rational only for high quality dynastic politicians and non dynastic ones. The selection bias of dynastic politicians are explored in this case as well. We show that, in this case, only dynastic candidates who are expected to do well because of their inherited reputation are observed by econometricians. Our result therefore suggests that not controlling for this self-selection effect would have a tendency to overestimate the success rate of dynastic candidates. To correct for this bias, future studies are needed to take the self-selection effect into account.

We also extend our basic model to study reputation and long run memory effects and show that, under our plausible assumptions, the main results obtained in the simpler model are robust. By doing so, we can explore an important channel for reputation building in dynasties. When candidates care about the political success of their descendants they will exert higher effort on average to ensure good performance independently on being low or high quality types. In this sense, a political dynasty can be efficient because reputation concerns push politicians to higher levels of effort, which in turn tends to benefit voters’ welfare.

Overall, also in the case where entry costs push out dynastic politicians likely to be of low type from running for political offices, one dynasty may cease existing anyway. A dynasty can in fact end either because the descendant is a bad politician or because one family member had bad luck even if her descendant could be a good politician. Therefore, in addition to explaining the stylized fact that dynastic candidates are much more likely to win elections than non-dynastic candidates, our model also provides a rationale that, in periods of great social and economic changes, dynastic candidates usually find themselves facing significant headwind in seeking political offices. For example, in the 2015–2016 US Republican primaries, voters deserted dynastic candidates in favor of candidates who sometimes have no experience in politics. The once front-runner Jeb Bush was relegated to insignificance soon after the political novice Donald Trump “entered into the fray” and had to quit early in the race. On the Democratic Party side, Hillary Clinton was once viewed as the inevitable
candidate, who none of the party’s establishment wished to challenge. Yet, she faced strong challenge from the once-deemed “fringe candidate” Bernie Sanders, and found herself fighting hard to survive and eventually win.

In the remainder of the paper, we outline prior literature in Section 2. We then describe our model in Section 3 and analyze the baseline model in Section 4. We extend our baseline model in Section 5 to consider candidates’ decision whether to run for office. In Section 6 we investigate the effect of reputational concern on politicians’ behavior. Section 7 discusses long run memory and some other factors. We then conclude in Section 8.

2 Literature

Political dynasties have been attracting social scientists’ attention and concerns at least since Mosca (1966), Michels (1915), and Pareto (1968). Dynastic candidates enjoy a competitive advantage in electoral races and are more likely to be succeeded by their relatives in political offices. Dynasties and dynastic politics are persistent and produce long term effects. Dal Bó et al. (2009) study US Congress data from 1789 to 1996 and suggest that local political connections and name recognition play an important role for setting political dynasties. They also find that, once established, political dynasties are persistent and thus power begets power. Feinstein (2010) explores data for the 1994-2005 U.S. Congress including also losing candidates, and finds that dynastic politicians enjoying “brand name advantages” possess a competitive edge over non-dynastic politicians. Past experience and the capacity to raise funds for political campaigning do not seem instead to have a significant impact on the probability of being elected. More recently, similar tests and results have been conducted for other countries, as in Rossi (2011) for Argentina, Daniele (2015) for Italy, Querubin (2013) for the Philippines.

There is a body of research exploring the socio-economic consequences of political dynasties. Daniele and Geys (2014) find evidence of a progressive reduction of dynasts’ human capital in terms of lower level of education in Italian municipalities.
Moreover, the dynastic advantage is not uniquely a political one. Folke et al. (2016) investigate data on Swedish municipalities and find that a parent occupying a high rank in the executive branch of the municipal government increases her children’s earnings by about 15 percent, children of politicians start later their tertiary education because they find earlier and more easily a job within the municipality. Asako et al. (2015) investigate the economic effects of political dynasties by using Japanese data from 1997 to 2007. Japanese dynasts enjoy higher electoral success and are better at channeling transfers to their own districts. On the other hand, the economic performance is more limited in dynastic districts. Similar results are in Bragança et al. (2015), who explore data on Brazilian municipalities finding that polities with higher rates of dynasts experience larger public spending, especially in infrastructures, but also that such higher levels of spending are not associated with better health, higher levels of education, or economic performance.

However, the presence of dynasties has positive sides. Dynastic politics creates incentives for politicians to be more forward looking and caring for the long-run, and limits moral hazard (Olson 2000). Besley and Reynal-Querol (2013) find that dynastic selection can play a role in improving economic performance when institutions for controlling politicians are weak and policy-making skills are persistent within a dynasty. Selected dynasties will survive only when their economic performance is strong enough and for this selection process better economic performance is observed among countries with dynastic leaders. Crowley and Reece (2013) explore the impact of dynastic governors in the US states for the 1950-2005 period; their results indicate that opportunistic behavior of incumbents in their last mandate is mitigated by dynasts’ concerns for the political success of their offspring. Dynastic governors thus increase incumbents’ accountability. Rivera (2015) uses electoral and political data during Victorian Britain and shows how parties rely on dynasts when their local organizations are weak.
3 Basic model

Time is discrete and the horizon is infinite. Every period there is a continuum of voters of measure one, and a number of potential candidates running for the office of governor. Voters as well as candidates live for one period only and exit at the end of the period. They are subsequently replaced by the same measure of voters and new candidates.

VOTERS: At the beginning of each period $t$ a new generation of voters of measure 1 enters. Upon entering a new voter meets an old voter with a small probability $\eta > 0$ who informs her about the policy $p(t-1)$ that was implemented in period $t-1$. We call a voter who knows the history $p(t-1)$ an informed voter while one who does not an uninformed voter. As a consequence, informed voters in period $t$ know what happened in period $t-1$, but do not know the history in period $t-2$ and before.\footnote{Allowing voters to observe the whole history does not affect our results, but makes the analysis more complicated. In that case we would need to worry about two or more dynastic candidates competing against each other, and voters will try to figure out which one is more likely to be the right candidate to vote for. We will relax this assumption and consider a more general case in Section 7.}

Voters are heterogeneous, with their most preferred policy depending on $\theta_i$, which follows a continuous distribution $\Phi(\cdot)$ over the unit interval. In addition, an individual’s preference over policies is a function of the state of the world. In particular, an individual has the payoff function

$$u_i = -(x - \theta_i)^2$$

where $x$ is the realized policy outcome. And the policy outcome is determined by the policy $p$ adopted and the state of the world $s$:

$$x = p - s.$$ 

The state of the world $s$ is drawn from a symmetric CDF, $F(s)$, over the interval $[-\lambda, \lambda]$. We assume that in addition to the knowledge of the history $p(t-1)$, an informed voter also knows the realized state $s$ in period $t$. But an uninformed voter does not observe $s$. 

$\textit{\footnote{\textcentering{}}}$.\textsuperscript{1}
**POLITICAL CANDIDATE:** In every period \( t \), \( n_t \) number candidates run for the office of governor. One candidate eventually becomes governor through an election with a two-round voting system.\(^2\)

Candidates are observationally identical except for one candidate, candidate 1, whose last name is recognized by voters because her ancestor was in office in period \( t - 1 \). Note that we will use the terms “ancestor” and “descendant” in a rather loose way. Ancestor is used to represent a member of the family who was in office in period \( t - 1 \), and descendant denotes another family member who may run for office in period \( t \). For this basic model, a candidate whose ancestor was in office in period \( t - 1 \) has a name recognition and is called dynastic candidate, while one whose ancestor was not in office in period \( t - 1 \) is a non-dynastic candidate.

In addition, candidates can be one of two types, good candidates are type-H and type-L is reserved for low quality candidates (we will explain the quality difference momentarily). The type of a candidate is unknown to anyone, including the candidate herself, and becomes known to the candidate only after the election. This assumption can be motivated by the fact that the campaign itself is a big test for the candidates, and one never knows how good or bad she is at the job without going through the long stressful process. But during the campaign process, candidates gradually learn about their own types by evaluating their performances in fund raising events, policy debates, town hall meetings, etc. However, it is common knowledge that each candidate is of type-H with probability \( \alpha \) and of type-L with probability \( 1 - \alpha \). Furthermore, if a candidate’s ancestor is known to be of type \( k \), the candidate is more likely to be of type \( k \) as well. Denoting by \( \tau_{j}^{t-1} \) the type of candidate \( j \)’s ancestor, we assume:

\[
\text{Prob}(\tau_{j}^{t} = H) = \gamma \alpha + (1 - \gamma) \text{Prob}(\tau_{j}^{t-1} = H),
\]

where \( \gamma \in (0, 1) \). This can be motivated by the argument that family characteristics and human capital may transfer between family members.

\(^2\)A two-round system is a voting system used to select a single winner in which each voter casts a single vote for her preferred candidate. If no candidate receives more than 50% of the votes in the first-round, a ballot round between the first two candidates decides the winner.
Informed voters learn $p(t-1) \& s$

All voters vote

Governor learns $k \&$ chooses $(e^*, p(s))$

Outcome $x$

After the election, the winning candidate becomes the governor and chooses her effort $e \in [0, 1)$ to find out the true state of the world and implement a policy accordingly. We assume that the probability of the state revealed to her is simply $e$. In choosing her effort to determine the true state, the governor incurs a cost of $C_k(e)$ depending on her type. We assume $C'_k(0) = C''_k(0) = 0$ and, for $e > 0$ and for $k \in \{H, L\}$:

$$C'_k(e) > 0, \quad C''_k(e) > 0, \quad C'_k(1) = \infty.$$  

Compared to a type-L candidate, a type-H candidate has a lower marginal cost of effort,

$$C'_H(e) < C'_L(e) \quad \forall \ e \in (0, 1).$$

Voting is costless and voters care only about their own payoffs. Each voter votes for the candidate from which he has the highest expected payoff. Voters who are indifferent between two or more candidates will randomly cast a vote for those candidates. We summarize the timing of events in Figure 1. The governor receives a payoff of $W$ from holding the office. She also cares about the well-being of citizens. In particular, social welfare enters into her payoff function. The expected payoff for a governor of type-$k$ is:

$$U^k = - \left[ c \int_0^1 (p - s - \theta_i)^2 d\Phi(\theta) + (1 - e) \int_{-\lambda}^{\lambda} (p - s - \theta_i)^2 dF(s) d\Phi(\theta) \right] + W - C_k(e).$$
Clearly, the optimal policy the governor will implement depends upon whether she learns the true state or not.

**Lemma 1.** The optimal policy is \( p^* = s + \bar{\theta} \), where \( \bar{\theta} \equiv E[\theta] \), if the governor knows the true state \( s \); and the optimal policy is \( p^e = \bar{\theta} \) otherwise.

**Proof.** First, when \( s \) is known, the optimization problem facing the governor becomes:

\[
\max_p U_j(s) = W - \int_0^1 (p - s - \theta_i)^2 d\Phi(\theta_i).
\]

The first-order condition gives:

\[
\int_0^1 (p^* - s - \theta_i) d\Phi(\theta_i) = 0.
\]

Hence, we have \( p^* = s + \bar{\theta} \).

Next, ex ante, the expected value of the state is zero and the optimal policy therefore should be the mean of \( \theta_i \), that is, \( p^e = \bar{\theta} \). \( \Box \)

Lemma 1 indicates that the expected payoff for the governor becomes

\[
U^k = W - \left[ \sigma_\theta^2 + (1 - e)\sigma_s^2 \right] - C_k(e),
\]

(2)

where \( \sigma_\theta^2 \) and \( \sigma_s^2 \) denote, respectively, the variances of \( \theta \) and \( s \). To see this, note that

\[
\int_0^1 (p^* - s - \theta_i)^2 d\Phi(\theta_i) = \int_0^1 (\bar{\theta} - \theta_i)^2 d\Phi(\theta_i) = \sigma_\theta^2.
\]

And

\[
\int_0^1 \int_{-\lambda}^\lambda (p^e - s - \theta_i)^2 dF(s) d\Phi(\theta) = \int_0^1 \int_{-\lambda}^\lambda [(\bar{\theta} - \theta_i)^2 - 2s(\bar{\theta} - \theta_i) + s^2] dF(s) d\Phi(\theta) = \sigma_\theta^2 + \sigma_s^2.
\]

The expression in equation (2) lends itself to an intuitive interpretation. The role of \( W \) may represent the ego-rent from occupying the desired political office. The role covered by \( C_k(e) \) has been explained before. The expression containing \( \sigma_\theta^2 \) and \( (1 - e)\sigma_s^2 \) shows that, because voters’ aggregate welfare enters politician’s utility function, the
more the policy positions are dispersed the more the policy choice will reduce the utility of those voters whose preferred position is far from the one implemented. The second term in (2) means that extreme values of the realized state imply large welfare costs when the candidate sets towards the “easy” policy $p^e$. However by exerting more effort, the governor increases the probability of implementing the policy that is socially optimal given the state of the world. At the limit, when the effort tends towards its maximum value, the cost for not implementing the optimal policy will tend to zero.

Note also that voters know about the interdependence of aggregate utility and the winning candidate’s utility. This simplifies the analysis in that policy announcement of implementing any other policy than $p^*(s)$ or $p^e$ would not be considered credible. Suppose we were to assume that, like voters, each candidate may know the true state $s$ with a small probability $\eta$ and also let candidates announce policies before voters cast their votes. A candidate who knows the state $s$ could announce that, if elected, she will implement $p^m(s)$, the policy preferred by the median voter. However, this will not win her the election because voters know that ex post, because of her utility function, $p^*(s)$ and $p^e$ are the only possible policy implementations.

On a related note, an alternative model, in which each candidate may know the true state with probability $\eta$, would still produce similar results as the simple model does. In the alternative model, a candidate who knows the true state may want to signal her type by announcing a different policy platform from other uninformed candidates. Nevertheless, without any effective means to prevent uninformed candidates from mimicking her policy platform, an informed candidate would not be able to credibly convince voters that she knows the true state while others do not. Hence, adding this extra feature does not change voters’ behavior. For simplicity we therefore assume that no candidate knows the state of the world, so there is no policy announcement stage.

In addition to the current payoff $U^k$, the governor may care about the effect of policy outcome on the prospect of her descendant being elected in future, provided she has any, and that her descendant may run in the next period. Denote by $\beta$ the
common discount factor of all candidates and $\mathcal{P}$ the probability that one’s descendant will run for office in period $t + 1$. To simplify notation, let $\delta = \beta \mathcal{P}$. And also let $\bar{V}_{t+1}$ be the ex ante expected payoff of the descendant if policy $p^*$ is successfully implemented and $V_{t+1}$ be the expected payoff if $p^c$ is implemented. Thus, we have the discounted payoff of the governor of type $k$ as:

$$V^k_t = W - [\sigma^2 + (1 - e)\sigma^2_s] - C_k(e) + \delta e\bar{V}_{t+1} + \delta(1 - e)\bar{V}_{t+1}.$$ 

4 No future concern: $\delta = 0$

In this section we consider the special case of $\delta = 0$. Of course, this happens only when the governor expects that none of her descendants would run for office in the future or simply because she did not care about their payoffs. This restriction makes the model simpler to explore, and allows us to focus on voting behavior by analyzing how voters’ vote differently in different states. Meanwhile, this case also provides a benchmark that we can compare to when the governor does care about the political success of her descendants should they choose to run for a political office.

Throughout this section, the equilibrium concept we use is subgame perfect equilibrium. We use the backward induction approach to solve the equilibrium. We first consider the governor’s optimal effort choice given her type. We then analyze voters’ choices, in particular the choices of informed voters, who are informed of history $p(t - 1)$ and the state $s$ and update their beliefs about a candidate’s type.

After winning the election, the governor optimally chooses effort $e$ to find out the state $s$ given her type. She then decides the policy to be implemented depending on whether she learns the true state or not.

**Proposition 1.** The governor of type-$H$ exerts higher effort than one of type-$L$ does, $e^*_H > e^*_L$. Thus, $p^*$ is more likely if a type-$H$ wins the election than if a type-$L$ does.

**Proof.** The optimal effort $e^*$ solves the following maximization problem:

$$\max_{e} U^k = W - [\sigma^2 + (1 - e)\sigma^2_s] - C_k(e).$$
Maximizing with respect to $e$ gives the first-order condition:

$$\sigma_s^2 - C_k'(e^*) = 0.$$  

So $e^*_k = C_k^{-1}(\sigma_s^2)$ where $C_k^{-1}(\cdot)$ is the inverse function of $C_k'(e)$. By the assumption on marginal costs, we conclude that

$$e^*_L < e^*_H,$$

and hence policy $p^*$ is more likely to be implemented when a type-H is elected than a type-L is.

Given the prior that the governor in period $t$ is of type-H with probability $\alpha_t$, the ex post belief that she is of type-H equals

$$Prob(\tau_j^t = H|p^*) = \frac{\alpha_t e^*_H}{\alpha_t e^*_H + (1 - \alpha_t) e^*_L}$$

if $p^*$ is implemented. But it equals

$$Prob(\tau_j^t = H|p') = \frac{\alpha_t(1 - e^*_H)}{\alpha_t(1 - e^*_H) + (1 - \alpha_t)(1 - e^*_L)}$$

if any policy $p'$, where $p' \neq p^*$, is implemented. This, of course, includes the case of $p' = p^e$. Hence, we assume that informed voters treat any off-the-equilibrium policy the same as $p^e$. We also make the following assumption:

**Assumption 1.** The parameters $\gamma$ and $C_k'$ are such that:

$$\frac{\gamma \alpha C_k'^{-1}(\sigma_s^2)}{\gamma \alpha C_k'^{-1}(\sigma_s^2) + (1 - \gamma \alpha) C_L'^{-1}(\sigma_s^2)} > \alpha,$$  

(3)

$$\frac{(\gamma \alpha + 1 - \gamma)(1 - C_H'^{-1}(\sigma_s^2))}{(\gamma \alpha + 1 - \gamma)(1 - C_H'^{-1}(\sigma_s^2)) + \gamma (1 - \alpha)(1 - C_L'^{-1}(\sigma_s^2))} < \alpha.$$  

(4)

Under Assumption 1, implementing $p^*$ leads to a posterior belief that the governor is of type-H with probability greater than $\alpha$ while implementing $p^e$ leads to a posterior belief that the governor is of type-H with probability less than $\alpha$.

To see this, note that the worst prior value $\alpha(t)$ can take is $\gamma \alpha$ and the most favorable prior value $\alpha_t$ can take is $\gamma \alpha + (1 - \gamma)$ given the transition of types in equation (1). And $C_k'^{-1}(\sigma_s^2)$ ($k = H, L$) is the equilibrium effort choices of type-$k$. This assumption
means that voters’ beliefs strongly react to observing the policy implemented in the previous period.

One may wonder how restrictive this assumption is. First, note that $C_k' - 1(\sigma_k^2) \equiv e_k^*$ and $e_H^* > e_L^*$. Thus, at the limit when $\gamma = 1$, i.e., correlation of types across family members is zero, the two conditions, (3) and (4) hold for sure.\(^3\) But a continuity argument would imply that for some $\gamma < 1$, the conditions would still hold. It is not hard to see that as the cost difference between high and low type increases, and thus, the difference in equilibrium efforts increases, the smaller $\gamma$ is needed for the conditions to hold.

**Lemma 2.** Uninformed voters randomly vote for a candidate.

*Proof.* Note that uninformed voters do not know the history $p(t-1)$. Thus, to an uninformed voter, a dynastic candidate and a non-dynastic candidate would be equally likely to be of type-H, that is, both dynastic candidate and non-dynastic candidate are of type-H with probability $\alpha$. In this case, an uninformed voter would be indifferent between a dynastic candidate and any non-dynastic candidate, and therefore vote randomly.

As a consequence, each candidate receives an equal share of the votes of the uninformed voters and the election is determined by informed voters. Also because of this, our results do not depend on the share of informed voters $\eta$ among the population of voters. In the following discussion, we therefore can look at the informed voters’ choices only.

**4.1 Symmetric distribution**

Here we assume that voters’ preferences follow a symmetric distribution. In this case, $\bar{\theta} = \theta^m = \frac{1}{2}$, where $\theta^m$ denotes the median voter’s preference. Because the mean and median preferences are the same, the majority of informed voters vote for the

\[^3\]In this case, condition (3) becomes $\alpha e_H^*/(\alpha e_H^* + (1 - \alpha)e_L^*) > \alpha$ while (4) becomes $\alpha(1 - e_H^*)/(\alpha(1 - e_H^*) + (1 - \alpha)(1 - e_L^*)) < \alpha$, which must be true since $e_H^* > e_L^*$.\]
candidate most likely to implement policy $p^*$. This is so because the optimal policy $p^*$ that maximizes social welfare is also the policy preferred by the median voter.

To see this, note that if the state $s > 0$, then all informed voters with $\theta_i \geq \frac{1-s}{2}$ would prefer $p^*$ to $p^e$ while the rest would rather see $p^e$ implemented. Whereas, if $s < 0$, all voters with $\theta_i \leq \frac{1-s}{2}$ prefer $p^*$ to $p^e$ while the rest prefer $p^e$. As a consequence, candidates considered more likely to implement $p^*$ would receive $\max\{1 - \Phi(\frac{1-s}{2}), \Phi(\frac{1-s}{2})\}$ share of votes of informed voters, and candidates less likely to implement $p^*$ would receive the rest of the votes.

**Proposition 2.** When a dynastic candidate runs against non-dynastic candidates, the outcome of the election depends on history: (i) the dynastic candidate wins if $p(t-1) = p^*$; (ii) a non-dynastic candidate wins otherwise.

**Proof.** Note that conditional on the governor successfully implementing policy $p^*$ in $t-1$, all voters would expect the dynastic candidate to be more likely to be of type-H than a non-dynastic candidate. In this case, because $\text{Prob}(\tau_j^{t-1} = H|p^*) > \alpha$, the probability that the dynastic candidate is of type-H is:

$$\alpha_t(p^*) = \gamma \alpha + (1 - \gamma) \text{Prob}(\tau_j^{t-1} = H|p^*) > \alpha.$$

The expected effort choice of the dynastic candidate after winning election is $e_1 = \alpha_t(p^*)e^*_H + (1 - \alpha_t(p^*))e^*_L$, and is higher than the expected effort of the non-dynastic candidate $e_2$, which equals $\alpha e^*_H + (1 - \alpha)e^*_L$. Thus, in the eyes of informed voters, the dynastic candidate is more likely to implement $p^*$, the policy of preferred by the majority of informed voters. As a result, she wins the election.

Following a similar reasoning, if her ancestor implemented $p^e$, then the probability that a dynastic candidate is of type-H falls below $\alpha$. Hence, she is less likely to implement $p^*$ than even a non-dynastic candidate. Consequently, the majority of the voters will vote for a non-dynastic candidate. This, of course, has to do with the result that the majority of informed voters prefer a non-dynastic candidate to a dynastic candidate. With the two-round voting system a non-dynastic candidate will win the election.

□
With symmetrically distributed policy preferences, the outcome is one-sided regardless of the state. The candidate deemed to be more likely of type-H always wins.

4.2 Asymmetric distributions

In real world, the median and mean of voters’ preferences may not be identical. Below we consider the case of an asymmetric distribution $\Phi(\theta)$. Recall that $\bar{\theta}$ and $\theta^m$ denote, respectively, the mean and the median of the distribution. We assume that $\bar{\theta} \neq \theta^m$ and that $|\bar{\theta} - \theta^m| < \lambda$. In this case, the dynastic candidate wins in some states but loses in the others. The reason for this is that, as long as the mean and the median policy positions differ, the socially optimal policy is not necessarily a winning policy. As a result, depending on the state, the fortune of the dynastic and non-dynastic candidates differs.

Recall that the payoff for a voter with preference $\theta_i$ from policy $p^*$ is $-(\bar{\theta} - \theta_i)^2$, whereas the payoff from $p$ is $-(\bar{\theta} - s - \theta_i)^2$. Consequently, voter $i$ will prefer $p^*$ to $p$ if and only if

$$(\bar{\theta} - \theta_i)^2 < (\bar{\theta} - s - \theta_i)^2$$

and will prefer policy $p$ to $p^*$ otherwise.

**Lemma 3.** Suppose $\bar{\theta} > \theta^m$, $p^*$ is the winning policy when $s \in [-\lambda, 0] \cup [2(\bar{\theta} - \theta^m), \lambda]$. Suppose $\bar{\theta} < \theta^m$, $p^*$ is the winning policy when $s \in [-\lambda, 2(\bar{\theta} - \theta^m)] \cup [0, \lambda]$.

**Proof.** To show the result we need to look at the problem faced by the median informed voter. First, we already know from the previous discussion that median informed voter votes for a candidate more likely to implement $p^*$ if and only if

$$(\bar{\theta} - \theta^m)^2 < (\bar{\theta} - s - \theta^m)^2.$$  \hspace{1cm} (5)

The reason we focus on the median voter is that, provided (5) holds, $p^*$ will be the preferred policy for the majority of informed voters. To see this, note that if (5) is true and $\bar{\theta} - \theta^m > 0$, then for all $\theta_i < \theta^m$, it is necessarily true that

$$(\bar{\theta} - \theta_i)^2 < (\bar{\theta} - s - \theta_i)^2.$$
Similarly, if (5) is true and \( \bar{\theta} - \theta^m < 0 \), then for all \( \theta_i > \theta^m \),

\[
(\bar{\theta} - \theta_i)^2 < (\bar{\theta} - s - \theta_i)^2.
\]

Next, suppose \( \bar{\theta} > \theta^m \), inequality (5) holds either when \( s \leq 0 \) or when \( s \geq 2(\bar{\theta} - \theta^m) \). Suppose \( \bar{\theta} < \theta^m \), the inequality holds either when \( s \geq 0 \) or when \( s \leq 2(\bar{\theta} - \theta^m) \). \( \square \)

A dynastic candidate with \( \alpha_t(p^*) \) is more likely to implement policy \( p^* \) while a non-dynastic candidate is more likely to implement policy \( p^f \). An immediate result that follows is that, ex ante before the state is revealed, the probability that a dynastic candidate wins is always greater than a non-dynastic candidate conditional on history \( p(t - 1) = p^* \).

**Proposition 3.** If the history is \( p(t - 1) = p^* \), ex ante, the dynastic candidate wins the election with probability \( \frac{3}{2} - F(2|\bar{\theta} - \theta^m|) \) > \( \frac{1}{2} \) while a non-dynastic candidate wins with probability \( F(2|\bar{\theta} - \theta^m|) - \frac{1}{2} \in (0, \frac{1}{2}) \).

**Proof.** We first consider the case of \( \bar{\theta} > \theta^m \). As previously shown, in this case, a dynastic candidate wins when \( s \leq 0 \) or when \( s \geq 2(\bar{\theta} - \theta^m) \). Given the assumption that \( F(s) \) is symmetric around zero, the probability that the dynastic candidate wins equals

\[
1 - \left[ F((2(\bar{\theta} - \theta^m)) - \frac{1}{2}) \right] = \frac{3}{2} - F(2(\bar{\theta} - \theta^m)).
\]

Next, if \( \bar{\theta} < \theta^m \), the dynastic candidate wins either when \( s \geq 0 \) or when \( s \leq 2(\bar{\theta} - \theta^m) \). The assumption of symmetric distribution of \( F(s) \) implies that the probability of dynastic candidate winning equals \( \frac{3}{2} - F(2(\theta^m - \bar{\theta})) \). \( \square \)

Before proceeding it is worthwhile pointing out that the probability of non-dynastic winning \( F(2|\bar{\theta} - \theta^m|) - \frac{1}{2} \) lies in the interval between 0 and \( \frac{1}{2} \), and is increasing in the difference \( |\bar{\theta} - \theta^m| \). By contrast, the probability that the dynastic candidate wins lies in the interval between \( \frac{1}{2} \) and 1 and is decreasing in the difference \( |\bar{\theta} - \theta^m| \). In the special case of \( \bar{\theta} = \theta^m \), as the symmetric distribution case in Section 4.1, \( F(0) = \frac{1}{2} \) and a dynastic candidate wins with probability one. On the other hand, as \( |\bar{\theta} - \theta^m| \) increases and approaches \( \lambda \), the probability that the dynastic candidate wins approaches \( \frac{1}{2} \).
In a similar vein, even if one’s ancestor performed badly while in office, the chance of a dynastic candidate winning is not zero. We summarize this in the following result.

**Proposition 4.** If the history is \( p(t - 1) = p^f \), ex ante, the dynastic candidate wins the election with probability \( F(2|\bar{\theta} - \theta^m|) - \frac{1}{2} > 0 \).

Given the history \( p(t - 1) = p^f \), the dynastic candidate is of type-H with probability less than \( \alpha \). Therefore, she is a “worse” candidate than a non-dynastic candidate. Yet, in some states, informed median voters would prefer \( p^f \) to \( p^* \) and, thus, would vote for the dynastic candidate whom they know is less likely to implement the socially optimal policy \( p^* \) than any unknown non-dynastic candidate.

**Corollary 1.** A shift in distribution \( \Phi(\theta) \) that expands the gap between \( \bar{\theta} \) and \( \theta^m \) lowers the probability of a dynastic candidate winning if the history is \( p(t - 1) = p^* \), and increases her chance of winning if the history is \( p(t - 1) = p^f \).

Given the results on both the symmetric and asymmetric distribution cases, it is therefore not surprising to observe that in real world, dynastic candidates are more likely to win elections when facing non-dynastic candidates. In the mean time, it is important to note that, with asymmetric distribution, even if one’s ancestor performed well in office, a dynastic candidate does not win for sure, and she may lose even if she is deemed as a “better candidate,” i.e., more capable and with a socially better policy, than a non-dynastic candidate.

This, helps to explain what happens to Jeb Bush in the US Republican primary in 2016. As a former governor with a record of social conservatism and pro-business policies in a big swing state, Jeb Bush was once regarded as the ideal Republican candidate and expected to dominate the contest. For example, in an Economist article (June 13, 2015), “Were he not called Bush, Jeb—the former governor of Florida—might now enjoy the aura of inevitability for Republicans that Hillary Clinton has among Democrats.” However, the Republican primary voters were so disillusioned with the political establishment after years of stagnant wage growth and increasing inequality,
that they would rather choose an outsider. Consequently he was soon overshadowed by the politically inexperienced Donald Trump and had to quit early in the race.

5 Entry decision

In the previous section we take the decision of whether to run for office as given and do not consider candidates’ entry choices. Here we explore such an extension and endogenize their entry decision. With entry decision, the timing of events changes to the one in Figure 2.

At the beginning of a period, each potential candidate $j \in \{1, \cdots, N\}$ decides whether to run for political office or not. To run for office, each candidate incurs a fixed entry cost $\xi > 0$ that is the same for all candidates. This cost can be interpreted as an opportunity cost of entering politics instead of choosing another career. A candidate will run only if her net payoff, which equals the expected payoff from running minus the entry cost, is above zero. We assume that the dynastic candidate knows history $p(t-1)$, that is, she knows whether her ancestor performed well or not, before making her entry decision.

Next, informed voters also learn about history and the state $s$. All voters vote subsequently. The governor chooses effort to learn about state $s$ and implement policies

\[ \text{Governor} \overset{\text{outcome } x}{\rightarrow} \{\varepsilon^*, p\} \]

\[ \text{Informed} \overset{\text{learn } s \& p(t-1)}{\rightarrow} \text{All voters} \overset{\text{vote}}{\rightarrow} \text{Candidates} \overset{\text{entry}}{\rightarrow} \]  

Figure 2: Entry and timing of events

---

4The Washington/ABC polls released on September 14, 2015 showed that only 36% of GOP primary voters prefer a candidate with experience in the political system whereas 58% want an outsider.
accordingly.

First, we note that, ex ante, before the state is revealed, the expected payoff of a governor of type-H and type-L are, respectively,

\[
\tilde{U}^H = W - \left[ \sigma^2_H + (1 - e^*_H)\sigma^2_s \right] - C_H(e^*_H), \\
\tilde{U}^L = W - \left[ \sigma^2_H + (1 - e^*_L)\sigma^2_s \right] - C_L(e^*_L)
\]  

(6)

conditional on her being of type-H and type-L. Clearly, a type-H governor has higher expected payoff than a type-L, i.e., \(\tilde{U}^H > \tilde{U}^L\).

In the first stage, candidates need to decide whether to run for office before knowing their own types. Each non-dynastic candidate is known, both to herself and to voters, to be of type-H with probability \(\alpha\). Thus, the ex ante expected payoff to a non-dynastic candidate from becoming the governor is equal to:

\[
EU^n = \alpha\tilde{U}^H + (1 - \alpha)\tilde{U}^L,
\]  

(7)

where \(\tilde{U}^k\) is as defined in (6).

By contrast, the dynastic candidate could be more or less likely to be of type-H depending on what has transpired in period \(t - 1\). To be specific, under Assumption 1, the expected payoff from becoming the governor to the dynastic candidate equals

\[
EU^d(p^*) = \alpha_t(p^*)\tilde{U}^H + (1 - \alpha_t(p^*))\tilde{U}^L > EU^n
\]  

(8)

conditional on history \(p(t - 1) = p^*\); but it equals

\[
EU^d(p^e) = \alpha_t(p^e)\tilde{U}^H + (1 - \alpha_t(p^e))\tilde{U}^L < EU^n
\]  

(9)

otherwise.

To run for office, a candidate needs to incur the fixed entry cost \(\xi\). We assume that:

**Assumption 2.** The entry cost is not too high, that is,

\[
EU^n \geq 2\xi.
\]

That is, conditional on the probability of being elected as governor is no less than \(\frac{1}{2}\), it is worthwhile running for office for a non-dynastic candidate. Thus, we have:
**Proposition 5.** Suppose $\theta_i$ follows a symmetric distribution. Under Assumption 2: (i) If $p^*$ is successfully implemented in period $t - 1$, the dynastic candidate runs for office in period $t$. (ii) If $p^e$ is implemented in $t - 1$, only non-dynastic candidates run for office in period $t$.

**Proof.** Proposition 2 shows that the dynastic candidate wins for sure conditional on $p^*$ being successfully implemented. In this case, her net payoff from running equals 
\[ \alpha_t(p^*)U^H + (1 - \alpha_t(p^*))U^L - \xi > 0, \]
so she should run for office. A non-dynastic candidate, however, loses for sure and the net payoff for a non-dynastic candidate from running equals $-\xi < 0$.

A dynastic candidate loses the election for sure conditional on $p^e$ being implemented, indicating negative net payoff from running.

With symmetric distribution, a candidate deemed less likely to implement the socially optimal policy has no chance of winning the election and therefore should not run. But with asymmetric distribution, the socially optimal policy is no longer the winning policy in all circumstances and there are states in which the majority of voters would prefer the alternative policy $p^e$ to $p^*$. In this case, the dynastic candidate does not win for sure even if the history is $p(t - 1) = p^*$. Meanwhile, she can still win election even if the performance of her ancestor is not good, i.e., $p(t - 1) = p^e$ as we explained before. Therefore, entry by non-dynastic candidate becomes viable when history is $p(t - 1) = p^*$. Of course, to ensure entry is a rational choice, the entry cost $\xi$ can not be too high relative to the probability of winning.

Let $\xi_1 = EU^n[F(2|\bar{\theta} - \theta^m|) - \frac{1}{2}]$ and $\xi_2 = EU^d(p^e)[F(2|\bar{\theta} - \theta^m|) - \frac{1}{2}]$. Recall that $EU^n$ is the expected payoff of a non-dynastic candidate from becoming the governor while $EU^d(p^e)$ is the payoff of a dynastic candidate conditional on $p^e$. So it is true that $\xi_1 > \xi_2$.

**Proposition 6.** Suppose the distribution $\Phi(\cdot)$ is asymmetric. (i) Non-dynastic candidates run for office in period $t$ regardless of the history $p(t - 1)$ when entry cost $\xi \leq \xi_1$; (ii) Dynastic candidate runs for office regardless the history $p(t - 1)$ when entry cost $\xi \leq \xi_2$.  

21
Since $\xi_2 < \xi_1$, the dynastic candidate will not run for office when $\xi \in (\xi_2, \xi_1]$ if the history is $p(t - 1) = p^*$. Hence, on average, a dynastic candidate whose ancestor did not perform well is less likely to run for office than a non-dynastic candidate.

Some remarks are worthwhile at this point. First, the above analysis shows that dynastic candidates are indeed more likely to win when they choose to run for office. This is true no matter whether the distribution of policy preferences is asymmetric no symmetric. On the other hand, most studies on dynasties may have overestimated the success rate of dynastic candidates relative to non-dynastic candidates. This is so as some (potential) dynastic candidates may have rationally chosen not to run because the poor performance of their ancestors negatively affects their electoral fortune.

Second, our result also suggests that, while political dynasties would be a persistent phenomenon in democratic societies where free elections to choose political leaders are in place, it might not be so for particular dynasties. A dynasty will usually end after a while. This happens when a low type descendant is ultimately ushered in, or because of uncertainty in policy implementation even when the descendant is of high type. Additionally, in times of great social and economic transformation, voters aspiring for change will abandon dynastic candidates in favor of non-dynastic candidates.

Whereas we assume the same entry cost for both dynastic and non-dynastic candidates, we can also extend to one in which dynastic candidates may have better outside option through family connection. In that case, a dynastic candidate would have higher entry cost than non-dynastic candidates. One prediction would be, if a dynastic candidate chooses to run for office, then she must expect to win with very high probability; otherwise, she should not run.

6 Reputation concerns: $\delta > 0$

A large literature on reputation (e.g., Holmström 1999) has showed that reputation concerns cause decision makers to sacrifice current gains for future benefits. In this section we show that concerns for family reputation play an important role in dynasties as well. In particular, concerns for the political careers and success of descendants
would increase the governor’s optimal effort regardless their types.

We assume here that reputational concerns arise only after a candidate wins the elections and knows her type. Before proceeding, we note that, when the governor chooses the optimal effort $e_t$ in period $t$, she expects that voters in $t + 1$ are more likely to vote for a candidate deemed to be of type-H with higher probability. This is true irrespective of whether the preference distribution $\Phi$ is symmetric or not.

To further simplify our analysis, we assume throughout this section that only dynastic candidates whose ancestor was the governor in period $t - 1$ and implemented $p^*$ successfully run for office in period $t$. This is not too restrictive as we have showed previously that, unless the entry cost is very low, a candidate whose ancestor implemented $p^e$ does not run for office in period $t$.

Previously we used $\alpha_t(p^*)$ to denote the probability of a dynastic candidate being of type-H conditional on the successful implementation of $p^*$ in the previous period. We let $\tilde{\alpha}_k$ be the probability of one’s descendant being of type-H conditional on the governor is of type $k$. Given the transition of types as specified in (1), we have:

\[ \tilde{\alpha}_H = \gamma \alpha + (1 - \gamma) > \tilde{\alpha}_L = \gamma \alpha. \]

When $\delta > 0$, the discounted payoff for a governor who is of type-$k$ and chooses $e$ equals:

\[ V^k_t = W - [\sigma^2 + (1 - e)\sigma^2] - C_k(e) + \delta e[\tilde{\alpha}_k \tilde{V}^H_{t+1} + (1 - \tilde{\alpha}_k) \tilde{V}^L_{t+1}]. \]

Here $\tilde{V}^k_{t+1}$ denotes the ex ante expected payoff of the descendant assuming she is of type $k$. The precise form of $\tilde{V}^k_{t+1}$ depends on the likelihood of her winning the election in period $t + 1$ and, thus, also on the distribution of voters’ preferences. With probability $e$, policy $p^*$ will be successfully implemented and the dynasty may be continued, in which case, her descendant may obtain $\tilde{V}^H_{t+1}$ or $\tilde{V}^L_{t+1}$ conditional on her type. But with probability $1 - e$, policy $p^e$ will be implemented and the dynasty will be terminated.

### 6.1 Symmetric distribution

We first consider the case of symmetric distribution. Again, in this case, the majority of voters prefer $p^*$ to $p^e$ regardless of the state. Therefore, conditional on history
\( p(t) = p^* \), the governor expects her descendant to win for sure in period \( t + 1 \). This implies:

**Lemma 4.** Assuming \( p^* \) was implemented in the last period, the ex ante expected payoffs of descendant of type \( k \in \{H, L\} \) are:

\[
\tilde{V}^H_{t+1} = V^H_t - \xi > \tilde{V}^L_{t+1} = V^L_t - \xi.
\]

**Proof.** To see this, note that conditional on \( p^* \) is implemented successfully, the descendant would be elected for sure when facing a non-dynastic candidate in \( t + 1 \). This is true as long as the ex ante probability that she is of type-H is greater than \( \alpha \) and, thus, the probability of her descendant being of type-H is \( \alpha_t(p^*) > \alpha \).

But conditional on her descendant being elected and their types being the same, she would face exactly the same problem as her ancestor. Thus, we have \( \tilde{V}^H_{t+1} = V^H_t - \xi \) and \( \tilde{V}^L_{t+1} = V^L_t - \xi \). This implies:

\[
V^H_t = W - \left[ \sigma^2 + (1 - e)\sigma_s^2 \right] - C_H(e) + \delta e[\hat{\alpha}_H V^H_t + (1 - \hat{\alpha}_H) V^L_t - \xi]
\]

and

\[
V^L_t = W - \left[ \sigma^2 + (1 - e)\sigma_s^2 \right] - C_L(e) + \delta e[\hat{\alpha}_L V^H_t + (1 - \hat{\alpha}_L) V^L_t - \xi].
\]

Because \( \hat{\alpha}_H = \gamma \alpha + 1 - \gamma > \hat{\alpha}_L = \gamma \alpha \), it must be true that \( \tilde{V}^H_{t+1} > \tilde{V}^L_{t+1} \). This concludes the proof. \( \square \)

This result implies that we can simplify the discounted payoff of the governor as:

\[
V^H = \frac{W - \sigma^2 - (1 - e_H)\sigma_s^2 - C_H(e_H) + \delta e_H(1 - \hat{\alpha}_H)V^L - \delta e_H \xi}{1 - \delta e_H \hat{\alpha}_H} \tag{11}
\]

\[
V^L = \frac{W - \sigma^2 - (1 - e_L)\sigma_s^2 - C_L(e_L) + \delta e_L \hat{\alpha}_L V^H - \delta e_L \xi}{1 - \delta e_L(1 - \hat{\alpha}_L)} \tag{12}
\]

**Proposition 7.** When \( \delta > 0 \) and \( \Phi(\cdot) \) is symmetric, the optimal effort \( \hat{e}^*_k \) is greater than \( e^*_k \) that is the optimal effort when \( \delta = 0 \). Moreover, optimal effort increases in \( \delta \), that is, for \( k = H, L \),

\[
\frac{\partial \hat{e}^*_k}{\partial \delta} > 0.
\]

24
Proof. To show the first part, we differentiate \( V^k \) with respect to \( e \). This gives:

\[
\frac{\partial V^H}{\partial e} = \frac{\sigma_s^2 - C'_H(e_H) + \delta(1 - \bar{\alpha}_H)V^L - \delta \xi}{(1 - \delta e_H \bar{\alpha}_H)} + \frac{\delta \bar{\alpha}_H[W - \sigma_b^2 - (1 - e_H)\sigma_s^2 - C_H(e_H) + \delta e_H(1 - \bar{\alpha}_H)V^L - \delta e_H\xi]}{(1 - \delta e_H \bar{\alpha}_H)^2},
\]

(13)

\[
\frac{\partial V^L}{\partial e} = \frac{\sigma_s^2 - C'_L(e_L) + \delta \bar{\alpha}_L V^H - \delta \xi}{1 - \delta e_L(1 - \bar{\alpha}_L)} + \frac{\delta(1 - \bar{\alpha}_L)[W - \sigma_b^2 - (1 - e_L)\sigma_s^2 - C_L(e_L) + \delta e_L \bar{\alpha}_L V^H - \delta e_L\xi]}{[1 - \delta e_L(1 - \bar{\alpha}_L)]^2}.
\]

(14)

If we let \( e_H = e^*_H \) and \( e_L = e^*_L \), then

\[
\frac{\partial V^H}{\partial e}|_{e^*_H} = \frac{\delta[\bar{\alpha}_H V^H + (1 - \bar{\alpha}_H)V^L - \xi]}{1 - \delta e^*_H \bar{\alpha}_H} > 0
\]

\[
\frac{\partial V^L}{\partial e}|_{e^*_L} = \frac{\delta[\bar{\alpha}_L V^H + (1 - \bar{\alpha}_L)V^L - \xi]}{1 - \delta e^*_L(1 - \bar{\alpha}_L)} > 0.
\]

The two inequalities follow from Assumption 2, which says that, conditional on winning the election for sure, a candidate would obtain strictly positive net payoff and, thus, is willing to incur the campaign cost to run for office. Hence, we conclude that the optimal effort \( \hat{e}^*_H > e^*_H \) and \( \hat{e}^*_L > e^*_L \).

To prove the second part, first we note that at \( \hat{e}^*_k, \frac{\partial V^k}{\partial e} = 0 \) for \( k = H, L \). After some transformation, the first order condition (13) becomes:

\[
\sigma_s^2 - C'_H(\hat{e}^*_H) + \delta(1 - \bar{\alpha}_H)V^L + \delta \bar{\alpha}_H V^H - \delta \xi = 0.
\]

Differentiating the two sides of the equality with respect to \( \delta \) we have:

\[
-C'_H \frac{\partial \hat{e}^*_H}{\partial \delta} + (1 - \bar{\alpha}_H)V^L + \bar{\alpha}_H V^H + \delta \bar{\alpha}_H \frac{\partial V^H}{\partial \hat{e}^*_H} \frac{\partial \hat{e}^*_H}{\partial \delta} - \xi = 0.
\]

Using the fact that \( \frac{\partial V^H}{\partial e}|_{\hat{e}^*_H} = 0 \), we can simplify to get

\[
\frac{\partial \hat{e}^*_H}{\partial \delta} = \frac{(1 - \bar{\alpha}_H)V^L + \bar{\alpha}_H V^H - \xi}{C'_H} > 0.
\]

Similarly we can show that

\[
\frac{\partial \hat{e}^*_L}{\partial \delta} = \frac{\bar{\alpha}_L V^H + (1 - \bar{\alpha}_L)V^L - \xi}{C'_L} > 0.
\]

This concludes the proof of the proposition. \( \square \)
Proposition 7 shows that concerns for family reputation and the prospect of one’s
descendant increases the governor’s optimal effort level. The more likely one’s de-
scendant will run for office in future (higher \( \delta \)), the more effort is the governor willing
to exert to maintain good reputation.

### 6.2 Asymmetric distribution

In this subsection we consider the case in which voters’ preferences are asymmet-
rically distributed. From the previous discussion we already know that, when \( \Phi(\theta_i) \)
is asymmetric, successfully implementing the socially optimal policy \( p^* \) in period \( t \)
does not guarantee the electoral success for one’s descendant in period \( t + 1 \). This is
particularly true in times when there is a significant gap between the socially opti-
mal policy \( p^* \) and the median voter’s preferred policy \( p^c \). Nevertheless we show that
a similar result as that in Proposition 7 still holds. Namely, concerns for the political
prospect of one’s descendant still prompts a governor to exert higher effort than she
would in the absence of the concern.

The main difference between the symmetric and asymmetric distribution cases
is that, conditional on successfully implementing \( p^* \), now the governor expects her
descendant to win the election with probability \( \pi = \frac{3}{2} - F(2|\bar{\theta} - \theta^m|) \). On the other
hand, provided that her descendant wins the election, she would still face the same
problem as the governor faced. Hence we have a similar result as Lemma 4.

**Lemma 5.** Assuming \( p^* \) was implemented in the last period, the ex ante expected pay-
offs of descendant of type \( k \in \{H, L\} \) are:

\[
\tilde{V}_{t+1}^H = \pi V_t^H - \xi > \tilde{V}_{t+1}^L = \pi V_t^L - \xi.
\]

Note that conditional on \( p^* \) in period \( t \), her descendant of type-\( k \) would win election
and become governor in period \( t + 1 \) with probability \( \pi \) and, thus, ex ante, has an
expected payoff of \( \pi V_t^k \). Next we show that her optimal effort still increases in the
degree she cares about future.
Proposition 8. Suppose $\delta > 0$ and $\Phi(\cdot)$ is asymmetric, the optimal effort $\hat{e}_k^*$ is greater than $e_k^*$ that is the optimal effort when $\delta = 0$. Moreover, the optimal effort increases in $\delta$, i.e., for $k = H, L,$

\[
\frac{\partial \hat{e}_k^*}{\partial \delta} > 0.
\]

Proof. First we note that in the case,

\[
V^H = W - \sigma_0^2 - (1 - e_H)\sigma_s^2 - C_H(e_H) + \delta \pi e_H (1 - \tilde{\alpha}_H) V^L - \delta e_H \xi
\]

\[
V^L = W - \sigma_0^2 - (1 - e_L)\sigma_s^2 - C_L(e_L) + \delta \pi e_L \tilde{\alpha}_L V^H - \delta e_L \xi
\]

Letting $e_H = e_H^*$ and $e_L = e_L^*$, we have:

\[
\frac{\partial V^H}{\partial e} \bigg|_{e_H^*} = \frac{\delta [\pi \tilde{\alpha}_H V^H + \pi (1 - \tilde{\alpha}_H) V^L - \xi]}{1 - \delta \pi e_H^* \tilde{\alpha}_H} > 0.
\]

\[
\frac{\partial V^L}{\partial e} \bigg|_{e_L^*} = \frac{\delta [\pi \tilde{\alpha}_L V^H + \pi (1 - \tilde{\alpha}_L) V^L - \xi]}{1 - \delta \pi e_L^* (1 - \tilde{\alpha}_L)} > 0.
\]

Both inequalities follow from Assumption 2 and the fact that $\pi \geq \frac{1}{2}$. Hence, we conclude that concern for the payoff of one’s descendant always increases effort.

Next, we differentiate the first order condition with respect to $\delta$ and make transformation, which gives:

\[
\frac{\partial \hat{e}_H^*}{\partial \delta} = \frac{\pi (1 - \tilde{\alpha}_H) V^L + \pi \tilde{\alpha}_H V^H - \xi}{C_H'} > 0.
\]

Similarly we can show that

\[
\frac{\partial \hat{e}_L^*}{\partial \delta} = \frac{\pi \tilde{\alpha}_L V^H + \pi (1 - \tilde{\alpha}_L) V^L - \xi}{C_L'} > 0.
\]

This concludes the proof of the proposition.

\[
\square
\]

7 Extension and discussion

Previously we have assumed that voters have short memory and reputation lasts only for one period. Thus, the reputation effect is short-lived and according to this definition a candidate whose ancestor was the governor in periods before $t - 1$ becomes
a non-dynastic candidate. For example neither Jeb Bush nor Hillary Clinton would be
dynastic candidates, since their “ancestors,” for the former brother George W. Bush,
for the latter husband Bill Clinton, were not in office immediately before they decided
to run. If so our example at the beginning of the paper would be misplaced, and
this could lead one to think that our results hold only under this quite restrictive
assumption. However, we hope readers will be convinced soon that this assumption
has been made purely for ease of exposition and it is not key to our results.

We relax the short memory restriction and take a candidate as a dynastic one
if she had ancestors in politics more than one generation before. We also discount
the dynastic advantage as the number of generations increases. From the voter’s
perspective what matters is not whether a candidate has name recognition, but rather
it is the extra information that can be used to determine the candidate’s type. We
denote by $\alpha(p^*_{t-h})$ with $h \geq 1$ a dynastic candidate with ancestor in office in period
$t-h$ who successfully implemented policy $p^*$. For simplicity we assume that the
posterior belief that their ancestors were type-H are all equal and denote it by $\alpha^*$,
that is, $\alpha^* = \text{Prob}(\tau_j(t-h) = H | p^*)$ for all $j$ and $h$. In this case, we have:

$$\alpha(p^*_{t-h}) = \alpha[1 - (1 - \gamma)^h] + (1 - \gamma)^h \alpha^*. $$

Clearly, as $h$ increases, $\alpha(p^*_{t-h})$ approaches $\alpha$.

Hence, we have the following results:

1. When dynastic and non-dynastic candidates compete against each other, on av-
   erage, the dynastic one $\alpha(p^*_{t-h})$ with the lowest $h$ is more likely to win. It is as
   if voters remember good past performance but put more weights to more recent
   good performance.

2. With symmetric distribution of $\Phi$, the dynastic candidate with lowest $h$ wins for
   sure.

3. With asymmetric distribution, either the dynastic candidate with the lowest $h$
   wins or a non-dynastic candidate wins, depending on the state $s$. 

28
The model developed in this paper is highly stylized. For ease of exposition we have assumed all candidates are ex ante identical. The simplification helps highlight the advantage a dynastic candidate has over an otherwise identical candidate. One may ask what happens in a richer model in which dynastic and non-dynastic candidates are from different parties that differ in their own policy preferences. Note that if that is the case, then what matters most to voters is the policy they would implement, not whether they are dynastic candidates or not. For example, the family name Clinton is only a plus for Hillary among Democratic voters, but could be a negative among Republican voters. Something similar can be said of the name of Bush. Hence, a model that also incorporates partisan differences would not give us significantly different results from what is obtained from this simple model.

Perhaps of equal importance, our analysis ignores the quality of the candidates themselves. Whereas agreeing with a candidate’s positions is the most important factor guiding voters’ vote, personality, or lack of it, is another important factor. Great campaigners can and do sway voters who may disagree with some of their views, while lackluster ones can disenchant even their natural supporters. On the other hand, introducing this additional heterogeneity would significantly complicate the analysis without shedding any new insights.

In addition, we assume there is only one policy voters care and voters’ preferences are exogenously given. In real world, there are a number of policies voters care. And depending on the policy that voters care the most in an election, the distribution of voters’ preference would be endogenous and changing over time. For example, after a major terrorist attack, voters tend to care more about national security. After long period of stagnant economic growth and increase inequality, voters may shift their views on trade and redistribution policy. However, an extension in that direction should not significantly affect our main results.

---

5For a paper on candidates’ character, see Kartik and McAfee (2007).
8 Conclusion

The model proposed here sheds light on some relevant features of the politics of families and dynasties. Most of all, we provide a formalization which is able to explain phenomena found in previous and current empirical literature. First of all, central is the relationship between voters’ political polarization, measured as the distance between the average and median voters ideal policies, and dynastic success. Second, the entry selection process plays an important role in keeping out from the race those dynasts that are believed to be more likely to be of low quality. Third, dynastic behavior in politics is related both to inherited and bequest name recognition and quality. A two generations dynamic model as the one proposed here seems thus to be the appropriate pathway to proceed for understanding better the emergence, persistence, and extinction of political dynasties on the one hand, and their effect on social political and economic outcomes on the other.

While dynastic history and perspectives are conducive to better policies in some cases, political polarization can still explain why low quality dynasts have a chance in winning anyway, and that the larger the polarization the more similar are the chances of winning between high and low quality dynasts. However, if entry costs keep out from running those dynasts believed to be of low quality, then the edge gained and measured in empirical paper could be explained by this selection bias, because only dynasts that are likely to be of good quality decide to run. The fact that dynasts are sometimes not conducive to good policies and outcomes thus can be explained both by low entry costs unable to deter low quality dynasts from the political race, or by large political polarization, which replaces good dynasts with bad ones.

References


