Political Dynasties and Economic Development: Theory and Evidence from India *

Anders Kjelsrud[1], Anirban Mitra[2], and Arnab Mukherji[3]

[1]Oslo Metropolitan University

[2]University of Kent

[3]IIM Bangalore

December 13, 2024

[Preliminary and Incomplete]

Abstract

Political dynasties exist in many democratic countries at different levels of government. How do such dynasties affect public policy? In particular, do they promote clientelism? We study these questions in the context of India where we map familial ties among politicians in the state legislature of 8 major states. Our theory predicts that the presence of dynastic candidates increases lobbying and clientelism by the wealthy. Our empirical analysis focuses on asset accumulation of candidates and local implementation of the largest workfare public program in the world (MNREGA). We find that State Assembly constituencies with political dynasties competing exhibit higher wealth among their top contestants, and that areas exposed to dynasties receive less MNREGA employment.

Keywords: Clientelism, Political dynasties, MNREGA

*We thank Camille Boudot-Reddy and seminar participants at the Oslo Metropolitan University and Birkbeck Business School for useful suggestions and comments. The usual disclaimer applies.

1 Introduction

Political dynasties exist in most countries. They appear in democratic and non-democratic countries alike, and at central and local levels of governance (Chandra, 2016; Dal Bó et al., 2009). The presence of such dynasties may affect policy – and economic development more broadly – through numerous channels. In this paper, we argue that one key channel is through changing the environment regarding clientelism and the amount of *implicit* quid pro quo possible between the political players and the local elites.

We first posit a simple theoretical model to outline how the *presence* of political dynasties affects the incentives of the local elite to engage in capture. The logic rests on two core ideas. First, the elites care about multiple elections (long-term horizon), and hence, they would like to form relational contracts with political candidates. The key assumption here is that transacting with a familiar politician over time allows greater returns on the their 'investment'. Second, a political dynasty – as an entity – is more likely to re-contest in future elections following a defeat as compared to a single candidate.

Combining these two features in a simple two-period model, we show that the presence of political dynasties increases the overall level of capture by the elite. The intuition behind this result is that the elite will wish to contribute to both candidates to hedge their risk, but owing to their desire of transacting with familiar faces, they will back the non-dynastic candidate *more* in the first period. This is so since the elite is aware that the dynasty is more likely to re-contest in the next period regardless of the outcome in the current period. This behaviour leads to more overall clientelism when dynastic candidates compete.

To test the empirical relevance of this potential channel, we collected novel data on political dynasties operating in Indian state-level elections. In particular, we map familial relations between politicians in 8 large states: Bihar, Haryana, Karnataka, Kerala, Punjab, Rajasthan, Tamil Nadu and West Bengal. We define candidates as dynastic if they have a family member who previously participated in an election at either the national or state level, and collected data on this as follows. We began with a set of leads on potential candidates whose family members had competed in earlier elections. Our research team then went to the capitals of each state in our sample to visit the offices of political parties active in those states. The aim of this was to validate the leads and to identify additional candidates. Further confirmation was sought through discussions with journalists and university professors actively writing about elections, as well as by reviewing media reports and public writings. This approach allowed us to identify not only close familial relationships, such as father-son, but also more distant ones, such as grandparent-grandchild, uncle-nephew, and others. We consider this data collection as a key contribution of the paper, as there is currently a lack of systematic evidence on political dynasties in India below the national level.

Clientelism is challenging to measure especially in contexts where funds from the elite to politicians are transferred either without official records or under misleading headings. In the Indian context, which possibly resonates with many developing countries, the implicit quid pro quo would entail personal pecuniary gains for the politician *and* for the elite. Specifically, the elite may be able to sway politicians to tilt policies in their favor at the expense of the (poor) median voter.

We capture this aspect of pecuniary gains to the politicians by utilising a particular feature of the Indian political system — namely, the mandatory disclosure of asset holdings as part of the "Right to Information Act" beginning in 2003. These disclosures provide a snapshot of the market value of a contestant's assets, just prior to the election for which they were filed. We begin by analyzing the asset values of the top two candidates in each State Assembly constituency (hereafter, AC) for elections held between 2005 and 2018. Strikingly, we find that dynastic candidates have almost twice as much wealth as other top candidates. We also find that non-dynastic politicians *competing against* the dynastic candidates have higher wealth than other comparable politicians within the same state. These descriptive patterns are consistent with our proposed theoretical mechanism, namely that the presence of a dynasty leads to more clientelism and that it raises the overall level of donations from local elites to politicians.

In our main empirical analysis, we shift focus to examining the policy impacts of such clientelism. Guided by the theory, our expectation is that ACs with active political dynasties would implement less pro-poor policies, reflecting a shift *away* from the preferences of the poor median voter. We test this by studying the local implementation of the world's largest public workfare program, the Mahatma Gandhi National Rural Employment Guarantee Act (henceforth, MNREGA). MNREGA is designed to employ low-skilled workers typically on public infrastructure development, and provides a legal entitlement of 100 days of paid employment per year for rural households. Although the implementation of MNREGA relies on self-selection and demand for such work, in practice the constraints are on the supply-side (see Dutta et al. (2014), Maiorano (2014) among others). Moreover, previous research has shown that state-level politicians (MLAs) have both incentives and opportunity to influence the program implementation (see e.g., Gulzar and Pasquale (2017)).

Using village council (Gram Panchayat) level data on MNREGA as the metric of pro-poor policy, we utilise the Delimitation Act of 2008 to exploit exogenous exposure to dynastic politicians. The aim of the Delimitaion Act was to equalise the population across constituencies and to demarcate the ones for mandated representation of Scheduled Castes and Tribes (SCs and STs). The consensus view is that this was done without much influence from politicians (e.g., influential incumbents did not experience any significant gains and the smaller and larger constituencies were the ones to experience most changes owing to the Act).[1]

We leverage the boundary changes to estimate plausible causal effects of political dynasties on the provision of pro-poor policy. Our approach is to compare areas that were originally within the same AC – and exposed to the same political dynasty – but that became part of different ACs after the Delimitation. Clearly, the choice of which post-Delimitation AC the dynasty decides to compete in is an endogenous decision. Our strategy, in a nutshell, is to "allocate" dynasties to the post-Delimitation ACs with the greatest population overlap with the former ACs. We validate this approach in two ways. We first show that the (greatest) population overlaps strongly predict where the dynasty ends up competing. We next show that villages within the area with the greatest population overlap are similar to villages in other parts of the pre-Delimitation AC on observable characteristics (i.e. the "treatment" and the"control" groups are balanced).

Using this empirical approach, we find robust and consistent evidence of lower supply of MNREGA in areas exposed to dynastic candidates post the Delimitation. Specifically, we find a reduction in the amount of wages paid, in the number of days worked and in the number of persons employed. These effects are robust to multiple alternative specifications.

Taken together, our findings on higher wealth and lower MNREGA implementation strongly corroborate our central hypothesis that the presence of political dynasties fuels greater capture by the elite.

Our paper contributes to several strands of earlier literature. There exists work on democratic capture by the elite or other interest groups by means of vote buying, voter co-optation, patronage networks, and the use of force or its threat (e.g., Bardhan and Mookherjee (2006)) at local or decentralised levels.[2] However, there is a paucity of research investigating the effects of political dynasties on clientelism and capture.

There has been a growing interest in understanding the role of hereditary leadership and political dynasties on economic growth and development. For instance, Besley and Reynal-Querol (2017) view hereditary leadership as a relational contract which improves policy incentives. They compile and analyse a unique dataset on leaders between 1874 and 2004 from 197 countries in which they are classified as hereditary leaders based on their family history. Their main empirical finding is that economic growth is higher in polities with hereditary leaders but only if executive constraints are weak. Another closely related study is George (2020). He compiles data on the biographical profiles of all 4807 MPs since India's

[1] See Iyer and Reddy (2013), Bardhan et al. (2020) and Kjelsrud et al. (2020) among others for details this and on the Delimitation Act itself.

[2] In a related vein, Besley et al. (2005) find that education increases the likelihood of selection to public office and reduces the odds that a politician uses political power opportunistically.

first parliament in 1952. His analysis offers three empirical findings. First, he finds that descendants worsen poverty and public good provision in villages they represent. Second, he uncovers that founders have positive effects on economic development. Third, he finds that dynastic politics generates a “reversal of fortune” development pattern, where places develop faster in the short run (because of positive founder effects), but are poorer in the long run (because negative descendant effects outweigh positive founder effects).

The rest of the paper is organised as follows. Section 2 describes our data on political dynasties in the 8 Indian State Assemblies. Section 3 contains a simple model which sets the ground for our empirical analysis. Sections 4 and 5 contain the details of empirical analysis — specifically, our identification strategy and the main results when examining the assets of MLAs and the implementation of MNREGA. Section 6 concludes. Other relevant details are contained in the Appendix.

2 Political dynasties in Indian state elections

We start this section by describing our data collection. In the current version of the paper, we focus on 8 large states: Bihar, West Bengal, Rajasthan, Punjab, Haryana, Karnataka, Kerala and Tamil Nadu. Collectively, these states account for about 45% of the Indian population and they span all parts of the country.[3] We end the section by providing descriptive statistics.

2.1 Data collection and definitions

In the current version of the paper, we focus on 8 large states of India, namely, Bihar, West Bengal, Rajasthan, Punjab, Haryana, Karnataka, Kerala and Tamil Nadu. Collectively, these states account for about 45% of the Indian population and they span all parts of India.[4] For each LA state election, we see candidate-specific details at the level of competition, i.e. the LA constituency, provided a valid election took place.[5] This Election Commission data captures the number of competitors in the constituency, their political parties, final vote share, and some contestant-specific attributes such as the father’s name, age, etc. We begin with this data to identify a contestant as a dynastic contestant (or not).

We *define* a candidate as *dynastic* candidates if the candidate had a family member that participated in an earlier election for the upper or lower house or in an earlier LA election

[3] In ongoing work we seek to expand the sample to other states.

[4] In ongoing work we seek to expand the sample to other states.

[5] Once elections have been announced, they may be called off before an election takes place in the instance of a death of one of the candidates, during or immediately after the election, if there is a report of violence, voter fraud, or if any other critical irregularity is reported. In such instances, elections and election results of that constituency are suspended, and the process is held later.

of that state. The first member of each dynasty thus, starts as a routine candidate, but when a member of her family competes in a subsequent LA election, then both the family member and the first member are seen as members of a political dynasty. We track family in a loose sense, and this includes, father-offspring, mother-offspring, husband-wife, siblings, uncle-nephews or nieces, and sometimes even in relationships with in-laws (father-in-law and son-in-law, mother-in-law and daughter-in-law, etc.).

While the candidates we focus on candidates who participated in LA elections from 2005 onwards, their family members may have competed in a prior election, anytime after India's independence. Thus, we scan elections from 1951 onward, i.e. the conduct of the first national elections in India till the LA election in which each candidate participated. The process of identifying a family member is more complex and we follow a multi-step process for candidates in each state.

To identify a familial relationship between one electoral candidate and another we began by first visiting each state capital included in the study. Our study team began by visiting the offices of each of the national political parties and regional parties that were active in that state. At each party office, we began with a set of leads on political candidates who had family members who had competed in an earlier election. We sought parallel confirmation by speaking with journalists, or university professors from different political science departments who were actively writing about elections. This mapping of the familial relationship was further explored by looking at media reports and public writing. For some of these familial relationships, for example, father-child, and father-child-grandchild, we were able to use self-reported data by the candidate on the father's name to check if our search had missed anyone. Our political office, journalist, and academia-based word-of-mouth search strategy picked up most of the father-child, father-child-grandchild instances that we could corroborate in the reported election data giving us confidence that this would also be complete for the other familial relationships that were reported.

2.2 Descriptive statistics

Table 1 shows the number of dynasties with at least one active descendant during the period 2000 to 2022. In total, we identify 235 dynasties across the 8 states. On average, these dynasties have 2.3 members, including the founder. The median number of members is 2. Hence, the typical political dynasty in our sample is small, with one founder and one descendant.

[Add anecdotes]

To give a sense of the prevalence and importance of the dynasties, we also plot the share of MLA positions they won over the time period. On total, 4.7% of the MLAs were descendants in a dynasties. The highest share is found in Haryana (11.6%) and the lowest in West Bengal (2.5%).

In Table 2, we plot the most common types of relationships with the dynasties. 53% of the descendants in our sample are sons of the founder. Grandson is the second most common connections with the founder (8.2%), followed by siblings (7.5%), spouse (7.2%) and daughter (6.4%).

Finally, in Table 3 we show gender and statistics on the electoral success of founders and descendants. Only 2.8% of the dynastic founders in our sample are females. In contrast, as much as 21% of the descendants are females. On average, the founders participated in 4.8 state elections, which is about twice the number for descendants. While the founders won 55% of the elections they participated in, the descendants won 46% of the elections. The higher popularity of the founders are also reflected in the average vote share of 43% versus 38% for the descendants.

Table 1: Number of active dynasties (2000-2022), by states

	#Dynasties	Mean #members	Dynastic MLA share
	(1)	(2)	(3)
Bihar	69	2.3	0.056
Haryana	17	3.1	0.116
Karnataka	22	2.3	0.034
Kerala	25	2.2	0.054
Punjab	16	2.5	0.058
Rajasthan	37	2.0	0.053
Tamil Nadu	29	2.1	0.034
West Bengal	20	2.3	0.025
All	235	2.3	0.047

Note: The table shows descriptive statistics of the dynasties in our sample. Column 1 shows the number of dynasties, Column 2 the average number of members (including the founder), while Column 3 shows the share of MLAs over the period 2000-2022 that are descendants in a political dynasty.

Table 2: Most common relationships with founder, shares

	(1)
Son	0.532
Grand children	0.082
Sibling	0.075
Spouse	0.072
Daughter	0.064
Daughter-in-law	0.041
Son-in-law	0.030
N	268

Table 3: Election results (1962-2022) and gender of founders and descendants

	Founders (1)	**Descendants** (2)
Elections run	4.787	2.424
Share winning	0.545	0.461
Mean vote share	0.433	0.380
Female	0.028	0.209
N	210	312

3 Theory

Our theory takes as a starting point the idea that donors (business groups, self-employed persons, etc.) contribute funds to some political parties. They do so in the hope of securing some pecuniary gains in the future (post elections). Clearly such donors are not concerned with elections in just a *particular* period; it is the long term gains they value. This brings to the fore the issue of the expected returns from this implicit *quid pro quo* between a donor and a political party. Given that these are informal contracts, the gains to a donor from transacting with a familiar face across different periods are higher than when such a relationship has to be forged with a new candidate every period. And this is precisely why having a political dynasty (being associated with a particular party) can make a palpable difference to these informal contracts.

This is the environment in which we set up a simple formal model to investigate the link between such forms of election funding and the presence of political dynasties.

3.1 The Model

Consider a typical electoral district. Let there be two political parties here, denoted by A and B. Start with the case where we have a political dynasty active in this district. W.l.o.g. let party A be associated with this family. We will subsequently turn to the case where there are no political dynasties.

What is the defining feature of a political dynasty? A political dynasty is able to field some member of the family as a candidate from party A in *every* election. For other parties (party B, in the context of our model), the fielded candidate need not be the same every time. In a sense, a political dynasty has an identity of its own which exists independently of the party identity. It is like a "single candidate" in a dynamic multi-period setting, while for party B the candidate potentially changes at every election. We will deal with this in detail later.

Consider a typical donor, D. To keep the setup as simple as possible, we assume that D lives for two periods.[6] In each period, there are elections in which party A's candidate is pitted against party B's. In each of these periods D faces the following decision problem: how much funds (if anything) to donate to each of the two parties' fielded candidates. Why would one donate? A contribution by D towards party i (where $i = A, B$) in period t (where $t = 1, 2$) has two effects. First, it brings to D a positive return, to be interpreted as economic gains. This only accrues if the recipient of the donation actually *wins* the election in period t. Secondly, it affects the relative chance of success for party i. So the marginal effect of

[6] The basic logic behind our results extends to a finite N–period problem for $N > 2$.

a positive contribution to party i in period t (weakly) increases the probability of party i's victory in the same period.

The return to D from contributing $x_i \geq 0$ to party i in period 1 is λx_i, conditional on party i winning, where $\lambda > 1$; otherwise D gets 0. This return should be thought of as pecuniary gains arising from awarding of contracts, concessions etc. Clearly, this is informal and hence depends upon the "personal rapport" D shares with the party i's candidate. This is also why D would prefer "transacting" with a familiar candidate.

Now turn to period 2. Here the return to D from contributing x_i to party i's candidate is λx_i (conditional on party i winning) *only if* the same candidate is fielded/a candidate from the same political dynasty is fielded. Otherwise, there is a dip in the return by a factor $\delta \in (0,1)$ and the return becomes $\lambda\delta x_i$. We assume $\lambda\delta > 1$ so that there is still gains to be had from contributing to a new candidate. What we want to emphasize is that the value of a "new relationship" is lower since trust needs to be built over time and through repeated interactions. This is plausible particularly in a setting like ours where formal contracts are not in place or infeasible.[7]

These monetary returns are translated into payoffs for D by means of the utility function $u(.)$ with standard properties. Specifically, we have $u' > 0$, $u'' < 0$ and $u'(0) \to \infty$. Additionally, we impose the following requirement:

$$\tau u'(\tau x) \leq u'(x) \ \forall \tau \in (0,1) \text{ and } x > 0.$$

This is effectively an upper bound on the curvature of $u(.)$ hence rules out D being "too" risk-averse. This assumption is clearly not without loss of generality but it simplifies the analysis considerably. Moreover, it is not particularly restrictive as several commonly used utility functions satisfy this.[8] There is a (strictly) convex cost of making these contributions. This cost should be thought of as the valuation of D's income net of contributions.[9]

The mapping from the contributions to the election outcome in any period is given by the following. Let p_t^A denote the probability of party A' victory in period t, for $t = 1, 2$. Then, we have:

$$p_t^A = \omega \frac{\rho x_t^A}{\rho x_t^A + x_t^B} + (1-\omega)p^A.$$

Here $\omega \in (0,1)$, $\rho \geq 0$ and p^A is a measure of $A's$ strength in elections independent of any contributions. In the situation where there is no political dynasty, a natural choice for ρ is

[7] This in the spirit of Ghosh and Ray (1996). They study cooperative behaviour in communities with limited flow of information. The equilibria in their setup are characterized by an initial testing phase, followed by cooperation if the test is successful.

[8] Take the iso-elastic utility function $u(x) = x^{(1-\sigma)}/(1-\sigma)$. All $\sigma \in (0,1)$ satisfy our restriction. Also, $u(x) = ln(x)$ meets the requirement.

[9] The strict convexity of the cost makes the objective functions (developed below) concave and hence guarantees a maxima to the constrained optimisation problems.

unity and for p^A is 1/2. Note, $\rho > 1$ is appropriate for strong candidates from a political dynasty. Also, $p^A > 1/2$ is indicative of party A being popular. The parameter ω determines the weight the contributions have in tilting the election one way or the other.

Additionally, we assume that whenever a candidate from a party *not* associated with a political dynasty loses an election in the first period, he is replaced with some exogenous probability $\pi \in (0, 1)$. For a dynastic party, there is *no* replacement in the sense that some member within the family is fielded. Recall, a new candidate (post-replacement) is viewed less favourably by D.

Figures 1 and 2 below depict the sequence of moves in the two possible worlds: one with a dynasty and one without.

Now we are in a position to outline D's objective function. Note, it involves payoffs from two periods. We begin with the second period. In period 2, exactly one of the two situations arise: the incumbent (i.e., the winner from period 1's election) is either from party A or from party B.

Consider the case where the incumbent is from party B. D chooses (x_2^A, x_2^B) to maximise the following:

$$p_2^A u(\lambda x_2^A) + (1 - p_2^A)u(\lambda x_2^B) - c(x_2^A + x_2^B)$$

where $c(.)$ captures the "cost" of contributions — in terms of forgone consumption of other goods — as discussed above. Moreover, $c(0) = 0$, $c' > 0$ and $c'' > 0$. This cost function implicitly puts a bound on the extent of the contributions. Additionally, we assume $c(.)$ is sufficiently convex to ensure the objective function is concave.[10] Notice, here that the return from either party winning is λ times the contribution as both candidates are the same as in the previous election; hence no dilution in the rate of return (no δ).

Let $(\overline{x}_2^A, \overline{x}_2^B)$ denote the arg max of the above problem.

Also, let $V_B \equiv p_2^A u(\lambda \overline{x}_2^A) + (1 - p_2^A)u(\lambda \overline{x}_2^B) - c(\overline{x}_2^A + \overline{x}_2^B)$ where p_2^A is evaluated at $(\overline{x}_2^A, \overline{x}_2^B)$.

Now consider the other possibility, namely, the case where the incumbent is from party A. Here D chooses (x_2^A, x_2^B) to maximise the following:

$$p_2^A u(\lambda x_2^A) + (1 - p_2^A)[\pi u(\delta\lambda x_2^B) + (1 - \pi)u(\lambda x_2^B)] - c(x_2^A + x_2^B)$$

Here, the return from party B's candidate winning depends upon whether it is the same candidate as before or if it is a new face fielded by B. As discussed earlier, the return to D in the latter case is lower (captured by the factor δ). Notice from the way the problem has been set up, D's choice of (x_2^A, x_2^B) is made *prior* to party B's deciding on whether to persist

[10]Note, the strict concavity of $u(.)$ is not sufficient to guarantee that $p_2^A u(\lambda x_2^A) + (1 - p_2^A)u(\lambda x_2^B)$ is concave in (x_2^A, x_2^B) as p_2^A also depends upon (x_2^A, x_2^B).

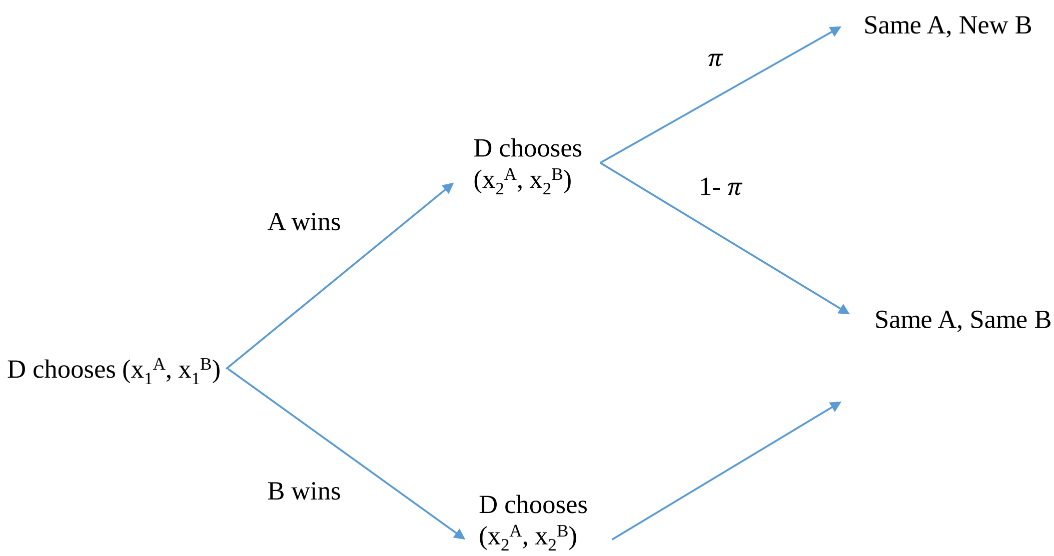


Figure 1: The sequence of moves in a dynastic setup.

with last period's losing candidate or try a new candidate. The idea is that parties seek contributions from donors well in advance whilst specific candidates are decided at dates much closer to elections.[11]

Let $(\underline{x}_2^A, \underline{x}_2^B)$ denote the arg max of the above problem.

Also, let $V_A \equiv p_2^A u(\lambda \underline{x}_2^A) + (1-p_2^A)[\pi u(\delta\lambda \underline{x}_2^B) + (1-\pi)u(\lambda \underline{x}_2^B)] - c(\underline{x}_2^A + \underline{x}_2^B)$ where p_2^A is evaluated at $(\underline{x}_2^A, \underline{x}_2^B)$.

Notice, by construction, $V_B \geq V_A$.

Now we return to D's problem in period 1. Here, D chooses his actions — (x_1^A, x_1^B) — with the aim of maximising the payoffs over *both* periods. More formally, the objective function is the following:

$$p_1^A[u(\lambda x_1^A) + V_A] + (1-p_1^A)[u(\lambda x_1^B) + V_B] - c(x_1^A + x_1^B)$$

where $p_1^A = \omega \frac{\rho x_1^A}{\rho x_1^A + x_1^B} + (1-\omega)p^A$. To keep thing simple, we are not discounting the payoffs from the second period.[12]

Notice, in the case of no political dynasty, it does not matter to D as to *which* party's candidate is the incumbent at the start of period 2. So for D, the second period problem

[11]This timing structure is not crucial to our results. We discuss later an alternative timing structure where D's choice of (x_2^A, x_2^B) is made *subsequent* to party B's decision on a specific candidate. The main results are robust to such timing issues.

[12]Introducing a discount factor would not affect the results in any qualitative manner.

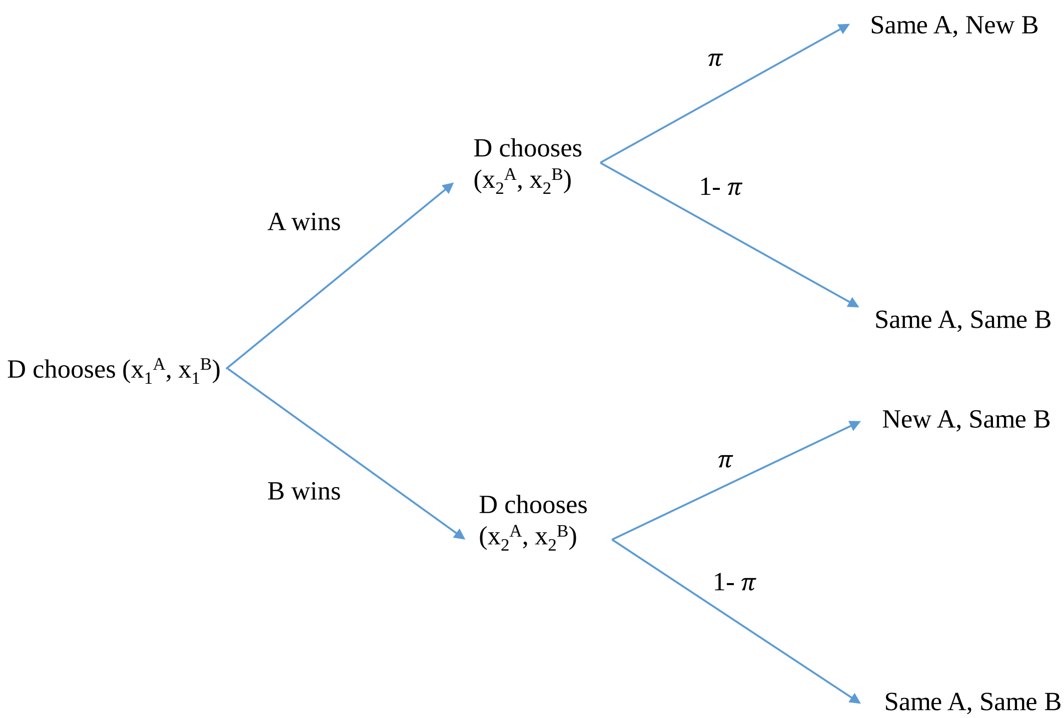


Figure 2: The sequence of moves in a setup without a dynasty.

is basically the one described for the case of the incumbent being from party A. So the payoff accruing to D in the second period is V_A. Therefore, D's problem in the first period simplifies to the choice of (x_1^A, x_1^B) for maximising

$$p_1^A u(\lambda x_1^A) + (1 - p_1^A) u(\lambda x_1^B) - c(x_1^A + x_1^B).$$

The careful reader will observe that this is essentially identical to D's problem when in period 2 and facing an incumbent from party B.

3.2 Equilibrium

The only "active" player in this setup is the donor D. The behaviour of the two parties is bound by pre-specified rules: change the candidate in the subsequent period with a fixed probability π in case of a loss in the initial period, otherwise continue with the same candidate. Also, for a dynastic party (party A), there is no change in the candidate in the sense that some member of the family contests the election under party A's banner.

This implies that the equilibrium is simply a profile of contributions by D in each period so that they maximise D's overall payoff summed over the two periods.

3.3 Baseline setup

As a natural benchmark, we begin with the case where the two parties are *ex-ante* symmetrical in all respects except that party A will be the dynastic party in the environment where a political dynasty *is* present. So we have $p^A = 1/2$ and $\rho = 1$.

We begin with the comparison of D's optimal choices in period 2 under the two different period 1 scenarios. First, we look at the case where party B is the winner in period 1.

3.3.1 Party B is the winner in period 1.

Recall, here D's objective function is

$$p_2^A u(\lambda x_2^A) + (1 - p_2^A) u(\lambda x_2^B) - c(x_2^A + x_2^B).$$

Our assumptions regarding $u(.)$ and $c(.)$ guarantee that the above is concave in (x_2^A, x_2^B). Hence, the FOCs are both necessary and sufficient for a maxima.

$$FOC(x_2^A) : \lambda p_2^A u'(\lambda x_2^A) + \frac{\partial p_2^A}{\partial x_2^A} \Big[u(\lambda x_2^A) - u(\lambda x_2^B) \Big] = c'(x_2^A + x_2^B)$$

$$FOC(x_2^B) : \lambda(1-p_2^A)u'(\lambda x_2^B) + \frac{\partial p_2^A}{\partial x_2^B}\Big[u(\lambda x_2^A) - u(\lambda x_2^B)\Big] = c'(x_2^A + x_2^B)$$

Also, we have

$$p_t^A = \omega \frac{x_t^A}{x_t^A + x_t^B} + (1-\omega)\frac{1}{2}$$

for $t = 1, 2$. This implies

$$\frac{\partial p_2^A}{\partial x_2^A} = \omega \frac{x_2^B}{(x_2^A + x_2^B)^2} > 0 \text{ and } \frac{\partial p_2^A}{\partial x_2^B} = -\omega \frac{x_2^A}{(x_2^A + x_2^B)^2} < 0.$$

Taken together, these yield:

$$\frac{\partial p_2^A}{\partial x_2^A} = -\frac{\partial p_2^A}{\partial x_2^B} . \frac{x_2^B}{x_2^A}.$$

Using this in the FOCs w.r.t x_2^A and x_2^B, we get:

$$\lambda p_2^A u'(\lambda x_2^A) - \alpha \frac{x_2^B}{x_2^A} = \lambda(1-p_2^A)u'(\lambda x_2^B) + \alpha \tag{1}$$

where $\alpha \equiv \frac{\partial p_2^A}{\partial x_2^B}[u(\lambda x_2^A) - u(\lambda x_2^B)]$.

Observation 1. *Given that party B wins in the first period, D will choose to donate the same amount to either party in the second period, i.e., $x_2^A = x_2^B \equiv x$. Also, x is the (unique) solution to $\lambda.u'(\lambda x) = 2c'(2x)$.*

This result is in line with one's intuition. In the final period — given that both A and B look symmetrical ex-ante — there is no reason to favour one over the other. Setting $x_2^A = x_2^B$ is optimal in the sense that D is smoothing "consumption" across the two possible outcomes.

Next we turn to the alternative scenario, namely, where party A is the first period winner.

3.3.2 Party A is the winner in period 1.

Recall, here D's objective function is

$$p_2^A u(\lambda x_2^A) + (1-p_2^A)[\pi u(\lambda\delta x_2^B) + (1-\pi)u(\lambda x_2^B)] - c(x_2^A + x_2^B).$$

Once again, our assumptions regarding $u(.)$ and $c(.)$ guarantee that the above is concave in (x_2^A, x_2^B). Hence, the FOCs are both necessary and sufficient for a maxima.

$$FOC(x_2^A) : \lambda p_2^A u'(\lambda x_2^A) + \frac{\partial p_2^A}{\partial x_2^A}\Big[u(\lambda x_2^A) - \pi u(\lambda\delta x_2^B) - (1-\pi)u(\lambda x_2^B)\Big] = c'(x_2^A + x_2^B)$$

$$FOC(x_2^B) : \lambda(1-p_2^A)\left[\pi\delta u'(\lambda\delta x_2^B)+(1-\pi)u'(\lambda x_2^B)\right]+\frac{\partial p_2^A}{\partial x_2^B}\left[u(\lambda x_2^A)-\pi u(\lambda\delta x_2^B)-(1-\pi)u(\lambda x_2^B)\right]$$
$$= c'(x_2^A + x_2^B)$$

Combining these with $\frac{\partial p_2^A}{\partial x_2^A} = -\frac{\partial p_2^A}{\partial x_2^B} \cdot \frac{x_2^B}{x_2^A}$, we get:

$$\lambda p_2^A u'(\lambda x_2^A) - \beta \frac{x_2^B}{x_2^A} = \lambda(1 - p_2^A)[\pi\delta u'(\lambda\delta x_2^B) + (1 - \pi)u'(\lambda x_2^B)] + \beta \tag{2}$$

where $\beta \equiv \frac{\partial p_2^A}{\partial x_2^B}[u(\lambda x_2^A) - \pi u(\lambda\delta x_2^B) - (1 - \pi)u(\lambda x_2^B)]$. Note, $\frac{\partial p_2^A}{\partial x_2^B} = -\omega . \frac{x_2^A}{(x_2^A + x_2^B)^2} < 0$.

Observation **2.** *Given that party A wins in the first period, D will choose to donate different amounts to the parties in the second period, i.e., $x_2^A \neq x_2^B$. Moreover, $x_2^A > x_2^B$.*

The intuition behind the above result is straight-forward. When faced with asymmetrical returns to similar contributions from the two options (A and B), there is a natural tendency for D to bias contributions towards the higher yielding option, namely, party A in this case. This relative favoritism towards A leads to a higher chance of success for A as compared to B which further reinforces D's incentives to back A more.

This leads to the next question as to what happens to *aggregate* contributions in the second period. Is it higher when party A wins in period 1 or when party B does? The following observation sheds some light on this.

Observation **3.** *Consider the aggregate contributions in the second period, i.e. $x_2^A + x_2^B$, by D. There exists a threshold $0 < \overline{\omega} < 1$ such that whenever $\omega \geq \overline{\omega}$, $x_2^A + x_2^B$ is higher in the scenario where party B wins in the first period as compared to where party A wins in the first period.*

Observation 3 informs us that as long as contributions have a sufficient impact on electoral outcomes, aggregate contributions will be higher for B being the first period winner rather than A. But why must that be? The basic idea is the following. When D knows that he can influence the outcome to a significant degree by varying the relative composition of his contributions, he will tend to raise x_2^A to a level higher than x_2^B. This clearly improves party A's chances of winning and for D reduces the likelihood that he receives a lower expected return (owing to the "depreciation factor" δ arising from meeting B's new candidate). The greater his influence on the odds (i.e, the higher is ω), the more willing he is to cut back on x_2^B. It is precisely *this* cutting back which leads to a lower aggregate contribution in this scenario as compared to when he faces the same candidates as in the first period (i.e., when

B wins in the first period).[13]

It is important to note that this threshold value of ω, namely $\overline{\omega}$ depends upon the factors which affect the expected return from contributing to the non-dynastic party (i.e., party B); specifically, on π and δ.

Corollary **1.** *The threshold $\overline{\omega}$ is lower for higher values of π and lower values of δ.*

It is clear that D enjoys a strictly higher payoff whenever the period 1 winner is from party B rather than from A. In other words, $V_B > V_A$.[14] It is immediate that the difference $V_B - V_A$ is increasing in π and decreasing in δ.

It is important to re-iterate that Observation 2 also applies to the case of a donor in period 2 in a setup devoid of political dynasties. This concludes the discussion of the various second period scenarios that may arise in this setup. We now turn to the analysis of the problem faced by D in the first period.

3.3.3 D's problem in period 1.

As mentioned earlier, in the situation where both parties are non-dynastic D's first period optimisation problem is the *same* as the one D faces when the incumbent is from party B in period 2. Therefore, the solution to this is the same as the one in that scenario. Hence, $x_1^A = x_1^B = x$ where x is as defined in Observation 1.

Now we turn to the situation where party A is associated with a political dynasty. Here, D chooses his actions — (x_1^A, x_1^B) — with the aim of maximising the payoffs over *both* periods. Therefore, the objective function is the following:

$$p_1^A[u(\lambda x_1^A) + V_A] + (1 - p_1^A)[u(\lambda x_1^B) + V_B] - c(x_1^A + x_1^B)$$

where $p_1^A = \omega \frac{x_1^A}{x_1^A + x_1^B} + (1 - \omega)\frac{1}{2}$ and $V_B > V_A$. Notice, here V_A and V_B are *independent* of the choice variables (x_1^A, x_1^B) and are thus treated as "parameters".

Once again our assumptions regarding $u(.)$ and $c(.)$ guarantee that the above is concave in (x_1^A, x_1^B). Hence, the FOCs are both necessary and sufficient for a maxima.

$$FOC(x_1^A) : \lambda p_1^A u'(\lambda x_1^A) + \frac{\partial p_1^A}{\partial x_1^A}\Big[u(\lambda x_1^A) - u(\lambda x_1^B) + V_A - V_B\Big] = c'(x_1^A + x_1^B)$$

[13] Of course, D's contribution to party A is higher in the former situation as compared to the latter but then owing to the strict concavity of $u(.)$ the rise in contributions towards A cannot fully compensate for the drop in contributions towards B.

[14] By construction, $V_B \geq V_A$. The two are equal only when $p_2^A = 1$ which is not possible since $\omega < 1$ and $p_A < 1$.

$$FOC(x_1^B): \lambda(1-p_1^A)u'(\lambda x_1^B) + \frac{\partial p_1^A}{\partial x_1^B}\Big[u(\lambda x_1^A) - u(\lambda x_1^B) + V_A - V_B\Big] = c'(x_1^A + x_1^B)$$

Let $\gamma \equiv \frac{\partial p_1^A}{\partial x_1^B}\Big[u(\lambda x_1^A) - u(\lambda x_1^B) + V_A - V_B\Big]$. Note, $\frac{\partial p_1^A}{\partial x_1^B} = -\omega.\frac{x_1^A}{(x_1^A+x_1^B)^2} < 0$.
So the FOCs can be combined to yield:

$$\lambda p_1^A u'(\lambda x_1^A) - \gamma\frac{x_1^B}{x_1^A} = \lambda(1-p_1^A)u'(\lambda x_1^B) + \gamma. \tag{3}$$

We now attempt to characterise D's optimal behaviour in the first period given that the different second period outcomes have been worked out. The main interest lies in understanding the following two issues: (i) which of the two parties are going to get favoured (in terms of contributions) and (ii) how much the aggregate contribution is in relation to when political dynasties are absent.
The following result sheds some light on both of these aspects.

Observation **4.** *In period 1 with a dynastic setup, D chooses (x_1^A, x_1^B) such that:*
(i) $x_1^A < x_1^B$, and
(ii) there exists $\tilde{\omega} < 1$ such that whenever $\omega \geq \tilde{\omega}$, we have $x_1^A + x_1^B \geq 2x$ where $2x$ is the total period 1 contribution under a no-dynasty scenario.

Of the two parts to Observation 4, the first one needs little explication. The result that party B's candidate gets relatively more support by D simply stems from the fact that the expected second period gains are *larger* for D when B's candidate wins in the first period; in more formal terms, $V_B > V_A$. In fact, higher this difference the larger the asymmetry in terms of D's contributions towards the two parties.[15]

The second part is (relatively) less obvious. This states that when contributions by D have "sufficient" influence over the electoral outcome, the aggregate first period contributions by D in a dynastic setup is (weakly) higher than when political dynasties are absent. Recall that in an environment sans political dynasties, D contributes *equally* to each party in period 1. Why so? This is because the second period payoff to D is the same regardless of which party wins in the first period. With party A being dynastic, were party B to win in period 1, the second period payoff to D is greater (since $V_B > V_A$). So when D has "sufficient" influence over the electoral outcome (i.e., $\omega \geq \tilde{\omega}$), he contributes a significant amount to B as he is assured that B will win with a large chance. In fact, this amount is high enough to ensure that the aggregate contribution exceeds that in the setup without political dynasties. Note, the differential gain driving this additional contribution towards B, namely $(V_B - V_A)$,

[15]As discussed earlier, a higher π and a lower δ exacerbates the difference between V_B and V_A.

is independent of (x_1^A, x_1^B) and thereby does not exhibit diminishing returns to a higher x_1^B; this helps to sustain the higher level of aggregate contribution here.

As regards the dependence of this threshold $\tilde{\omega}$ on the parameters of the model, an exact analogue of Corollary 1 applies. So, $\tilde{\omega}$ is lower for higher values of π and lower values of δ.[16]

Finally, we are in a position to attempt a period by period comparison of a dynastic setup versus a non-dynastic setup in terms of aggregate contributions.

3.3.4 Overall Contributions: A comparison.

Based on the analysis so far, we are able to comment on the following question:

Is there a greater amount of aggregate contributions in the scenario where a political party is associated with a dynasty than not?

We are able to provide an unambiguous answer in the situation where the ability of D's contributions to affect the electoral outcomes is not insignificant.

Proposition **1.** *The aggregate contributions, i.e., $\sum_{t=1}^{2}(x_t^A + x_t^B)$ are unambiguously higher in a situation where one party is associated with a political dynasty than when neither party is, provided $\omega \geq max\{\varpi, \tilde{\omega}\}$.*

The result stated in Proposition 1 offers a clear ranking of the aggregate contributions in the different scenarios as long as these contributions have some degree of influence on the election outcomes. In particular, if contributions would have *no* effect on the relative electoral success of a party then Proposition 1 is agnostic about the ranking.

This brings us to the crucial question as to what *really* determines the weight carried by these contributions on a party's success probability. And this — in turn — begs the question as to what the funds are used for. Presumably these funds are used to cover campaigning costs, organizing election rallies and the like; additionally, they may be used for vote-buying. Moreover, in our setup, we are looking at private donors. So these contributions will be more effective when there is little or no state-sponsored funding for elections. Hence, we expect these private contributions to play an important role in the context where either:

(i) there is minimal or no public funds available for covering election-related expenses by the parties, or

(ii) the requirements for disclosures by the donors and the recipients are non-existent/lax.

Both (i) and (ii) leave sufficient scope (and incentives) for private actors to engage in such implicit contracts with the contesting candidates. It is important to recognize that such is the state of affairs in most developing countries and, in particular, in India.

[16] The proof of this is omitted for brevity. It follows directly from observing that γ is increasing in $V_B - V_A$ and this latter term is increasing in π and decreasing in δ.

4 Asset holdings of politicians

Above we have argued that constituencies where a dynasty is present (and competing) are likely to receive more donations. In this section, we test empirically whether the asset holdings of politicians are consistent with this prediction.

As part of the "Right to Information Act" of 2003, it became mandatory for politicians to reveal their asset holdings before running for office. We make use of the candidate affidavits digitalized by the Shrug database (Asher et al., 2019), which gives us a snapshot of contestant's assets and liabilities just prior to the election. This data is perceived as to give a reasonably accurate picture of politicians' wealth level (see e.g. Fisman et al., 2014).

The affidavits data covers the time period 2004 to 2017, and spans three elections for most of our 8 states. We match the candidate names listed in the affidavits records to data on election outcomes taken from the Indian Elections Dataset (Agarwal et al., 2021). This matching gives us positions and vote shares for each candidate. In total, we match more than 80% of winners and runner-ups to an affidavits file. See Appendix XXX for more details on this matching.

We focus on the winner and runner-up in each constituency and start by running the following simple regression:

$$\text{Asinh(Net Assets)}_{ist} = \alpha_1 Dynastic_{ist} + \alpha_2 CompetingWithDynastic_{ist} + \sigma_{st} + \epsilon_{ist}, \quad (4)$$

where $\text{Asinh(Net Assets)}_{ijt}$ denotes the inverse hyperbolic sine of the net asset holding of politician i (running in a constituency in state s) just prior to the election in year t.[17] $Dynastic_{ist}$ denotes a candidate belonging to a dynasty, while $CompetingWithDynastic$ denotes a candidate that competes against a dynasty. We define this as someone that stood for election in a constituency where a dynastic candidate also competed and ended at least top five. σ_{jt} denotes state times election year fixed effects. We use fixed effects at this level, rather than at the AC level, as our main interest is how asset holdings of politicians vary *across* constituencies. See however Appendix XXX for a similar specification with AC-level fixed effects.

We present estimates of Equation 4 in Column 1 of Table 4. Strikingly, dynastic candidates ending top-2 in an election have almost twice as high asset values than other top-2 candidates within the same state (98%). Consistent with our theory, we also find that candidates competing with a dynastic candidate have about 20% higher assets value as compared to other top-2 candidates. In Columns 2-3, we restrict the sample to highly contested con-

[17] We use the inverse hyperbolic sine instead of log to deal with zeros and negative values. Following Fisman et al. (2014), we calculate net assets by subtracting total liabilities from the report asset values.

stituencies, defined as a win margin of no more than 10 percentage points (Column 2) or no more than 5 percentage points (Column 3). As can be seen, these restrictions barely change the estimate for dynastic candidates, while the difference between candidates that compete with dynasties versus other candidates become larger.

We next present estimates using net asset holdings prior to the *next* election as outcome, controlling for present asset holdings. We use the following specification:

$$
\begin{aligned}
&\text{Asinh(Net Assets Future)}_{ist} = \alpha_1 Dynastic_{ist} + \alpha_2 CompetingWithDynastic_{ist} \\
&+ \alpha_3 Incumbent_{ist} + \alpha_4 (Dynastic_{ist} \times Incumbent_{ist}) \\
&+ \alpha_5 (CompetingWithDynastic_{ist} \times Incumbent_{ist}) \\
&+ \beta \text{Asinh(Net Assets)}_{ist} + \sigma_{st} + \epsilon_{ist},
\end{aligned}
\tag{5}
$$

The specification captures differences among candidates in terms of changes in wealth and the α_3-coefficient can therefore be interpreted as an estimate of the private gain of holding office. Note that we only observe the affidavits of candidates that re-run in the next election.[18]

We present estimates of Equation 5 in Columns 4-6 of Table 4. Three points stand out. First, the estimates in Column 4 suggest that incumbent politicians obtain a 48 percentage points greater increase in their asset values over the election cycle as compared to runner-up candidates from the initial election. This translates into an annual asset growth premium of about 8 percent. In Columns 5-6, we restrict the sample to highly contested constituencies. This reduces the estimated incumbency advantage to 35% (Column 5, win margin $\leq$ 10%-points) and 20% (Column 6, win margin $\leq$ 5%-points). Note that this latter estimate is very similar to what Fisman et al. (2014) found using a sample of all large Indian states and one election cycle.

Second, the asset growth premium of being an incumbent is dwarfed by the premium of belonging to a dynasty. In the full sample, we find that dynastic candidates have 70% higher growth in their asset holdings, equivalent to an annual premium of as much as 11%. Note that the effect cannot be explained by dynasties having an additional incumbency advantage. In fact, the interaction between dynasties and incumbent has a negative sign, although it is far from being statistically significant.

Third, non-dynastic candidates that compete against dynastic candidates also seem to have higher asset accumulation as compared to other candidate. The point estimate in all samples is around 20%, but imprecisely estimated and not statistically different from zero (p-value of 0.240 in Column 4). As for dynasties, the positive coefficient is not driven by incumbents as the interaction term between incumbents and candidates competing with

[18]See Fisman et al. (2014) for a discussion on how this might affect the estimates.

dynasties numerically balance the main effect.

To sum up, the results in this section show that dynastic candidates are richer than other top candidate as measured by the net value of their asset holdings. They also have a greater assets accumulation over the election cycle. Moreover, and consistent with our theoretical prediction, we find that non-dynastic candidates competing with dynasties hold more assets than other top candidates.

Table 4: Regression: Net assets values

Dep. var:	IHS(Net Assets)			IHS(Net Assets Future)		
	All margins	Margin ≤10	Margin ≤5	All margins	Margin ≤10	Margin ≤5
	(1)	(2)	(3)	(4)	(5)	(6)
Dynastic	0.983***	1.033***	1.038***	0.701**	0.798***	0.509
	(0.124)	(0.150)	(0.136)	(0.236)	(0.255)	(0.315)
Competing with dynasty	0.194**	0.272**	0.337***	0.222	0.198	0.188
	(0.075)	(0.118)	(0.098)	(0.197)	(0.271)	(0.368)
Incumbent				0.464***	0.352***	0.204**
				(0.100)	(0.087)	(0.081)
Dynastic×Incumbent				-0.106	-0.266	0.164
				(0.242)	(0.284)	(0.331)
Competing with dynasty ×Incumbent				-0.213	-0.145	-0.100
				(0.305)	(0.454)	(0.535)
Observations	8175	4453	2418	3397	1965	1131
R^2	0.374	0.377	0.396	0.385	0.405	0.437

Robust standard errors clustered at state-year are shown in parentheses. *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent.

5 Pro-poor policy: NREGA

5.1 Background

The workfare program operates on a foundational principle of self-selection: every rural household interested in work is legally entitled to participate. Thus, on paper it is designed to be fully demand-driven, with the following procedures to aggregate demand for work through the political chain. The implementation of the program is delegated to Panchayati Raj institutions, which are local governance bodies comprising political councils at the district, block, and Gram Panchayat levels. Gram Panchayats, in collaboration with block and district administrations, are tasked with creating a list of potential projects before the start of each financial year. These projects are intended to serve as a reservoir of work opportunities, activated based on local demand. Households apply for job cards through their local Gram Panchayat and can request work as needed. These work requests are forwarded by the Gram Panchayat to the block administration and then to the district administration for final project approval.

In practice, however, the program functions more as a supply-driven initiative. This is evident in the significant variation in its implementation across and within states. A growing body of evidence suggests that this variation is largely attributable to unmet demand for employment (Dutta et al., 2014; Ministry of Rural Development, Government of India, 2012).

What drives the regional variation in program implementation? While the central government funds the majority of the program, state governments are required to cover a portion of project costs and administrative expenses. Consequently, states' fiscal capacity and overall administrative efficiency may therefore play a role. Previous research suggest that the incentives and motivations of local politicians and bureaucrats also play a significant role (see e.g., Gulzar and Pasquale, 2017; Gupta and Mukhopadhyay, 2016).

5.2 Data

We scraped data on local NREGA implementation from the *MGNREGA Public Data Portal*. This data source provides information at the level of Gram Panchayats for the financial year of 2011-12 and onwards. We mainly make use of the following variables: the total amount disbursed to workers' bank and post office accounts, the number of days worked and the number of workers. We use the total amount disbursed to workers as our main variable for NREGA implementation, as it covers both wages and workdays.

The NREGA data provides names of districts, blocks and Gram Panchayats but does not have Census identification numbers. We therefore match the dataset with the Census based on location names. We first match district and block names based on a combination of fuzzy matching and manual checking. We then match Gram Panchayat names within each district and block based on fuzzy matching. This procedure follows Asher and Novosad (2017); Gulzar and Pasquale (2017); Kjelsrud et al. (2020). In total, we are able to match around 75% of the Gram Panchayats in the NREGA data to the Census. Having obtained Census village identifiers for the NREGA dataset, we merge the dataset to geo-coded maps of the all Indian villages from the ML InfoMap. We next overlay this with similar maps of pre- and post-Delimitation ACs. For each matched Gram Panchayats in the NREGA data we thus have information on which Census villages they cover, and which pre- and post-Delimitation AC they belongs to. This information is necessary for our identification (see below). Finally, we construct a set of control variables using the Census of India for 2001. This dataset includes basic population characteristics and information about a large number of public amenities.

5.3 Identification and specifications

We identify the effect of the *presence* of dynasties by exploiting variation induces by the Delimitation of 2008.

The Delimitation changed boundaries of ACs within districts (as well as Parliamentary constituencies within states). The aim of the redistricting was to equalize population across constituencies and to reserve constituencies for scheduled castes and scheduled tribes in proportion to updated measures of their population shares. The whole process started in mid-2004 and was based on population characteristics from the Census of India 2001. The redistricting was lead by an independent three-member commission. The commission was assisted by ten associate members in each state, consisting of five MPs and five MLAs. The associate members had no formal voting power. Previous research suggest that the boundary changes were politically neutral and find no evidence of gerrymandering (Bardhan et al., 2020; Iyer and Reddy, 2013; Kjelsrud et al., 2020).

We explore the boundary changes to identify plausibly causal estimates of the effect of the *presence* of political dynasties. We illustrate our approach in Figure 3. Imagine two pre-Delimitation ACs, A and B, and that there is a political dynasty present in A, but not in B. Suppose next that the Delimitation creates two new ACs, A' and B'. Our approach is to compare the implementation of NREGA in the two areas marked with the hatched lines in the figure. Prior to the Delimitation, both areas belonged to the same AC and they were exposed to the same political dynasty. After the Delimitation, however, the areas are represented by different MLAs, and plausibly, only one areas is exposed to the dynasty.

We do not know ex ante which post-Delimitation AC the dynasty from A will run in. As this is an endogenous choice, we "allocate" dynasties to the post-Delimitation ACs with the greatest population overlap with the former ACs (A' in the illustration, if we assume that the population is uniformly distributed). Given that the boundary changes plausibly were orthogonal to factors influencing the implementation of NREGA, this approach identifies credible evidence of the effect of the presence of a dynasty.

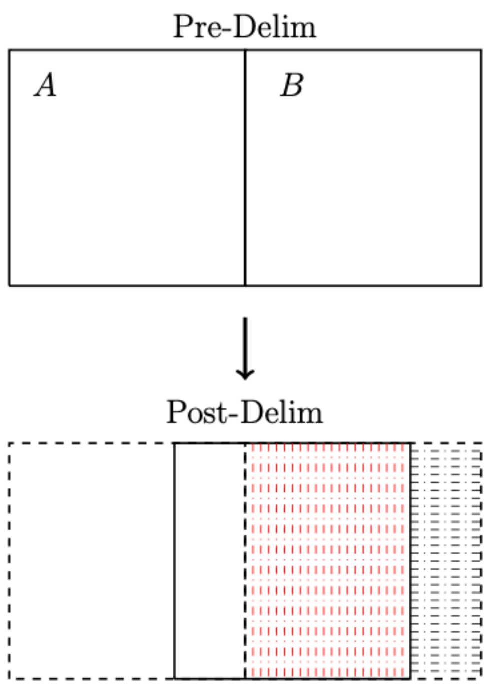
We implement this using a regression with fixed effects for pre-Delimitation ACs:

$$NREGA_{ijkt} = \beta PredDynastic_k + \sigma_j + \sigma_t + X'_{ij} + \epsilon_{ijkt}, \tag{6}$$

where σ_j denotes the pre-Delimitation AC fixed effects, $NREGA_{ijk}$ captures the local NREGA implementation in Gram Panchayat i, from pre-Delimitation AC j and post-Delimitation AC k, in year t; and $PredDynastic_k$ is a binary variable capturing whether post-Delimitation AC k is predicted to by dynastic based on having the largest population overlap with the constituency the dynasty competition in prior to the Delimitation. σ_t denotes a set of fixed effects for financial years (or possibly year×state fixed effects), while X'_{ij} is a vector of Gram Panchayat level controls constructed from the Census of India. The coefficient of interest is β, which captures the effect of the presence of the dynasty on NREGA implementation. Standard errors are clustered at the level of pre-×post-Delimitation ACs.

We pin down the time frame of the analysis based on the date of the first post-Delimitation in each state . In the baseline specification, we use the four years preceding this election (i.e. roughly the election term). For instance, in Bihar the first post-election took place in October-November 2010. The sample for Bihar therefore covers the years 2011-2012 to 2014-2015. See Table A1 for a full list of election years for each states in our sample.

Figure 3: Illustration of the identification



Note: The figure illustrates our main identification.

5.4 Validation

We validate the above empirical approach in two ways. First, we test whether $PredDynastic_k$ predicts in which constituency the dynasties compete after the Delimitation. To do this, we run a regression similar to Equation 6, but by changing the outcome to the binary variable, $PostDynasticAC_{ijks}$, denoting whether the dynasty run in AC j:

$$PostDynasticAC_{ijks} = \beta_1 PredDynastic_{ijks} + \sigma_{js} + X'_{ijs} + \epsilon_{ijks} \tag{7}$$

Estimates of Equation 7 are presented in Table 5. As can be seen, the binary variable capturing the population overlap strongly predicts were the dynasty ends up running after the boundary changes.

Our identification relies on the assumption that $PredDynastic_k$ is unrelated to unobserved factors determining NREGA implementation once we include the pre-Delimitation constituency fixed effects. For this to be the case, the allocation of villages to new ACs should be as good-as-random (again, within pre-Delimitation constituencies). As a second validation, we therefore look for potential differences in Gram Panchayats. We run two different balancing tests, presented in Columns 1 and 2 of Table 6.

We first test whether villages that are redistricting differ from those that are not, using the following specification:

$$y_{vjs} = \beta_1 ReDistricted_{ijks} + \sigma_{js} + \epsilon_{ijks}, \tag{8}$$

where y_{vjs} denotes a Gram Panchayat level characteristics from the Indian Census and $ReDistricted_{ijks}$ is binary variable denoting whether the Gram Panchayat was redistricted. We define redistricting as ending up in a post-Delimitation AC other than the majority of the population in the pre-Delimitation AC. σ_{js}, as before, denotes a set of pre-Delimitation AC fixed effects.

As a second test, closer to our main specification, we check whether Gram Panchayats that are predicted to be in a dynastic AC after the Delimitation differ from the other Gram Panchayat in their pre-Delimitation AC, using the following regression:

$$y_{vjs} = \beta_1 PredDynastic_{ijks} + \sigma_{js} + \epsilon_{ijks}. \tag{9}$$

We present the estimates in Table 6. Column 1 shows estimates of Equation 8. Only one of the coefficients are statistically significant at a 10% significant level (access to Primary Health Centres). We also check whether the Census covariates are jointly significant. To do this, we regress the redistricting dummy on all the controls in addition to the fixed effects.

The F-test from this regression is 1.32, which implies that the rich set of controls are unable to predict which Gram Panchayats that were redistricted (within the pre-Delimitation ACs).

Column 2 similarly presents estimates of Equation 9. Most coefficients are insignificant and close to zero. Only 2 out of the 13 estimated coefficients are statistically significant at the 10%-level (population share of SCs/STs and access to paved roads). We also test the joint significance by placing *PredDynastic* on the left-hand side and the controls on the right-hand side. This leads to a F-statistics of 1.29, which implies that the full set of covariates are unable to predict which Gram Panchayats that end up in areas likely exposed to dynasties (again within the pre-Delimitation ACs).

Table 5: Actual versus predicted dynastic ACs

Dep.var.:	Dynastic AC Post	
	(1)	(2)
Predicted Dynastic	0.397***	0.396***
	(0.087)	(0.087)
Observations	41275	41275
R^2	0.071	0.070
Controls		Y
Pre-AC FEs	Y	Y

Robust standard errors clustered at Pre-Delim ACs are shown in parentheses.
*** significant at 1 percent, ** significant at 5 percent,
* significant at 10 percent.

5.5 Results

Table 7 presents our main result on NREGA implementation, using Equation 6. In the odd numbered columns we do not add the Gram Panchayats Census controls, which we do in the even numbered columns.

As a first outcome, we use a binary variable denoting whether or not Gram Panchayats had *any* NREGA project during the four years following the election (Columns 1-2). The estimated coefficients suggest that areas exposed to a dynasty have a 3.1 percentage points smaller chance of receiving any NREGA project over the election cycle (p-value=0.11). Notice that almost all Gram Panchayats had at least one project (sample mean of 0.94), leading to limited variation in this outcome.

We next explore the intensive margin by restricting the sample to Gram Panchayats with at least some NREGA work and use three measures of the amount of work provided. In

Table 6: Balance test

Dep.var.:	ReDistricted (.3168) (1)	Predicted Dynastic (.0475) (2)	Mean (3)
Log population	-0.0156	0.0052	8.0166
	(0.0100)	(0.0423)	
Population share ST/SC	-0.0038	0.0167*	0.2662
	(0.0034)	(0.0090)	
Literacy rate females	0.0016	0.0021	0.4299
	(0.0016)	(0.0061)	
Literacy rate males	0.0009	0.0015	0.6078
	(0.0015)	(0.0057)	
Primary school	0.0006	-0.0059	0.9095
	(0.0026)	(0.0072)	
Middle school	-0.0094	0.0084	0.4524
	(0.0057)	(0.0191)	
Secondary school	-0.0048	-0.0051	0.2132
	(0.0042)	(0.0179)	
PHC	-0.0046*	-0.0025	0.0837
	(0.0026)	(0.0085)	
PHC sub-centre	0.0089	-0.0119	0.2402
	(0.0061)	(0.0140)	
Electricity	-0.0024	-0.0130	0.6884
	(0.0065)	(0.0280)	
Bus/train connection	0.0018	-0.0160	0.6839
	(0.0074)	(0.0222)	
Paved road	0.0071	-0.0300*	0.8165
	(0.0059)	(0.0160)	
N	41275		

Robust standard errors clustered at Pre-Delim ACs are shown in parentheses. *** significant at 1 percent,
** significant at 5 percent, * significant at 10 percent.

Columns 3-4 we use the log of total amounts dispersed to workers bank account as outcome. We find a negative effect of 22%, significant at the 1%-level. Similarly, Columns 5-8 show that the exposure to a dynasty reduces the number of NREGA working days by 20% (Column 6) and the number of NREGA workers by 15% (Column 8). All estimates are practically similar whether or not we included the Gram Panchayat-level controls constructed from the Indian Census. This is not surprising, given the seemingly balanced sample (see Table 6).

Table 7: NREGA and the presence of dynasties

Dep.var.:	Any NREGA		Log Amounts		Log Days		Log Persons	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PredDynastic	-0.029	-0.031	-0.214**	-0.223***	-0.193**	-0.204**	-0.142*	-0.154**
	(0.019)	(0.020)	(0.085)	(0.081)	(0.092)	(0.087)	(0.074)	(0.070)
Observations	41275	41275	38600	38600	38933	38933	38934	38934
R^2	0.000	0.017	0.001	0.056	0.001	0.063	0.001	0.100
Controls		yes		yes		yes		yes
Dep.var.mean	0.94	0.94	14.09	14.09	9.30	9.30	6.00	6.00

All regressions includes Pre-Delimitation AC fixed effects. Robust standard errors clustered at Pre-Delim ACs are shown in parentheses. *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent.

5.6 Robustness checks

6 Conclusion

In this paper, we examine the economic effects of a relatively underexplored informal institution — political dynasties, which are prevalent in many democracies. We begin by analyzing this phenomenon through a formal lens, using a simple two-period model that deliberately abstracts from any differences in competence between dynastic and non-dynastic politicians.[19] The model demonstrates that as long as elites prioritize long-term prospects and dynasties have a greater ability to position their members in elections, the prevalence of political dynasties leads to higher levels of clientelism.

We then evaluate the empirical relevance of our theory by conducting an extensive data collection effort. This involved identifying familial links between politicians serving in the state legislative assemblies of eight major Indian states. The resulting dataset represents a key contribution of this paper.

[19]This can be easily incorporated into our framework by a suitable choice of parameters.

By integrating our dataset with publicly available information on politicians' asset holdings, we demonstrate that the top contestants in constituencies where political dynasties are active have higher wealth level. This pattern is not solely driven by dynastic candidates being wealthier. Remarkably, non-dynastic candidates who lose to dynastic competitors are significantly wealthier than those who lose to other non-dynastic candidates.

Finally, we examine whether the presence of dynasties shifts policy in a pro-rich direction. To do this, we focus on a quintessential pro-poor policy: the public workfare program MNREGA. To identify plausible causal effects, we exploit the electoral boundary changes introduced by the Delimitation Act of 2008. Our findings reveal a clear pattern: areas more likely to be exposed to a political dynasty following the Delimitation experience significantly less MNREGA employment. This is evident in lower total payments, fewer days worked, and a reduced number of persons employed.

References

Agarwal, A., N. Agrawal, S. Bhogale, S. Hangal, F. R. Jensenius, M. Kumar, C. Narayan, and B. U. Nissa (2021): "TCPD Indian Elections Data v2. 0. Trivedi Centre for Political Data, Ashoka University," .

Asher, S., T. Lunt, R. Matsuura, and P. Novosad (2019): "The socioeconomic high-resolution rural-urban geographic dataset on India (SHRUG)," *URL: https://doi. org/10.7910/DVN/DPESAK.*

Asher, S. and P. Novosad (2017): "Politics and Local Economic Growth: Evidence from India," *American Economic Journal: Applied Economics*, 9, 229–273.

Bardhan, P. and D. Mookherjee (2006): "Decentralization, Corruption and Government Accountability," in *International Handbook on the Economics of Corruption*, Cheltenham, UK: Edward Elgar Publishing.

Bardhan, P. K., S. Mitra, D. Mookherjee, and A. Nath (2020): "How Do Voters Respond to Welfare vis-à-vis Public Good Programs? An Empirical Test for Clientelism," Federal Reserve Bank of Minneapolis.

Besley, T. and M. Reynal-Querol (2017): "The logic of hereditary rule: theory and evidence," *Journal of Economic Growth*, 22, 123–144.

Chandra, K. (2016): "Democratic dynasties," *Democratic dynasties: State, party and family in contemporary Indian politics*, 12–55.

Dal Bó, E., P. Dal Bó, and J. Snyder (2009): "Political dynasties," *The Review of Economic Studies*, 76, 115–142.

Dutta, P., R. Murgai, M. Ravallion, and D. Van de Walle (2014): *Right to Work? Assessing India's Employment Guarantee Scheme in Bihar*, The World Bank.

Fisman, R., F. Schulz, and V. Vig (2014): "The private returns to public office," *Journal of Political Economy*, 122, 806–862.

George, S. E. (2020): "Like Father, Like Son? The Effect of Political Dynasties on Economic Development," Mimeo.

Gulzar, S. and B. J. Pasquale (2017): "Politicians, Bureaucrats, and Development: Evidence from India," *American Political Science Review*, 111, 162–183.

Gupta, B. and A. Mukhopadhyay (2016): "Local Funds and Political Competition: Evidence from the National Rural Employment Guarantee Scheme in India," *European Journal of Political Economy*, 41, 14–30.

Iyer, L. and M. Reddy (2013): *Redrawing the Lines: Did Political Incumbents Influence Electoral Redistricting in the World's Largest Democracy?*, Harvard Business School.

Kjelsrud, A., K. Moene, and L. Vandewalle (2020): "The Political Competition over Life and Death – Evidence from Infant Mortality in India," Graduate Institute Geneva Working Paper.

Maiorano, D. (2014): "The Politics of the Mahatma Gandhi National Rural Employment Guarantee Act in Andhra Pradesh," *World Development*, 58, 95–105.

Ministry of Rural Development, Government of India (2012): *MGNREGA Sameeksha, An anthology of research studies on the Mahatma Gandhi National Rural Employment Guarantee Act, 2005, 2006-2012*, New Delhi: Orient BlackSwan.

A Details on model

Proof. [OBSERVATION 1.] The sign of α depends upon the relative sizes of x_2^A and x_2^B. Setting $x_2^A = x_2^B$ satisfies the FOCs w.r.t x_2^A and x_2^B. To see why, observe the following: $x_2^A = x_2^B$ implies $\alpha = 0$. To satisfy equation (1), it must be that $p_2^A = 1 - p_2^A$. This is indeed true when $p^A = 1/2$ and $\rho = 1$ since $x_2^A = x_2^B$ leads to $p_2^A = 1/2$.

The FOC w.r.t x_2^A for $x_2^A = x_2^B \equiv x$ yields $\lambda.u'(\lambda x) = 2c'(2x)$. The LHS of this equation is falling in x and the RHS is rising. Moreover $\lambda.u'(0) > 2c'(0)$, thus completing the proof. $\square$

Proof. [OBSERVATION 2.] The proof proceeds in two steps.

Step 1: $x_2^A \neq x_2^B$.

Suppose $x_2^A = x_2^B$. Call the common value z. Clearly, $p_2^A = 1/2$. Also, $\beta < 0$. Therefore, equation (2) becomes:

$$(\lambda/2).u'(\lambda z) - \beta = (\lambda/2).[\pi\delta u'(\delta\lambda z) + (1-\pi)u'(\lambda z)] + \beta.$$

But $\beta < 0$ and $u'(\lambda z) \geq \pi\delta u'(\delta\lambda z) + (1-\pi)u'(\lambda z)$ contradicts the equality above. This completes Step 1.

Step 2: $x_2^A > x_2^B$.

Suppose not. In light of Step 1, this implies x_2^A must be strictly lower than x_2^B.

Let $x_2^A = a$ and $x_2^B = b$ where $a < b$. Then $p_2^A = \frac{1}{2} - \omega.\frac{(b-a)}{2(a+b)} < 1/2$. Consider the following deviation by D. Suppose now D sets $x_2^A = b$ and $x_2^B = a$. Clearly, this should not yield an increase in D's payoff. Denote D's original period 2 payoff (from $x_2^A = a$ and $x_2^B = b$) by $\Gamma(a,b)$ and the one from the deviation by $\Gamma(b,a)$. Note,

$$\Gamma(b,a) = \left(\frac{1}{2} + \omega.\frac{(b-a)}{2(a+b)}\right)u(\lambda b) + \left(\frac{1}{2} - \omega.\frac{(b-a)}{2(a+b)}\right)[\pi u(\lambda\delta a) + (1-\pi)u(\lambda a)] - c(a+b)$$

and

$$\Gamma(a,b) = \left(\frac{1}{2} - \omega.\frac{(b-a)}{2(a+b)}\right)u(\lambda a) + \left(\frac{1}{2} + \omega.\frac{(b-a)}{2(a+b)}\right)[\pi u(\lambda\delta b) + (1-\pi)u(\lambda b)] - c(a+b).$$

Hence,

$$\Gamma(b,a) - \Gamma(a,b) = \left(\frac{1}{2} + \omega.\frac{(b-a)}{2(a+b)}\right).\pi[u(\lambda b) - u(\lambda\delta b)] - \left(\frac{1}{2} - \omega.\frac{(b-a)}{2(a+b)}\right).\pi[u(\lambda a) - u(\lambda\delta a)].$$

Note, $u(\lambda z) - u(\lambda\delta z)$ is (weakly) increasing in z given our assumption $\tau u'(\tau x) \leq u'(x)\ \forall\tau \in (0,1)$ and $x > 0$. Therefore, $u(\lambda b) - u(\lambda\delta b) \geq u(\lambda a) - u(\lambda\delta a)$ and we have $\Gamma(b,a) - \Gamma(a,b) > 0$. This contradicts that $x_2^A = a$ and $x_2^B = b$ is optimal. Hence, Step 2 is established thus completing the proof. $\square$

Proof. [OBSERVATION 3.] Start with the scenario where party A wins in the first period. Denote the optimal choice of D's contribution here by (x_2^A, x_2^B).

Suppose $x_2^A + x_2^B \geq 2x$ where x is defined as in Observation 1. Hence, $c'(x_2^A + x_2^B) \geq c'(2x)$ by the strict convexity and increasing nature of $c(.)$. So from the FOCs w.r.t x_2^B, we have :

$$\lambda(1-p_2^A)[\pi\delta u'(\lambda\delta x_2^B) + (1-\pi)u'(\lambda x_2^B)] + \beta \geq \lambda.(1/2).u'(\lambda x).$$

Given that $p_2^A > 1/2$ and $\beta < 0$, this implies $x_2^B < x$.

Hence, $x_2^A > x$ to maintain $x_2^A + x_2^B \geq 2x$. From the FOCs, we also get:

$$\lambda p_2^A u'(\lambda x_2^A) + \lambda(1-p_2^A)[\pi\delta u'(\lambda\delta x_2^B) + (1-\pi)u'(\lambda x_2^B)] + \beta(1 - \frac{x_2^B}{x_2^A}) = 2c'(x_2^A + x_2^B).$$

Given that $\beta < 0$ and $x_2^A > x_2^B$, we have:

$$\lambda p_2^A u'(\lambda x_2^A) + \lambda(1-p_2^A)[\pi\delta u'(\lambda\delta x_2^B) + (1-\pi)u'(\lambda x_2^B)] > 2c'(x_2^A + x_2^B).$$

Notice, $u'(\lambda x_2^B) \geq \pi\delta u'(\lambda\delta x_2^B) + (1-\pi)u'(\lambda x_2^B)$ by our assumption on $u(.)$. Hence,

$$\lambda p_2^A u'(\lambda x_2^A) + \lambda(1-p_2^A)u'(\lambda x_2^B) > 2c'(x_2^A + x_2^B).$$

Rewrite x_2^A as $x+\psi$ and x_2^B as $x-\epsilon$. In order to have $x_2^A + x_2^B \geq 2x$, we require $\psi \geq \epsilon$. So the LHS of the relation above can be written as:

$$\lambda p_2^A u'(\lambda(x+\psi)) + \lambda(1-p_2^A)u'(\lambda(x-\epsilon)).$$

Now compare this with $\lambda p_2^A u'(\lambda x) + \lambda(1-p_2^A)u'(\lambda x)$. For $\psi \geq \epsilon$, there is clearly an unique threshold for p_2^A strictly lower than 1 (call it $\overline{p}$) such that any $p_2^A \geq \overline{p}$, we have:

$$\lambda p_2^A u'(\lambda x) + \lambda(1-p_2^A)u'(\lambda x) \geq \lambda p_2^A u'(\lambda(x+\psi)) + \lambda(1-p_2^A)u'(\lambda(x-\epsilon)).$$

Recall equation (2) which states

$$\lambda p_2^A u'(\lambda x_2^A) - \beta\frac{x_2^B}{x_2^A} = \lambda(1-p_2^A)[\pi\delta u'(\lambda\delta x_2^B) + (1-\pi)u'(\lambda x_2^B)] + \beta.$$

Also, $p_2^A = \omega\frac{x_2^A}{x_2^A+x_2^B} + (1-\omega)\frac{1}{2}$. Consider an increase in the parameter ω. All else constant, this leads to a rise in the LHS since β falls and p_2^A rises. Moreover, the RHS falls owing to the same factors. So for equation (2) to hold, x_2^B must fall. Moreover, as $\omega \to 1$ we have $x_2^B \to 0$. So, as $\omega \to 1$ we have $p_2^A \to 1$. Given the continuity of p_2^A in ω, we know there exists a threshold for ω (call it $\overline{\omega}$) such that $\omega \geq \overline{\omega}$ implies $p_2^A \geq \overline{p}$. Additionally, the threshold is unique owing to the strictly monotonic relation between ω and p_2^A. Also, $\overline{\omega} < 1$ since $\overline{p} < 1$.

Hence for any $\omega \geq \overline{\omega}$, we have:

$$\lambda p_2^A u'(\lambda x) + \lambda(1-p_2^A)u'(\lambda x) \geq \lambda p_2^A u'(\lambda(x+\psi)) + \lambda(1-p_2^A)u'(\lambda(x-\epsilon)).$$

Recall that $\lambda p_2^A u'(\lambda x)+\lambda(1-p_2^A)u'(\lambda x) = 2c'(2x)$ and $\lambda p_2^A u'(\lambda(x+\psi))+\lambda(1-p_2^A)u'(\lambda(x-\epsilon)) > 2c'(x_2^A+x_2^B)$.

So, we infer $c'(2x) > c'(x_2^A + x_2^B)$. Hence, $2x > x_2^A + x_2^B$ which contradicts our initial supposition that $x_2^A + x_2^B \geq 2x$ and thus completes the proof. □

Proof. [COROLLARY 1.] By equation (2), we have:

$$\lambda p_2^A u'(\lambda x_2^A) - \beta\frac{x_2^B}{x_2^A} = \lambda(1-p_2^A)[\pi\delta u'(\lambda\delta x_2^B) + (1-\pi)u'(\lambda x_2^B)] + \beta.$$

A higher value of π ceteris paribus implies a lower β and hence a fall in the RHS while generating a rise in the LHS. For equalisation of the two sides, there has to be a fall in x_2^B which leads to a rise in p_2^A. Hence, a given level of ω is associated with a higher level of p_2^A for a higher π. This brings down the threshold $\overline{\omega}$ corresponding to $\overline{p}$.
Analogous arguments apply for lower values of δ. □

Proof. [OBSERVATION 4.]

Part (i):

Suppose $x_1^A = x_1^B$. This implies $\gamma > 0$ as $\frac{\partial p_1^A}{\partial x_1^B} < 0$ and $V_B > V_A$. Also, $p_1^A = 1/2$. This clearly violates equation (3). Hence, $x_1^A \neq x_1^B$.

Now suppose $x_1^A > x_1^B$. Let $x_1^A = a$ and $x_1^B = b$ where $a > b$. Then $p_1^A = \frac{1}{2} + \omega.\frac{(a-b)}{2(a+b)} > 1/2$. Consider the following deviation by D. Suppose now D sets $x_1^A = b$ and $x_1^B = a$. Clearly, this should not yield an increase in D's payoff.

Denote D's original period 1 payoff (from $x_1^A = a$ and $x_1^B = b$) by $\Gamma(a, b)$ and the one from the deviation by

$\Gamma(b,a)$. Note,

$$\Gamma(b,a) = \left(\frac{1}{2} - \omega.\frac{(a-b)}{2(a+b)}\right)[u(\lambda b) + V_A] + \left(\frac{1}{2} + \omega.\frac{(a-b)}{2(a+b)}\right)[u(\lambda a) + V_B] - c(a+b)$$

and

$$\Gamma(a,b) = \left(\frac{1}{2} + \omega.\frac{(a-b)}{2(a+b)}\right)[u(\lambda a) + V_A] + \left(\frac{1}{2} - \omega.\frac{(a-b)}{2(a+b)}\right)[u(\lambda b) + V_B] - c(a+b).$$

Taking the difference between them yields:

$$\Gamma(b,a) - \Gamma(a,b) = \omega.\frac{(a-b)}{(a+b)}.[V_B - V_A] > 0.$$

This contradicts that $x_2^A = a$ and $x_2^B = b$ is an optimal choice. So $x_1^A < x_1^B$ is established.

Part (ii):

Suppose $x_1^A + x_1^B < 2x$. Then $c'(2x) > c'(x_1^A + x_1^B)$.
By the FOC wrt x_1^B, $\lambda(1-p_1^A)u'(\lambda x_1^B) + \gamma < c'(2x)$.
Given that $x_1^A < x_1^B$, we have $p_1^A < 1/2$ and $\gamma > 0$. By our assumption on $u(.)$,

$$\lambda u'(2\lambda x) \geq \lambda.\frac{1}{2}u'(\frac{1}{2}.2\lambda x) = c'(2x).$$

Therefore, given that $\gamma > 0$ there exists a $\underline{p} > 0$ such that for any $p_1^A \leq \underline{p}$, we have:

$$\lambda(1-p_1^A)u'(\lambda 2x) + \gamma \geq \lambda u'(2\lambda x).$$

Hence, for any $p_1^A \leq \underline{p}$ we have: $\lambda(1-p_1^A)u'(\lambda 2x)+\gamma \geq c'(2x)$. And so for any $p_1^A \leq \underline{p}$, $\lambda(1-p_1^A)u'(\lambda x_1^B)+\gamma < c'(2x)$ implies $x_1^B > 2x$.

Now consider equation (3) which states $\lambda p_1^A u'(\lambda x_1^A) - \gamma\frac{x_1^B}{x_1^A} = \lambda(1-p_1^A)u'(\lambda x_1^B) + \gamma$.
An increase in ω, ceteris paribus, leads to a rise in the RHS since p_1^A falls and γ rises. For the same reasons, the LHS decreases. Hence, x_1^A must fall to restore equality. Moreover, as $\omega \to 1$, $x_1^A \to 0$ and hence $p_1^A \to 0$. By the continuity of p_1^A, there is a threshold $\tilde{\omega} < 1$ such that whenever $\omega \geq \tilde{\omega}$, we have $p_1^A \leq \underline{p}$. Therefore we can claim that whenever $\omega \geq \tilde{\omega}$, $\lambda(1-p_1^A)u'(\lambda x_1^B) + \gamma < c'(2x)$ implies $x_1^B > 2x$. But this implies $x_1^A + x_1^B > 2x$ which contradicts our initial supposition and hence completes the proof. □

Proof. [PROPOSITION 1.] Start with the situation where neither part A nor B are associated with a political dynasty. Consider D's problem in the second period. Either party A or party B would have won in period 1. So w.l.o.g. D is facing the problem as when party A wins in period 1; so the solution is given in Observation 2. Furthermore, D's problem in period 1 is the same as one the facing (a hypothetical) D in a dynastic setup in period 2 where party B has won in period 1. So $x_2^A = x_2^B \equiv x$ where x is the (unique) solution to $\lambda.u'(\lambda x) = 2c'(x)$.

Now consider the case where party A is associated with a dynasty while party B is not. Here D's problem in the second period depends upon which party won in the previous period. With probability p_1^A it is party A, which by Observation 4, is less than 1/2. This yields to D the payoff V_A and is identical to the period 1 payoff of D in a situation where neither part A nor B are associated with a political dynasty. Also, probability $(1-p_1^A) > 1/2$ it is party B, which yields to D the payoff V_B $(> V_A)$. By Observation 3 we know that the contributions in the latter situation exceed that under the former as long as $\omega \geq \overline{\omega}$. This establishes that the aggregate contributions in period 2 is higher in a dynastic setup than otherwise.
To compare D's first period contribution under a dynastic setup with one without any dynasty, one needs to consult Observation 4.

Combining the above completes the proof. □

B Extra tables

Table A1: Election years by states

State	Election year
Bihar	2000, 2005, 2010, 2015, 2020
Karnataka	1999, 2004, 2008, 2013, 2018
Kerala	2001, 2006, 2011, 2016, 2021
West Bengal	2001, 2006, 2011, 2016, 2021
Punjab	2002, 2007, 2012, 2017, 2022
Haryana	2000, 2005, 2009, 2014, 2019
Tamil Nadu	2001, 2006, 2011, 2016, 2021
Rajasthan	1998, 2003, 2008, 2013, 2018