**Electoral Uncertainty, Income Inequality and the "Middle Class"**[1]

BY ANIRBAN MITRA AND SHABANA MITRA

Preliminary and Incomplete

September 2012
(First draft: July 2011)

ABSTRACT

We investigate how increased electoral competition — by influencing the proposed policies of competing parties — affects (i) income inequality and (ii) the size of a "middle-class" in society. In our model, parties can invest in different public goods which translate into higher incomes for a heterogeneous electorate. Our theory shows that the effect of close elections on either inequality or polarization crucially depends upon the degree of concavity of the income–from–investment function. Specifically, if the degree of concavity is "low" then close elections *reduce* inequality; the relation is *reversed* if the degree of concavity is "high". We check for these relationships using data from the Indian national elections which are combined with household-level consumption expenditure data rounds from NSSO (1987-88 and 2003-04) to yield a panel of Indian districts. We find that districts which have experienced tight elections exhibit lower inequality and polarization. This robust empirical finding, in conjunction with our theory, suggests that the marginal gains from investing in public goods in India is fairly significant.

*JEL codes:* D72, D78, O20

*Keywords:* Income distribution, polarization, political economy, targeting.

# 1 Introduction

In any society, uncertainty over electoral outcomes is likely to influence the proposed policies of competing political parties, as long as they are office–motivated. Recognizing that these policies are able to differentially impact the incomes of an income–wise heterogeneous electorate, leads us to ask the following questions. How do close elections — via

[1] We are grateful to James Foster and Debraj Ray for their support and encouragement. We would like to thank Alejandro Corvalan, Bård Harstad, Kalle Moene, Natalija Novta, Stephen Smith, Kjetil Storesletten and participants at the Development Reading groups at GWU and NYU, and the ESOP seminar participants at the University of Oslo for providing useful comments and suggestions. Of course, all remaining errors are solely ours.

Corresponding author: Anirban Mitra (University of Oslo). Email: anirban.mitra@econ.uio.no

their effect on equilibrium policies — affect the income distribution in society? In particular, does greater electoral uncertainty *reduce* or *exacerbate* existing income disparities? Also, does this lead to an increase in the size of the "middle class" or not? These are precisely the questions we attempt to address in this paper by first constructing a theoretical framework and then conducting a related empirical exercise with data from India.

There is a strong relationship between the size of the middle class and the degree of income polarization in society (see Esteban and Ray (2010) for a comprehensive discussion). The degree of income polarization is a measure of the extent of clustering in society along income lines. In particular, a high degree of income polarization is suggestive of society dominated by two income groups — the "haves" and the "have-nots" and thus a *smaller* middle-class.[2] This is, in principle, quite different from income inequality and from an empirical standpoint most measures of the two concepts often diverge over various ranges.

The literature so far has focused on how resources are targetted to districts which are "non-partisan" — and hence electorally more unpredictable — as opposed to districts which are strongly inclined towards some party. Models of political competition which have directly addressed questions of this nature (see for e.g., Lindbeck and Weibull (1987), Dixit and Londregan (1996, 1998), etc.) have generally concluded that "swing" districts get more targeted resources in the aggregate. Such theoretical findings have been empirically investigated. For instance, Arulampalam et al (2009) find evidence, in the case of India, of the central government making transfers to state governments on the basis of political considerations. They find that a state which is both aligned and swing in the last state election is estimated to receive 16% higher transfers than a state which is unaligned and non-swing. Bardhan and Mookherjee (2010) investigate political determinants of land reform implementation in the Indian state of West Bengal since the late 1970s. Their findings are consistent with a quasi-Downsian theory stressing the role of opportunism (re-election concerns) and electoral competition.

In this paper, we take a step towards investigating which groups *within* the swing districts get the larger share of the benefits. This paper is closely linked to the literature on "clientelism". Bardhan and Mookherjee (1999) provide a theoretical framework which can deal with the issue of the connection between electoral competitiveness and clientelism. However, their main focus is the relationship between the degree of decentralization and clientelism, whereas here we concentrate on the link between electoral competitiveness and inequality/polarization.

Our model begins with the premise that political parties can commit to policies which differentially affect the incomes of individuals in society. Suppose there are three categories of public goods available in society — one that disproportionately benefits the poor, another which disproportionately benefits the rich and finally a pure public good which benefits all groups equally. One could think of these benefits as augmenting the incomes of the citizens. For simplicity, we assume that there are three income groups in society —

[2]The measure of polarization posited by Foster and Wolfson (1992, 2010) is well-disposed towards capturing the size of the middle class. In fact, this is the measure we use extensively in the paper.

the *poor*, the *middle-income* and the *rich*. [3] Prior to elections, each of the two political parties can commit to a certain level of investment in each of these public goods. Of course, investment is costly — typically, it requires the party candidate to lobby the central government for funds, monitor the progress of the projects and so on. Also, in the spirit of Lindbeck and Weibull (1987) and Dixit and Londregan (1996) there is a constituency-level bias in favor of one party which is drawn from some distribution known to all. This bias can be interpreted as the ideological preferences of the voters in the constituency.[4] Given that each party wishes to maximize plurality *net* of investment costs, in equilibrium both parties end up proposing identical investment platforms provided they face the same investment cost function.

The model also delivers — in line with Lindbeck and Weibull (1987) and Dixit and Londregan (1996)— that as the level of electoral uncertainty increases in the district, the equilibrium level of investment in each type of public good increases; in other words, there is greater transfer to the electorally competitive districts.

The main result from our model is that the effect of close elections on either inequality or polarization crucially depends upon the the degree of concavity of the income–from–investment function. Specifically, if the degree of concavity is "low" then close elections *reduce* inequality; the relation is *reversed* if the degree of concavity is "high".

The key intuition behind the above result comes from the following observation. Increased electoral uncertainty forces each party to expend more effort into investment and so they allocate investment among the three goods in such a manner that gets them the highest possible return in terms of votes. The fact that the poor outnumber the rich guarantees that the investment in the pro–poor goods exceeds that in the pro–rich policies at any point in time. However, when the district gets more uncertain electorally, the issue of how to allocate additional investment among the three types of goods essentially determines which group gains the most. Depending upon the extent of concavity of the income–from–investment function, the biggest gainers could be the poor or the rich. Specifically, if the income–from–investment function exhibits a 'low' degree of concavity then there is disproportionately greater income gain for the poor than for the rich, thus reducing inequality and in some cases polarization. On the other hand, if the degree of concavity of the income–from–investment function is above some threshold, then the rich gain more income than the poor and inequality (and sometimes polarization) actually increases as a result of increased electoral uncertainty.

We then check for the relationship between close elections and income inequality and polarization for the case of national elections in India. Our main variable representing electoral "swing" is the actual *margin of winning*; in other words, we look at the difference between the percentage vote shares of the two parties that obtain the highest number of

[3] We allow for a continuous distribution of incomes in society but only require that the the different income-earners can be sorted into the three broad income groups; so within each group, there is some heterogeniety of incomes.

[4] Typically using any such probalistic-voting setup helps guarantee an equilibrium in pure strategies, which is something clearly desirable in this context.

votes in any constituency. The two NSS consumer expenditure rounds we utilize have almost 16 years between them and these intervening years have been witness to several national elections.[5] In the baseline specification, we take an average of the winning margins over several elections prior to each of the NSS expenditure rounds to get a measure of the electoral volatility of the districts.

We also experiment with alternative variables for electoral swing; for example, we restrict attention to the most recent election that took place before the relevant NSS expenditure round, rather than an average over several prior elections. The results we get are robust to such variations — more "swing" districts exhibit lower (expenditure) inequality. The pattern persists when we replace winning margin by simply the vote share of the winning party. There is evidence of a similar relationship between polarization and electoral uncertainty. Inter-quartile differences in expenditure (normalized by the average level of expenditure) are also positively associated with higher winning margins.

In sum, the empirical findings clearly suggest that greater electoral uncertainty reduces existing income disparities and promotes the growth of the middle class. One way to interpret the empirical results would be by means of our theory. Our interpretation is that in India, the income–from–investment function is not 'excessively' concave; in other words, the expected marginal gains from investing in public goods in India is fairly significant.

The remainder of the paper is organized as follows. Section 2 presents a simple model of electoral competition which describes the impact of uncertainty in election outcomes on equilibrium policy platforms and hence on the resulting income distribution in society. Section 3 describes the data, the empirical strategy and findings and Section 4 concludes.

## 2 The model

Suppose that a nation is composed of $N$ districts where $N$ is "large". In every district there is a unit mass of individuals who differ in terms of their incomes. For simplicity, assume that there are three distinct income groups in society — the poor (denoted by $p$), the middle-income (denoted by $m$) and the rich (denoted by $r$). Let $Y(.)$ represent the cdf of incomes in society and let $y_m$ and $y_r$ be two (exogenously given) income levels with $0 < y_m < y_r$ such that anyone with income lower than $y_m$ falls into group $p$ and anyone with income between $y_m$ and $y_r$ falls into group $m$. All individuals earning at least $y_r$ constitute the group $r$. Also, let $n_i$ denote the mass of group $i$ for $i = p, m, r$. We assume that $n_p > n_r$. Note that the way the income groups (and their corresponding sizes) have been defined makes it clear that the *median* income earner in a district could belong to either group $p$ or group $m$. However, it is reasonable to proceed with the presumption that the median income-earner belongs to group $m$; in a sense, it provides a natural interpretation of the

[5] The two rounds are the 43rd round (conducted in 1987-88) and the 61st round (conducted in 2003-04). Also, national elections take place once every 5 years. Sometimes, they are more frequent. For instance, when the incumbent government fails a "vote of confidence" (a sign that the ruling party has the support of the majority of the national legislators) and is forced to resign, fresh elections are called.

notion of a “middle class”.

There are two political parties $A$ and $B$ who field candidates for election. Each candidate proposes some (non-negative) allocation of investment in public goods. We assume that there are three categories of public goods in society:

(i) *Pro-poor* public goods: Fix some level of investment in this good, say $I_p$. Any additional investment in this good generates some positive benefits for all income groups. However, the marginal benefit to group $p$ outweighs that to group $m$ which in turn exceeds that to group $r$.

(ii) *Pro-rich* public goods: Fix some level of investment in this good, say $I_r$. Any additional investment in this good generates some positive benefits for all income groups. However, the marginal benefit to group $r$ outweighs that to group $m$ which in turn exceeds that to group $p$.

(iii) *Pure* public goods: Fix some level of investment in this good, say $I_m$. Any additional investment in this good generates equal positive benefits for all income groups.

It is not difficult to cite examples of each type of public good particularly in the context of developing countries. Foster and Rosenzweig (2001) posit a model with three kinds of public goods — irrigation facilities (pumps, tanks, tubewells), roads and schools — which differentially affect the welfare of the landed (and hence better-off) and landless (and hence poor) households. Specifically, they show that public expenditure on road-construction programs primarily benefit landless households by increasing local labor demand and the public purchase of irrigation facilities increases agricultural production and thus raises land rents which boost the incomes of the landed households. In our model, we could think of public expenditure on schools as a pure public good.

We will now make precise how investment in each of the three public goods affects any citizen’s payoff. Suppose the level of investment in the three goods are given by $\mathbf{I} \equiv (I_p, I_m, I_r) \geq \mathbf{0}$.

For an individual in group $p$, this generates a payoff (over and above her initial exogenously given income) given by $w_p(\mathbf{I}) = \lambda\beta(I_p) + \beta(I_m) + \mu\beta(I_r)$.

For an individual in group $m$, this generates a payoff (over and above her initial exogenously given income) given by $w_m(\mathbf{I}) = \beta(I_p) + \beta(I_m) + \beta(I_r)$.

For an individual in group $r$, this generates a payoff (over and above her initial exogenously given income) given by $w_r(\mathbf{I}) = \mu\beta(I_p) + \beta(I_m) + \lambda\beta(I_r)$.

By the nature of the public goods defined above, it must be that $\lambda > 1 > \mu \geq 0$. We assume that $\beta(0) = 0$, $\beta'(0) = \infty$, $\beta'(x) > 0$ for every $x > 0$ and $\beta'' < 0$. See Figure 1 for a useful illustration.

Now we turn to the question of how investment is each of these public goods is financed. One could think of there being a proportional tax on income which is uniformly applied to all districts. This gives us the total national funds for such investment which is then divided among the $N$ districts. It would be interesting in its own right to study how the bargaining between the districts determine the allocation of funds to each district; we do not do so here as our focus is on a different matter. We will simply assume that each district gets an equal share of the national funds for investment in these public goods; call it $R$.

Note, investment in any public good is costly in the sense that it requires effort by the party's candidate. One could think of this cost as monitoring costs of the projects, implementation costs, etc. Alternatively, one could think of this as "leaky buckets" in the sense that not all the funds get translated into fruitful investment, a part of it it lost or stolen. We assume that the cost of investment by party $j$ is given by $c(\hat{I}_j)$ where $\hat{I}_j = I_p^j + I_m^j + I_r^j$ for $j = A, B$. So, the cost function is common to both parties.

Also, we assume that it takes a linear form so that $c(\hat{I}) = c\hat{I}$ where $c > 0$. To capture the notion that the parties are actually cynical or "lazy", we have that $c$ is sufficiently high so that each party would never actually choose to spend all the available funds on the public goods. In some sense, we are abstracting from the matter of inter-district competition for funds and focusing on the case where paucity of funds in any district is not really an issue. This is plausible since in most developing countries the complaint is often that funds lie unused or are embezzled rather than being put to the purposes they were intended for. In some sense, we are taking the approach that more funds could be generated through greater taxation but the parties — either because they are plain lazy or corrupt — choose not to, since delivering a higher amount of public goods is too costly for them, after a certain point. In this spirit, we proceed as if the constraint $I_p + I_m + I_r \leq R$ is never binding.

An individual's preferences over candidates (and their proposed policies) are described as follows. First, individual $v$ exhibits a bias $a_v$, positive or negative, for party $A$. The corresponding payoff from $B$ is normalized to be zero, so $a_v$ is really a difference. This ideological bias can stem from many things, say the parties stand on issues other than public goods provision and so on. Moreover, we assume that every individual draws this bias from the *same* distribution with cdf $F(.)$ and corresponding density $f$ positive everywhere on the real line. Thus, it is this $F$ function which captures the ideological leanings or partisan nature of the constituency.

The timing is as follows. Both parties $A$ and $B$ field their respective candidates each of whom proposes a vector of investments, i.e. $\mathbf{I}_j \equiv (I_p^j, I_m^j, I_r^j) \geq \mathbf{0}$ for $j = A, B$. Each voter then draws her bias from $F$ (note, $F$ is public information but each individual's realization is observed by the individual alone) and then votes for the party who promises her higher utility. The party with the highest number of votes is declared the winner and the winner's proposed platform is implemented. Note, there is full-commitment from each party's side in keeping with the Downsian tradition.

Suppose $(\mathbf{I}_A, \mathbf{I}_B)$ is offered by party $A$ and party $B$, respectively. Consider a voter $v$ who belongs to income group $i$ where $i = p, m, r$. She will vote for $A's$ candidate if

$$w_i(\mathbf{I}_A) + a_v > w_i(\mathbf{I}_B).$$

Note, voter $v$ will vote for $B's$ candidate if the opposite inequality holds and will be indifferent in case of equality.

From the perspective of the party, an individual's vote is stochastic. The probability that she will vote for party $A's$ candidate is given by

$$1 - F(w_i(\mathbf{I}_B) - w_i(\mathbf{I}_A)).$$

Call it $p_i$ (note, it is the same for every voter $v$ in group $i$). The expected plurality for party $A$ is proportional to $\sum_i n_i p_i$.[6] Parties care about maximizing their respective expected plurality but are also sensitive to the cost of investment. In particular, given $B's$ platform $\mathbf{I}_B$, party $A$ will choose $\mathbf{I}_A$ to maximize:

$$\sum_i n_i p_i - c(\hat{I}_A).$$

Similarly, party $B$ will take $\mathbf{I}_A$ as given and choose $\mathbf{I}_B$ to maximize:

$$1 - \sum_i n_i p_i - c(\hat{I}_B).$$

Assuming that $F(w_i(\mathbf{I}_B) - w_i(\mathbf{I}_A))$ is convex in $\mathbf{I}_A$ (for any $\mathbf{I}_B$) and concave in $\mathbf{I}_B$ (for any $\mathbf{I}_A$) for each group $i$ — in the same vein as Lindbeck and Weibull (1987) — is sufficient to guarantee the existence of best-response functions for each of the two parties.

This sets up the ground for our first result which is stated in Proposition 1 below.

PROPOSITION **1.** *There is a unique equilibrium of this game. Moreover, the equilibrium is symmetric with each party offering platform* $\mathbf{I}^* \equiv (I_p^*, I_m^*, I_r^*)$ *where* $\mathbf{I}^*$ *satisfies the following equations (1)—(3):*

$$f(0)\beta'(I_p)[n_p\lambda + n_m + n_r\mu] = c \tag{1}$$

$$f(0)\beta'(I_m) = c \tag{2}$$

$$f(0)\beta'(I_r)[n_p\mu + n_m + n_r\lambda] = c \tag{3}$$

*Moreover,* $I_p^* > I_r^*$.

*Proof.* The proof is established in a few steps. First we show that there exists a unique $\mathbf{I}^* \equiv (I_p^*, I_m^*, I_r^*)$ which satisfies equations (1)—(3) and also $I_p^* > I_r^*$. Then we establish that that both parties offering $\mathbf{I}^*$ is an equilibrium of this game. The final step establishes the uniqueness of the equilibrium.

[6] To be precise, the expected plurality is given by $\sum_i n_i[p_i - (1-p_i)] = \sum_i n_i[2p_i - 1]$.

Step 1 (existence of $\mathbf{I}^* \equiv (I_p^*, I_m^*, I_r^*)$):

This is straightforward since $\beta' > 0$, $\beta'' < 0$ and $\beta'(0) = \infty$. Also, from equations (1) and (3), we get:

$$\beta'(I_p^*)[n_p\lambda + n_m + n_r\mu] = \beta'(I_r^*)[n_p\mu + n_m + n_r\lambda]$$

Re-arranging terms, we get:

$$\frac{\beta'(I_p^*)}{\beta'(I_r^*)} = \frac{n_p\mu + n_m + n_r\lambda}{n_p\lambda + n_m + n_r\mu}$$

Note that the RHS of the above equation is strictly less than unity since $\lambda > \mu$ and $n_p > n_r$. By the strict concavity of $\beta$ we have that $I_p^* > I_r^*$.

Step 2 (establishing that ($\mathbf{I}^*, \mathbf{I}^*$) constitutes an equilibrium):

Now we return to the basic problem each political party faces. Given $\mathbf{I}_B$, Party $A$ chooses $\mathbf{I}_A$ to maximize:

$$\sum_i n_i p_i - c(\hat{I}_A)$$

where

$$p_i = 1 - F(w_i(\mathbf{I}_B) - w_i(\mathbf{I}_A)).$$

Let

$$d_i \equiv w_i(\mathbf{I}_B) - w_i(\mathbf{I}_A).$$

The first order conditions are the following:

$$FOC(I_p^A) : \beta'(I_p^A)[f(d_p)\lambda n_p + f(d_m)n_m + f(d_r)\mu n_r] = c \tag{4}$$

$$FOC(I_m^A) : \beta'(I_p^A)[f(d_p)n_p + f(d_m)n_m + f(d_r)n_r] = c \tag{5}$$

$$FOC(I_r^A) : \beta'(I_r^A)[f(d_p)\mu n_p + f(d_m)n_m + f(d_r)\lambda n_r] = c \tag{6}$$

Now, given $\mathbf{I}_A$, Party $B$ chooses $\mathbf{I}_B$ to maximize:

$$1 - \sum_i n_i p_i - c(\hat{I}_B)$$

It is easily checked that $B$'s problem yields first-order conditions which are analogous to Party $A$'s. They are:

$$FOC(I_p^B) : \beta'(I_p^B)[f(d_p)\lambda n_p + f(d_m)n_m + f(d_r)\mu n_r] = c \tag{7}$$

$$FOC(I_m^B) : \beta'(I_p^B)[f(d_p)n_p + f(d_m)n_m + f(d_r)n_r] = c \tag{8}$$

$$FOC(I_r^B) : \beta'(I_r^B)[f(d_p)\mu n_p + f(d_m)n_m + f(d_r)\lambda n_r] = c \tag{9}$$

Suppose Party $B$ offers $\mathbf{I}^*$. Then using Party $A$'s FOCs one can check that offering $\mathbf{I}^*$ constitutes a best-response for Party $A$. Similarly, when Party $A$ offers $\mathbf{I}^*$, offering $\mathbf{I}^*$ constitutes a best-response for Party $B$ (using Party $B$'s FOCs). Thus, we have established that both parties offering $\mathbf{I}^*$ is an equilibrium.

Step 3 (uniqueness):

Suppose $(\mathbf{I}_A, \mathbf{I}_B)$ is an equilibrium different from $(\mathbf{I}^*, \mathbf{I}^*)$. Comparing the FOCs for $I_p$ for the two parties yields that $I_p^A = I_p^B$. Similar arguments apply for $I_m$ and $I_r$. This implies that $\mathbf{I}_A = \mathbf{I}_B$. The strict concavity of $\beta$ guarantees that $\mathbf{I}_A = \mathbf{I}_B = \mathbf{I}^*$.

Combining Steps (1)—(3) establishes the proposition. ■

Like in the previous literature (for instance, see Arulampalam et al (2009)) we interpret the density of the bias evaluated at 0, namely $f(0)$, to be an index of how *swing* or non-partisan the constituency happens to be. To see this in a more intuitive sense, consider density functions which are symmetric and unimodal. Now consider two constituencies $s$ and $t$ where $f_s(0) > f_t(0)$. This is roughly equivalent to saying that constituency $s$, in relation to $t$, has a higher proportion of citizens who are ideologically equidistant (or detached) from either party. Thus, $s$ is more swing than $t$ and so the former constituency can be more unpredictable in terms of election results. This suggests that competition should be tighter in $s$ as compared to $t$. This, in turn, leads to $s$ being more favored by the competing parties than $t$. In fact, in line with the findings of the previous literature, this is what is stated in Proposition 2 below.

PROPOSITION **2.** *Increased electoral uncertainty, as captured by a rise in $f(0)$, results in higher aggregate public good investment. Moreover, investment in every type of public good is increased.*

*Proof.* Let there be a rise in $f(0)$. From equations (1)—(3) given in the previous proposition, it is clear that investment in every type of public good rises since $\beta(.)$ is strictly concave. ■

## 2.1 The Effect on Income Inequality and Polarization

Before proceeding in any further, we presume a generic functional form for the income–from–investment function, $\beta(x)$. In particular, let $\beta(x)$, be of the following isoelastic form:

$$\beta(x) = \frac{x^{1-\sigma}}{1-\sigma}$$

for $\sigma > 0, \sigma \neq 1$.

So far we have not imposed any restrictions on $\lambda$ and $\mu$, except that $\lambda > 1 > \mu \geq 0$. We have not specified how high (low) $\lambda$ ($\mu$) has to be in relation to unity. But it is quite plausible that group $p$ ($r$) might find its own type of good much more important than either the pure public good or group $r'$s ($p'$s) preferred type of good. In other words, it seems quite likely that $\lambda - 1 \geq 1 - \mu$. So from now on we impose the following assumption:

*Assumption P:* $\lambda - 1 \geq 1 - \mu$.

Let electoral uncertainty increase as defined above, i.e., by a rise in $f(0)$. Let $\tilde{\mathbf{I}}$ represent the corresponding platform proposed by both parties in equilibrium.

Consider the change in the incomes of the members of group $i$ for $i = p, m, r$.

$$\Delta y_p = \lambda[\beta(\tilde{I}_p) - \beta(I_p)] + [\beta(\tilde{I}_m) - \beta(I_m)] + \mu[\beta(\tilde{I}_r) - \beta(I_r)]$$
$$\Delta y_m = [\beta(\tilde{I}_p) - \beta(I_p)] + [\beta(\tilde{I}_m) - \beta(I_m)] + [\beta(\tilde{I}_r) - \beta(I_r)]$$
$$\Delta y_r = \mu[\beta(\tilde{I}_p) - \beta(I_p)] + [\beta(\tilde{I}_m) - \beta(I_m)] + \lambda[\beta(\tilde{I}_r) - \beta(I_r)]$$

This implies the following relationships:

$$\Delta y_p - \Delta y_m = (\lambda - 1)[\beta(\tilde{I}_p) - \beta(I_p)] + (\mu - 1)[\beta(\tilde{I}_r) - \beta(I_r)]$$
$$\Delta y_m - \Delta y_r = (1 - \mu)[\beta(\tilde{I}_p) - \beta(I_p)] + (1 - \lambda)[\beta(\tilde{I}_r) - \beta(I_r)]$$

*Claim 1. For $\beta(x) = \frac{x^{1-\sigma}}{1-\sigma}$ where $\sigma \in (0,1)$, it must be that $\beta(\tilde{I}_p) - \beta(I_p) > \beta(\tilde{I}_r) - \beta(I_r)$.*

*Proof.* For $\beta(x) = \frac{x^{1-\sigma}}{1-\sigma}$, equations (1) and (3) imply:

$$\frac{\beta'(I_p)}{\beta'(I_r)} = [I_r/I_p]^{\sigma} = \frac{n_p\mu + n_m + n_r\lambda}{n_p\lambda + n_m + n_r\mu} \equiv \rho < 1$$

Note, $\beta(\tilde{I}_i) - \beta(I_i) = \frac{1}{1-\sigma}[\tilde{I}_i^{1-\sigma} - I_i^{1-\sigma}]$ for $i = p, m, r$. Therefore,

$$\beta(\tilde{I}_p) - \beta(I_p) > \beta(\tilde{I}_r) - \beta(I_r) \Leftrightarrow \tilde{I}_p^{1-\sigma} - I_p^{1-\sigma} > \tilde{I}_r^{1-\sigma} - I_r^{1-\sigma}$$

Note,

$$\tilde{I}_p^{1-\sigma} - I_p^{1-\sigma} = \tilde{I}_r^{1-\sigma}[(\tilde{I}_p/\tilde{I}_r)^{1-\sigma} - (I_p/\tilde{I}_r)^{1-\sigma}] = \tilde{I}_r^{1-\sigma}[\rho^{1-1/\sigma} - \rho^{1-1/\sigma}(I_r/\tilde{I}_r)^{1-\sigma}]$$
$$= \tilde{I}_r^{1-\sigma}\rho^{1-1/\sigma}[1 - (I_r/\tilde{I}_r)^{1-\sigma}].$$

On the other hand,

$$\tilde{I}_r^{1-\sigma} - I_r^{1-\sigma} = \tilde{I}_r^{1-\sigma}[1 - (I_r/\tilde{I}_r)^{1-\sigma}].$$

Therefore,

$$\beta(\tilde{I}_p) - \beta(I_p) > \beta(\tilde{I}_r) - \beta(I_r) \Leftrightarrow \rho^{1-1/\sigma} > 1.$$

But $\rho^{1-1/\sigma} > 1$ for any $\sigma \in (0,1)$ since $\rho < 1$. ∎

Now it is immediate that $\Delta y_p > \Delta y_r$ for any $\sigma \in (0,1)$. Note, $\Delta y_p - \Delta y_r = (\lambda - \mu)[(\beta(\tilde{I}_p) - \beta(I_p)) - (\beta(\tilde{I}_r) - \beta(I_r))]$ which is positive since $\lambda > \mu$ and by Claim 1.

*Claim 2. Suppose $\lambda - 1 \geq 1 - \mu$ and $\sigma \in (0,1)$. Then it must be that $\Delta y_p > \Delta y_m$.*

*Proof.* Note that

$$\Delta y_p - \Delta y_m = (\lambda - 1)[\beta(\tilde{I}_p) - \beta(I_p)] + (\mu - 1)[\beta(\tilde{I}_r) - \beta(I_r)].$$

Given Claim 1, we have $\Delta y_p > \Delta y_m$ as long as $\lambda - 1 \geq 1 - \mu$. ∎

This leaves us with two possibilities.

(i) $\Delta y_p \geq \Delta y_m \geq \Delta y_r$ with at least one inequality strict.

(ii) $\Delta y_p \geq \Delta y_r \geq \Delta y_m$ with at least one inequality strict.

Suppose $\tilde{Y}$ represents the new income distribution while the orginal one is given by $Y$. We first examine case (i).

*Claim 3. Suppose $\Delta y_p \geq \Delta y_m \geq \Delta y_r$ with at least one inequality strict. Then it must be that income inequality as measured by any Lorenz-consistent measure must have undergone a reduction on moving from $Y$ to $\tilde{Y}$. Also, polarization must have reduced.*

*Proof.* Start from $\tilde{Y}$. Now remove an income of $\Delta y_m$ from every person. This essentially brings the $m$–group back to their original position under the original income distribution $Y$. But both the other groups have moved towards group $m$ since $\Delta y_p > \Delta y_m > \Delta y_r$.

Now compare $(\Delta y_p - \Delta y_m)n_p$ and $(\Delta y_m - \Delta y_r)n_r$.

Suppose the first term is weakly smaller. Now, consider the folowing regressive transfer. Take $(\Delta y_p - \Delta y_m)$ from the original $p$–group and transfer to the original $r$–group, while dividing equally among the donors and recipients. This brings the $p$–group back to their original position as under $Y$ whereas the $r$–group is either at their original position under $Y$ or are poorer than under $Y$. Call this new distribution $Y'$. Note, $Y'$ is clearly more unequal than $\tilde{Y}$ by the Dalton principle. Moreover, $Y'$ has the same or lower inequality than $Y$. Therefore, income inequality as measured by any Lorenz-consistent measure must have undergone a *reduction* on moving from $Y$ to $\tilde{Y}$.

Suppose now that $(\Delta y_p - \Delta y_m)n_p > (\Delta y_m - \Delta y_r)n_r$. Now, consider the folowing regressive transfer. Take $(\Delta y_m - \Delta y_r)n_r$ from the original $p$–group and transfer to the original $r$–group, while dividing equally among the donors and recipients. This brings the $r$–group back to their original position as under $Y$ whereas the $p$–group is richer than under $Y$. Call this new distribution $Y''$. Note, $Y''$ is clearly more unequal than $\tilde{Y}$ by the Dalton principle. Moreover, $Y''$ has lower inequality than $Y$. Therefore, income inequality as measured by any Lorenz-consistent measure must have undergone a *reduction* on moving from $Y$ to $\tilde{Y}$.

To see the effect on income polarization, suppose $\hat{Y}$ represents the new income distribution with all income normalized so that the median income under the two distributions remain the same (median–normalization). Since $\Delta y_p > \Delta y_m > \Delta y_r$ and the median income–earner lies in group $m$, the mass of population earning between $y_m$ and $y_r$ is larger under $\hat{Y}$ than under $Y$. This clearly implies a growth of the middle class and reduced (income) polarization in terms of the Foster–Wolfson polarization measure.[7] ∎

We next examine case (ii).

*Claim 4. Suppose $\Delta y_p \geq \Delta y_r \geq \Delta y_m$ with at least one inequality strict. Also, suppose $n_p \geq n_m$. Then there exists a $\underline{\sigma} \in (0,1)$ such that for any $\sigma \in (0, \underline{\sigma}]$, increased electoral uncertainty leads to a reduction in income inequality as captured by any Lorenz–consistent measure. The effect on income polarization is however ambiguous.*

*Proof.* Start from $\tilde{Y}$. Now remove an income of $\Delta y_r$ from every person. This essentially brings the $r$–group back to their orginal position under the original income distribution $Y$. But the $p$–group and the $m$–group have moved closer to each other. First compare $(\Delta y_p - \Delta y_r)n_p$ and $(\Delta y_r - \Delta y_m)n_m$.

Suppose $(\Delta y_p - \Delta y_r)n_p \geq (\Delta y_r - \Delta y_m)n_m$. Now consider the following regressive transfer. Take $(\Delta y_r - \Delta y_m)n_m$ from the original $p$–group and transfer to the original $m$–group, while dividing equally among the donors and recipients. This brings the $m$–group back to their original position as under $Y$ whereas the $p$–group is the same or richer than under $Y$. Call this new distribution $Y''$. Note, $Y''$ is clearly more unequal than $\tilde{Y}$ by the Dalton principle. Moreover, $Y''$ has either the same or lower inequality than $Y$. Therefore, income inequality as measured by any Lorenz-consistent measure must have undergone a *reduction* on moving from $Y$ to $\tilde{Y}$.

Now suppose $(\Delta y_p - \Delta y_r)n_p < (\Delta y_r - \Delta y_m)n_m$. Given that $n_p \geq n_m$, this means $(\Delta y_p - \Delta y_r) < (\Delta y_r - \Delta y_m)$. Recall that $\frac{\beta(\tilde{I}_p)-\beta(I_p)}{\beta(\tilde{I}_r)-\beta(I_r)} = \rho^{1-1/\sigma}$. It is easily checked that $(\Delta y_p - \Delta y_r) < (\Delta y_r - \Delta y_m)$ is equivalent to $\rho^{1-1/\sigma} < \frac{2\lambda-\mu-1}{\lambda+1-2\mu}$. Note, $\rho^{1-1/\sigma}$ is falling in $\sigma$. Moreover, $\rho^{1-1/\sigma} > 1$ for any $\sigma \in (0,1)$. In fact, $\rho^{1-1/\sigma} \to \infty$ as $\sigma \to 0$. By the continuity of $\rho$, this implies there is some $\sigma < 1$, call it $\underline{\sigma}$ such that

$$\rho^{1-1/\underline{\sigma}} = \frac{2\lambda - \mu - 1}{\lambda + 1 - 2\mu}.$$

Thus for any $\sigma \in (0, \underline{\sigma}]$, it must be that $(\Delta y_p - \Delta y_r)n_p \geq (\Delta y_r - \Delta y_m)n_m$. Hence, there must be a reduction in income inequality as captured by any Lorenz–consistent measure.

However, the effect on income polarization is ambiguous since the the $p$–group and the $m$–group have moved closer to each other, in relation to group $r$. ∎

Combining the results in Claims 1—4, we get the following proposition.

[7] For a graphical demonstration, see Figure 2.

PROPOSITION **3.** *Suppose the income–from–investment function, $\beta(x)$, is of the following generic isoelastic functional form:*

$$\beta(x) = \frac{x^{1-\sigma}}{1-\sigma}$$

*for $\sigma > 0, \sigma \neq 1$.*

*Also, suppose $n_p \geq n_m$ and $n_p > n_r$. Then, under Assumption P, there exists a $\underline{\sigma} \in (0,1)$ such that for any $\sigma \in (0, \underline{\sigma}]$, increased electoral uncertainty leads to a reduction in income inequality as captured by any Lorenz–consistent measure. The effect on income polarization is however ambiguous.*

Next, we turn to the case where $\sigma \to 1$. Recall, $lim_{\sigma \to 1} \frac{x^{1-\sigma}}{1-\sigma} = ln(x)$. Note, in this case, we have

$$\frac{\beta'(I_p)}{\beta'(I_r)} = I_r/I_p = \frac{n_p\mu + n_m + n_r\lambda}{n_p\lambda + n_m + n_r\mu} = \rho < 1.$$

This leads to $\beta(\tilde{I}_p) - \beta(I_p) = \beta(\tilde{I}_r) - \beta(I_r)$. Noting that $\lambda - 1 \geq 1 - \mu$, we conclude that

$$\Delta y_p = \Delta y_r \geq \Delta y_m$$

with strict inequality whenever $\lambda - 1 > 1 - \mu$. This, in turn, implies that the effect on income inequality and polarization is ambiguous. The following proposition summarizes this result.

PROPOSITION **4.** *Suppose the income–from–investment function, $\beta(x)$, is given by*

$$\beta(x) = ln(x).$$

*Then, under Assumption P, the effect of increased electoral uncertainty on income inequality, as captured by any Lorenz–consistent measure, is ambiguous. The same is true of income polarization.*

This leads us to the question as to how increasing electoral uncertainty affects inequality and polarization when $\sigma$ exceeds unity. It turns out that the effect of increased electoral competition on income inequality can be quite different for high values of $\sigma$.

First, note that $\sigma > 1$ implies that

$$\frac{\beta(\tilde{I}_p) - \beta(I_p)}{\beta(\tilde{I}_r) - \beta(I_r)} = \rho^{1-1/\sigma} < 1.$$

*Claim 5. The group which clearly benefits the most from increased electoral competition is the rich. In other words, $\Delta y_r > max\{\Delta y_m, \Delta y_p\}$.*

*Proof.* $\Delta y_r - \Delta y_p = (\lambda - \mu)[(\beta(\tilde{I}_r) - \beta(I_r)) - (\beta(\tilde{I}_p) - \beta(I_p))] > 0$ since $\frac{\beta(\tilde{I}_p) - \beta(I_p)}{\beta(\tilde{I}_r) - \beta(I_r)} < 1$. Also,

$$\Delta y_r - \Delta y_m = (\lambda - 1)(\beta(\tilde{I}_r) - \beta(I_r)) - (1 - \mu)(\beta(\tilde{I}_p) - \beta(I_p))$$

$$\geq (1-\mu)[(\beta(\tilde{I}_r) - \beta(I_r)) - (\beta(\tilde{I}_p) - \beta(I_p))] > 0.$$

This establishes the claim. ∎

So there are two possibilities.

(i) $\Delta y_r \geq \Delta y_m \geq \Delta y_p$ with at least one inequality strict.

(ii) $\Delta y_r \geq \Delta y_p \geq \Delta y_m$ with at least one inequality strict.

As before, let $\tilde{Y}$ represent the new income distribution while the orginal one is given by $Y$. We first examine case (i).

*Claim 6. Suppose $\Delta y_r \geq \Delta y_m \geq \Delta y_p$ with at least one inequality strict. Then it must be that income inequality as measured by any Lorenz-consistent measure must have undergone an increase on moving from $Y$ to $\tilde{Y}$. Also, polarization must have increased.*

*Proof.* The proof follows directly from arguments analogous to those made in the proof of Claim 3. ∎

Before turning to case (ii), we would like to point out some properties of $\rho \equiv \frac{n_p\mu + n_m + n_r\lambda}{n_p\lambda + n_m + n_r\mu}$. Recall that $n_p = 1 - n_m - n_r$. Hence, we can rewrite $\rho$ as the following ratio:

$$\frac{\mu + n_m(1-\mu) + n_r(\lambda - \mu)}{\lambda + n_m(1-\lambda) + n_r(\mu - \lambda)}.$$

Note, that holding $n_m$ fixed, any increase in $n_r$ leads to an increase in $\rho$ since $\lambda > \mu$. The same holds for any increase in $n_m$ while holding $n_r$ fixed since $\lambda > 1 > \mu \geq 0$. Thus, a reduction (increase) in $n_p$ relative to $n_m$ and $n_r$ leads to an increase (decrease) in $\rho$.

We now turn to case (ii).

*Claim 7. Suppose $\Delta y_r \geq \Delta y_p \geq \Delta y_m$ with at least one inequality strict. Let the district have a 'high' share of poor people in the sense that $\rho < \frac{\lambda+1-2\mu}{2\lambda-\mu-1}$ and that $\frac{n_m}{n_r} \leq 1$. Then there exists a $\overline{\sigma} > 1$ such that for any $\sigma > \overline{\sigma}$, increased electoral uncertainty leads to an increase in income inequality as captured by any Lorenz–consistent measure. The effect on income polarization is however ambiguous.*

*Proof.* Start from $\tilde{Y}$. Now remove an income of $\Delta y_p$ from every person. This essentially brings the $p$–group back to their orginal position under the original income distribution $Y$. But the $m$–group and the $r$–group have moved away from each other. If we can show that this resulting distribution — call it $Y'$ — can be reached from $Y$ by means of a set of regressive transfers, then we are done. First compare $(\Delta y_r - \Delta y_p)n_r$ and $(\Delta y_p - \Delta y_m)n_m$.

Suppose $(\Delta y_p - \Delta y_m)n_m \leq (\Delta y_r - \Delta y_p)n_r$. Now start with the original distribution $Y$ and consider the following regressive transfer. Take $(\Delta y_p - \Delta y_m)n_m$ from the $m$–group and give it to the $r$–group, while dividing equally among the donors and recipients. Note, this brings the $m$–group to the same position as under $Y'$. Moreover, group $r$ is now either at the same position as under $Y'$ or they are poorer. This implies that $Y'$ — and hence $\tilde{Y}$ — is more unequal as compared to $Y$.

Next, suppose that $(\Delta y_p - \Delta y_m)n_m > (\Delta y_r - \Delta y_p)n_r$. Given that we have $\frac{n_m}{n_r} \leq 1$, this implies $\Delta y_p - \Delta y_m > \Delta y_r - \Delta y_p$. This, in turn, is equivalent to

$$\rho^{1-1/\sigma} > \frac{\lambda + 1 - 2\mu}{2\lambda - \mu - 1}$$

for any $\sigma > 1$. Note, we have assumed that $\rho < \frac{\lambda+1-2\mu}{2\lambda-\mu-1}$. Observe that $\rho^{1-1/\sigma}$ is falling in $\sigma$. In particular, $\rho^{1-1/\sigma} \to \rho$ as $\sigma \to \infty$. Hence by the continuity of $\rho$, there exists a $\overline{\sigma} > 1$ such that for any $\sigma > \overline{\sigma}$ it must be that $\rho^{1-1/\sigma} < \frac{\lambda+1-2\mu}{2\lambda-\mu-1}$. This contradicts $\rho^{1-1/\sigma} > \frac{\lambda+1-2\mu}{2\lambda-\mu-1}$ for every $\sigma > 1$ and rules out $(\Delta y_p - \Delta y_m)n_m > (\Delta y_r - \Delta y_p)n_r$. ■

Combining the results in Claims 5—7, we get the following proposition.

PROPOSITION **5.** *Suppose the income–from–investment function, $\beta(x)$, is of the following generic isoelastic functional form:*

$$\beta(x) = \frac{x^{1-\sigma}}{1-\sigma}$$

*for $\sigma > 0, \sigma \neq 1$.*

*Let the district have a 'high' share of the poor in the sense that $\rho < \frac{\lambda+1-2\mu}{2\lambda-\mu-1}$ and that $n_p > n_r \geq n_m$.*

*Then, under Assumption P, there exists a $\overline{\sigma} > 1$ such that for any $\sigma > \overline{\sigma}$, increased electoral uncertainty leads to an increase in income inequality as captured by any Lorenz–consistent measure. The effect on income polarization is however ambiguous.*

In principle, our theory recognizes that the relationship between electoral competitiveness and inequality (or polarization) depends upon various factors. Our framework clearly identifies the degree of concavity of the income–from–investment function — $\beta(.)$ — as one such crucial factor. To summarize, if the degree of concavity of the income–from–investment function is relatively low (close to linear) then one might expect increased electoral competition to reduce inequality and polarization. On the other hand, an income–from–investment function with a high degree of concavity may result in increased electoral competition actually increasing inequality and polarization. Now, to empirically identify the (average) income–from–investment function in any particular society is a real challenge. All the same, it would be important to know more about this function from a public policy perspective.

Interestingly, our model — though quite simple in its structure — offers a way to interpret empirical relations between close elections (a proxy for electoral competitiveness) and income inequality. If the observed association is negative, then we might use our model to intrepret that the returns-from-investment function does *not* exhibit a high degree of concavity in the sense that $\sigma$ is not too high (strictly less than $\overline{\sigma}$). On the other hand, a positive association between close elections and income inequality suggests the opposite — i.e., the degree of concavity of the income–from–investment function is above a certain threshold in the sense that $\sigma$ is not too small (strictly greater than $\underline{\sigma}$). In a sense, our theory — aside from identifying the relationship between electoral competitiveness and inequality (or polarization) — offers an indirect way of getting at the shape of the income–from–investment function.

One needs to bear in mind that we have only identifed two (separate) *sufficient* conditions for cleanly identifying the relationship between electoral competition and inequality: one which guarantees a *negative* relationship (Proposition 3) and one which guarantees a *positive* one (Proposition 5). It would be more interesting — although no doubt, more challenging — to identify conditions which are both necessary and sufficient.

We now turn to our findings with regard to data from India.

# 3 Empirical Analysis

## 3.1 Data

We need to combine data on incomes with the data on election outcomes. In the case of India, nationally representative data on personal incomes is hard to obtain since a vast majority of Indian households (primarily residing in rural parts) are exempt from payment of income taxes (see Banerjee and Pikkety (2003)). However, there is data on consumer expenditure in India which is publicly available; thus consumer expenditure serves as an excellent proxy for income in our analysis. These data are collected by the National Sample Survey Organization (NSSO) . The National Sample Survey (NSS) is a large-scale consumer expenditure survey which is conducted quinquennially and covers the entire nation; the unit of observation is a household. The recall period used is 30 days, i.e., the surveyed households are asked to provide information on consumption expenditure incurred over the past 30 days. For the current study we use the 43rd and 61st rounds of the NSS. The 43rd round was conducted during July 1987-June 1988 and the 61st round was conducted during July 2004-June 2005. Alongside information on consumer expenditure, the survey also collects data on other socio-economic characteristics of the (surveyed) households such as religion, caste, education, etc.

This information on household expenditure is combined with election data obtained from the Election Commission of India. We use the data for the parliamentary (or federal level) elections from 1977 to 2004. During this period, 11 such general elections took place in India. Our theory requires us to use some measure of the electoral competitiveness of the district — the "swing" nature, so to speak. We primarily utilize the difference in percentage vote shares of the two parties that obtain the highest number of votes in any constituency. This is in line with Arulampalam et al (2009). We use the winning margin and the vote share of the winning party — each averaged over the 3–4 elections prior to each expenditure round — in turn to capture the extent of electoral competition in the district.

We also use the information about whether the constituency had a shift away from or to a Member of Parliament (MP) from the Indian National Congress party.[8] The use of a more refined measure of swing which takes into account movements to and from different parties is not possible for the following reason. There has been an immense proliferation of political parties at both the state and central levels, most of it arising from the splitting up of the main existing national or even regional parties. Moreover, various coalitions — *ad hoc* and otherwise — became popular from the 1980s onward. This makes it very difficult to say whether there really has been an effective shift of regime when say person X wins the same seat first as a candidate of party L and then as a candidate of party R. Given the way the nature of politics and political parties evolved during this period, we chose to proceed with a rather conservative division of parties into "Congress" and "Non-Congress" camps and recorded the movements of a district between these camps over the different election periods.

A brief word about the Indian political system is in order. The Indian Parliament is bicameral in nature. However, the Lok Sabha is the popularly elected House and is *de facto* more powerful than the other House (Rajya Sabha). The popularly elected Members of Parliament (MP) enjoy a five-year term after which fresh Lok Sabha elections are held. There were 518 (Lok Sabha) constituencies in 1971. This went up to 542 after a Delimitation order in 1976 and then to 543 in 1991.

Population is the basis of allocation of seats of the Lok Sabha. As far as possible, every state gets representation in the Lok Sabha in proportion to its population as per census figures. Hence, larger and more populous states have more seats in the Lok Sabha as compared to their smaller and sparsely-populated counterparts. For example, Uttar Pradesh (a north Indian state) with a population of over 166 million has 80 Lok Sabha seats while the state of Nagaland with a population of less than 2 million has only one Lok Sabha seat.

The NSSO expenditure rounds allow identification of the sureveyed household upto the district to which it belongs; no finer identification is possible. However, it is often the case that a single district houses more than one electoral constituency; this is especially true for more populous districts. Given the nature of our hypotheses, we have to restrict attention to only single-constituency districts, i.e. to those places where a district corresponds to just one single constituency [9]. In our sample, there are 179 such districts which we follow for two time periods.

Tables 1 and 2 provide the summary statistics of the variables used in the analysis. Com-

[8] The Indian National Congress party is one of the most prominent national parties in India (established in 1885) and is closely associated with the Indian freedom struggle.

[9] In a district with several constituencies, the link between electoral competitivess and polarization (or inequality) cannot be clearly established. For example, any change in polarization (or inequality) in any one of the constituencies (presumably as a response to electoral competition in that constituency) does not necessarily reflect a similar change in polarization (or inequality) in the district overall.

paring the election data across the two periods, we see that elections clearly became more competitive over the years. For instance, in the elections prior to 1988, the average margin of victory varied between 4% and 51%. On the other hand, in the elections between 1988 and 2004 the average margin was never higher than 31% for any constituency. Between the two periods, both poverty and inequality have fallen on average across the districts suggestive of a trend towards a secular balanced growth. Notably, polarization as measured by the Foster-Wolfson index registers a decline – on average – when comparing across the two periods; this is suggestive of the growth of the "middle-class" over time. Altogether, these tables clearly indicate that there was a lot of dynamism both on the income distribution frontier and in the political scene in India during the period of our study.

We now move on to the details of our empirical strategy for the identification of the relevant parameters.

## 3.2 Empirical Specification

Our data provides a two-period panel spanning 1987-88 and 2004-05. We use a linear fixed effects specification for the empirical exercise. Specifically, for every district $d$ in time period $t$, we have :

$$y_{dt} = \alpha_d + \gamma_t + \beta \mathbf{X}_{dt} + \rho \mathbf{Z}_{dt} + \epsilon_{dt}$$

where $y_{dt}$ is a measure of inequality or polarization, $\mathbf{X}_{dt}$ includes a vector of variables describing the political climate in the district, (like average margin in the last 3-4 elections, etc.). $\mathbf{Z}_{dt}$ is the set of demographic and geographic controls such as the population share of the district, percentage of Hindus in the district, literacy rates and average monthly per capita expenditure for the district. $\alpha_d$ represents the district fixed effects while $\gamma_t$ captures the time effect. Also, $\epsilon_{dt}$ is the error term in this panel specification.

The primary results are collected below.

## 3.3 Results

We first turn to the relationship between electoral uncertainty and inequality in income (in our case, proxied by consumer expenditure). As discussed above, we construct several measures to capture the extent of political competition in a district. The primary proxy for electoral uncertainty exploits the difference in percentage vote shares of the two parties that obtain the highest number of votes in any constituency. This is in the spirit of Arulampalam et al (2009).

We use (i) the winning margin and (ii) the vote share of the winning party — each averaged over the 3—4 elections prior to each expenditure round — in turn to capture the extent of electoral competition in the district. The average margin in the previous 3 to 4 general elections is used as the primary variable to describe how closely the elections have been in a district. We use the average across the previous few elections to ensure that we are not capturing any effect specific to the particular election. The vote share of the winner is also used as a measure of electoral competition. Clearly, the higher the percentage of votes obtained by the winner, the lower the degree of electoral competitiveness in the district.

### 3.3.1 Main results

Table 3 gives the results for our benchmark case, the effect of average margin on the gini coefficent. We find that an increase in the political competition (a lower average margin) implies lower inequality. This result is robust to adding controls including the poverty headcount and poverty gap measures. Table 4 reflects that the result is robust to using the average vote share of winner in place of the average margin.

Additionally, we use the normalized inter-quartile range as a proxy for the level of inequality and also for the size of the middle-class. Even then we see that a higher average margin results in greater difference between the two income quartiles thus normalized (see Table 5).

Next, we turn to the relationship between electoral uncertainty and income polarization. Table 6 shows that polarization is also higher when there is lesser political competition as measured by the average margin. Recall, this is essentially saying that greater political competetion in a district is *positively* associated with a larger middle–class in the district.

### 3.3.2 Robustness checks

Rather than using the average values for the proxies of electoral competition in the previous 3—4 elections, one could also use the margin and vote share of winner from the prior election. We do so and our results are similar to the earlier ones (see Table 7 and Table 8).

Another way to capture the idea of a swing district would be the following. One could possibly identify whenever there is a change in the political party which wins the election in the district. However in 1977 (the first election year we look at) there were only 20 recognized political parties which contested the elections. By 1999 the number of recognized political parties had risen to 47. This significant rise in the number of political parties was not merely a case of greater participation of the general populace in the political domain — it was more the case that several political parties were created by the splintering of existing politcal parties. Therefore, for the time horizon we consider, we are unable to track whether there was a swing away from a particular political party or that merely a segment of the old party came back into power.

The only political party which has remained relatively "stable", in the sense of somewhat maintaining its core identity, is the Indian National Congress. Given the way the nature of politics and political parties evolved during this period, we chose to proceed with a rather conservative division of parties into "Congress" and "Non-Congress" camps and recorded the movements of a district between these camps over the different election periods. Therefore, as our additional measure of political regime change, we use whether or not the district moved away from/towards a Congress MP. We create a dummy variable which takes the value of 1 if there was a change to or from a Congress MP in the district and 0 otherwise. Note, the *swing congress* variable is a very crude measure of the district's electoral volatility and it exhibits much less variation vis-a-vis our other measures of electoral competition.

Table 9 contains some of the results using this *Swing Congress* variable. Note, that the *swing congress* variable exhibits a negative effect on the degree of inequality in the district as captured by the gini coefficent; this is substantively similar to our previous results which used other measures of electoral uncertainty. Table 9 shows that in all specifications the marginal effect of *Swing Congress* on inequality is significant and negative.

The first column in Table 10 reveals a strong negative correlation between *Swing Congress* and polarization in accordance with our previous findings. This effect is robust to the inclusion of several controls; see columns (2) through (5). Therefore, these results reiterate our basic findings.

#### 3.3.3 Concerns

We discuss two of the main concerns that arise in an empirical exercise of the kind undertaken here. The first one is endogeneity due to reverse causality. One could argue that the middle-class votes in a certain way so as to make the political contest close. The second is the issue of migration as a result of political transfers/ public goods provision. We briefly discuss each issue below.

The first concern regarding the voting behavior of the middle class implies the following — it presumes that the members of middle-class vote in a markedly coordinated fashion, which is perhaps not plausible in the Indian context. Bardhan et al (2008) study political participation and targeting of public sevices in the Indian state of West Bengal. In their words "...the difference in reported registration rates and turnouts were modest, more similar to the European patterns rather than the steep asymmetries in the United States. With regard to voting disturbances, there was no clear correlation with socio-economic status." They also find that attendance rates (in political meetings, such as rallies, election meetings called by political parties) did not exhibit any marked unevenness across different land classes. So this does not seem to pose a serious problem. Also, in all of the regressions presented so far, we look at the effect of elections on *subsequent* polarization (and inequality) — so that there is enough of a time lag with elections preceding the corresponding expenditure rounds.

To further bolster our case against reverse causality, we estimate the relationship between electoral uncertainity and inequality (or polarization) using a two–step linear GMM technique which does not assume strict exogeneity; instead only sequential exogeniety is assumed for estimating the relevant parameters. This method, therefore allows for the possibility of past inequality (or polarization) affecting current electoral outcomes. Some of these results are presented in Table 11 and Table 12. We gain re-assurance from the fact that the results with this GMM technique are robust across the different measures of electoral uncertainty and are very much in line with our previous findings.

As to the second concern — namely, migration as a response to political transfers/public goods provision — we can take some comfort in the fact that migration rates in India are rather low in comparison with other developing nations. In fact, Munshi and Rosenzweig (2009) explicitly state that "Among developing countries, India stands out for its remarkably low levels of occupational and spatial mobility." They delve into the proximate causes behind this phenomenon and using a unique panel dataset (identifying sub-caste (*jati*) membership) find that the existence of sub-caste networks that provide mutual insurance to their members play a key role in restricting mobility.

Taking stock of our entire empirical findings, we find that there is a serious relationship between the degree of electoral competition in a district and the nature of redistribution pursued therein. More specifically, we find that districts which have experienced tighter elections tend to evince lower levels of inequality and polarization suggesting that the middle class thrives where political parties are perceived to be relatively balanced in the eyes of the voters.

# 4 Conclusion

In this paper, we study how the degree of electoral competition affects the level of income inequality and the growth of a middle-income group in society. We build a theory based on the traditional two-party Downsian framework with ideological voters in the spirit of Lindbeck and Weibull (1987). Here, the political parties can *a priori* commit to some levels of investment in three different kinds of public goods — one that disproportionately benefits the poor, another which disproportionately benefits the rich and finally a pure public good which benefits all groups equally. One way to interpret these benefits would be to presume that they boost the incomes of the citizens, via some income–from–investment function. The main result from our model is that the effect of close elections on either income inequality or polarization crucially depends upon the degree of concavity of the income–from–investment function. Specifically, if this function exhibits a low degree of concavity then close elections *reduce* inequality; the relation is *reversed* if the function has a high degree of concavity.

The previous literature has stressed the role of political competition in directing tranfers and have generally concluded that "swing" districts get more targeted resources in the aggregate. To the best of our knowledge, no other work has looked at the effect of increased political competetion on the distribution of incomes in society.

We use data from the Indian parliamentary elections which are combined with household-level consumption expenditure data rounds from NSSO (1987-88 and 2003-04) to yield a panel of Indian districts. India has had a vibrant democracy since the nation's independence in 1947. Although there have been several political parties since the 1950s, the national elections had been by and large dominated by the Indian National Congress (INC) party. However, since the 1980s there have been a tremendous proliferation of political parties both at the state and the national levels. In fact, 1977 was witness to a non-Congress led government at the centre for the first time since India's independence. Although the INC continues to be a major player in national elections till this day, it no longer enjoys the kind of monopoly it did prior to the mid-1960s. Moreover, a majority of elections in the 1990s resulted in "hung Parliaments" meaning that no single party obtained a clear majority of seats and thus began the era of coalitional politics in India. Our period of study corresponds to the time *after* the INC had lost its quasi-monopoly in the political arena. So our data is from the phase where national elections were more intensely fought. All of these factors contribute in making India an interesting candidate for testing our hypotheses.

Our main variable representing electoral "swing" is the actual margin of winning which is the difference between the percentage vote shares of the two parties that obtain the highest number of votes in any constituency. Using this variable as our baseline measure of electoral volatility of a district, we obtain that a district which has experienced close elections tends to exhibit lower income inequality. The same is true in case of income polarization. We repeat our analysis with alternative variables for electoral swing; for example, we restrict attention to the most recent election that took place before the relevant NSS expenditure round (rather than take an average over several prior elections). The results we get are robust to such variations — more "swing" districts exhibit lower (expenditure) inequality. The pattern persists when we replace winning margin by simply the vote share of the winning party. There is evidence of a similar relationship between polarization and electoral uncertainty. Inter-quartile differences in expenditure (normalized by the average level of expenditure) also tend to be higher where winning margins are wider.

Overall our empirical findings — in the context of India — clearly suggest that greater electoral uncertainty reduces existing income disparities and promotes the growth of the middle class. How is one to interpret these findings? One way to do so would be by means of our theory. After all, our theoretical framework is supposed to capture some key elements of electoral competition which are particularly relevant for developing countries. Hence, our interpretation is that in India, the income–from–investment function is not 'excessively' concave; in other words, the expected marginal gains from investing in public goods in India is fairly significant.

It is important to point out that the notion of a middle-class adopted here is fairly "local" in the following sense: the middle-class in a district is some group whose earnings correspond to any given income band around the median income-earner in *that* district. Alternatively, one could think of a middle-class at the level of the *nation* (rather than a district) and then study the proportion of people in each district which falls in this "national middle-class". One could investigate how district-level political competition affects the size of this "national middle class" in every district. We plan to explore this question in future work.

In a way our results seem to highlight some interesting features of the electoral mechanism. The key issue here is the presence of people who are highly ideologically inclined towards some political party or the other. A party which rides to victory on the back of large popular support feels less inclined to cater to the toiling masses. After all, if the electorate likes the party to begin with, why should the latter bother working hard to reduce existing disparities? However, if one extends this to a dynamic setting, the voters would potentially change their opinion over time about the inactive (and ineffective) incumbent party. The problem often is that the opponent party — the challenger, so to speak — may not be much of a viable alternative. However, the very realization that perhaps each political party is *ex-ante* as good as the other should drive this voter bias close to nil in expected terms thus inducing better promises (and action) from both parties in future. This would be an interesting avenue to explore.

The fact is that parties themselves change their stand and nature over time. This makes any kind of convergence on part of voter biases quite unlikely. Incidentally, voter biases in regions tend to persist over time. For example, in the context of the US, New York has traditionally been a Democrat stronghold. In India as well, this kind of party loyalty is fairly common — for e.g., West Bengal (a state in eastern India) had been under the rule of a Left-led coalitional government for over 30 years. There may be clientelistic relationships which develop between incumbents and certain sections of the voters (see Bardhan and Mookherjee (1999)) which create such long spells of governance by a party; perhaps longer than what a dynamic extension of our simple model (with updating of voter biases) would predict.

Finally, it would be interesting to explore how different political parties have re-invented themselves over time and what impact has this had on their loyalists — perhaps the conservatives of today would have been liberal half a century ago. A more holistic view of the interplay between party evolution and changing voter loyalities could provide meaningful insights to policy-making.

# References

ARULAMPALAM, W., S. DASGUPTA, A. DHILLON, AND B. DUTTA (2009): "Electoral goals and center-state transfers: A theoretical model and empirical evidence from India," *Journal of Development Economics*, 88(1), 103 – 119.

BANERJEE, A., AND R. PANDE (2007): "Parochial Politics: Ethnic Preferences and Politician Corruption," Harvard University, Kennedy School of Government, typescript.

BANERJEE, A., AND T. PIKETTY (2003): "Top Indian Incomes, 1956-2000," .

BANERJEE, A., AND R. SOMANATHAN (2007): "The political economy of public goods: Some evidence from India," *Journal of Development Economics*, 82(2), 287 – 314.

BARDHAN, P., S. MITRA, D. MOOKHERJEE, AND A. SARKAR (2008): "Local Democracy in West Bengal: Political Participation and Targeting of Public Services.," *Boston University Working Paper, February 2008*.

BARDHAN, P., AND D. MOOKHERJEE (1999): "Relative Capture at Local and National Levels: An Essay in the Political Economy of Decentralization," *Working paper, Institute for Economic Development, Boston University, November 1999.*

——— (2010): "Determinants of Redistributive Politics: An Empirical Analysis of Land Reforms in West Bengal, India," *American Economic Review*, 100(4), 1572–1600.

BHAGWATI, J. N., AND T. N. SRINIVASAN (2002): "Trade and Poverty in the Poor Countries.," *American Economic Review*, 92(2), 180 – 183.

DIXIT, A., AND J. LONDREGAN (1996): "The Determinants of Success of Special Interests in Redistributive Politics," *The Journal of Politics*, 58(4), pp. 1132–1155.

——— (1998): "Fiscal federalism and redistributive politics," *Journal of Public Economics*, 68(2), 153 – 180.

ESTEBAN, J.-M., AND D. RAY (1994): "On the Measurement of Polarization," *Econometrica*, 62(4), pp. 819–851.

——— (2010): "Comparing Polarization Measures," *forthcoming, M. Garfinkel and S. Skaperdas (eds), Oxford Handbook of the Economics of Peace and Conflict, Oxford University Press 2011.*

FOSTER, A., AND M. ROSENZWEIG (2001): "Democratization, Decentralization and the Distribution of Local Public Goods in a Poor Rural Economy.," *Working Paper, Department of Economics, Brown University*.

FOSTER, J. E., AND A. F. SHORROCKS (1991): "Subgroup Consistent Poverty Indices," *Econometrica*, 59(3), 687–709.

FOSTER, J. E., AND M. C. WOLFSON (1992): "Polarization and the decline of the middle class: Canada and the U.S.," .

——— (2010): "Polarization and the decline of the middle class: Canada and the U.S.," *Journal of Economic Inequality*, 8, 247–273, 10.1007/s10888-009-9122-7.

LINDBECK, A., AND J. W. WEIBULL (1987): "Balanced-Budget Redistribution as the Outcome of Political Competition," *Public Choice*, 52(3), 273–297.

LIZZERI, A., AND N. PERSICO (2001): "The Provision of Public Goods under Alternative Electoral Incentives," *The American Economic Review*, 91(1), 225–239.

MUNSHI, K., AND M. ROSENZWEIG (2009): "Why is Mobility in India so Low? Social Insurance, Inequality, and Growth," *Brown University Working Paper, March 2009*.

MYERSON, R. B. (1993): "Incentives to Cultivate Favored Minorities Under Alternative Electoral Systems," *The American Political Science Review*, 87(4), 856–869.

SEN, A. K. (1997): "On Economic Inequality," (enlarged edition with a substantial new annex co-written with James. E. Foster), Clarendon Press, 1997.

TOPALOVA, P. (2005): "Trade Liberalization, Poverty And Inequality: Evidence From Indian Districts," Working Papers id:222, esocialsciences.com.

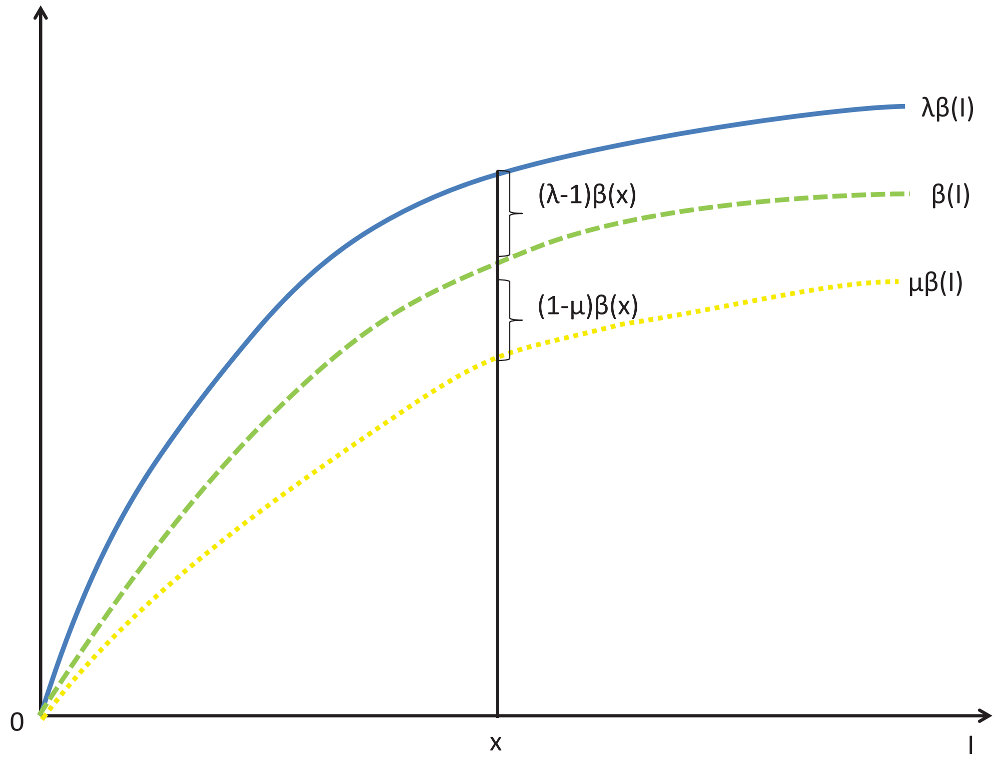


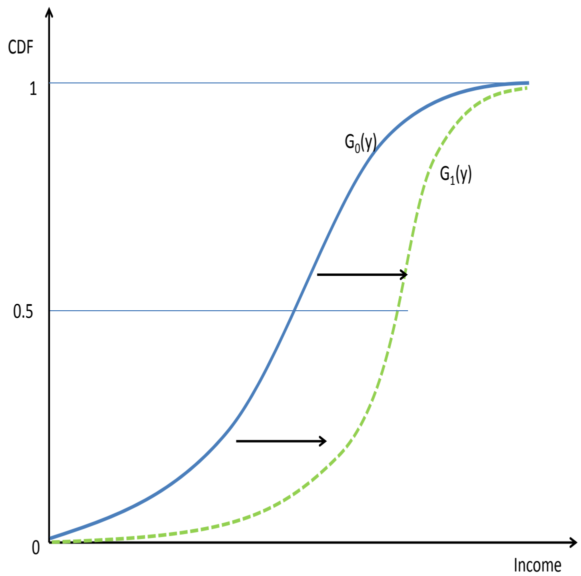

Figure 1: The Returns-from-Investment curves.

| Variable | Mean | Std Dev | Min | Max |
|---|---|---|---|---|
| Vote share of winner (in previous election) | 54.347 | 8.299 | 35.910 | 81.080 |
| Margin (in previous election) | 22.055 | 14.025 | 0.030 | 64.080 |
| Swing Congress | 0.240 | 0.428 | 0.000 | 1.000 |
| Average margin | 23.587 | 8.837 | 3.930 | 51.370 |
| Average Vote share of winner | 55.356 | 5.126 | 41.907 | 74.405 |
| Average pce | 190.047 | 51.251 | 89.037 | 350.526 |
| Literacy rate (%) | 42.122 | 14.362 | 14.077 | 90.848 |
| Population (%) | 0.199 | 0.078 | 0.042 | 0.408 |
| Rural population (%) | 80.662 | 14.001 | 21.622 | 100.000 |
| Headcount ratio (%) | 34.209 | 17.831 | 4.023 | 83.317 |
| Poverty gap ratio (%) | 8.344 | 5.756 | 0.365 | 30.985 |
| Gini (%) | 30.038 | 5.065 | 15.816 | 47.209 |
| Hindu population (%) | 84.448 | 18.017 | 0.779 | 100.000 |
| SC/ST (%) | 29.877 | 15.195 | 0.261 | 88.176 |
| Inter-quartile range/mean pce | 0.531 | 0.083 | 0.272 | 0.810 |
| Polarization (FW measure) | 0.127 | 0.025 | 0.057 | 0.209 |

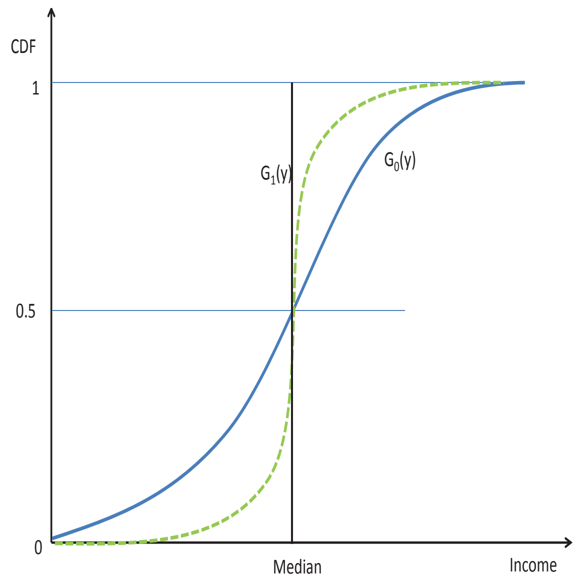
Table 1: Descriptive Statistics (1987-88). *Notes:* The information on electoral outcomes is from the Election Commission of India statistical reports. The national elections were held in 1977, 1980 and 1984-85. The data on the consumer expenditure and other demographic characteristics comes from the NSS 43rd round which was conducted during 1987-88.

| Variable | Mean | Std Dev | Min | Max |
|---|---|---|---|---|
| Vote share of winner (in previous election) | 48.794 | 8.189 | 26.540 | 69.830 |
| Margin (in previous election) | 11.270 | 8.970 | 0.190 | 40.660 |
| Swing Congress | 0.291 | 0.455 | 0.000 | 1.000 |
| Average margin | 11.433 | 5.820 | 2.573 | 30.433 |
| Average Vote share of winner | 46.211 | 6.183 | 29.837 | 62.823 |
| Average pce | 657.746 | 215.621 | 341.888 | 1,452.527 |
| Literacy rate (%) | 58.687 | 13.402 | 27.327 | 97.063 |
| Population (%) | 0.206 | 0.082 | 0.044 | 0.475 |
| Rural population (%) | 80.609 | 13.969 | 20.470 | 97.989 |
| Headcount ratio (%) | 23.024 | 16.060 | 0.000 | 67.986 |
| Poverty gap ratio (%) | 4.331 | 3.744 | 0.000 | 17.941 |
| Gini (%) | 26.162 | 5.514 | 11.621 | 43.083 |
| Hindu population (%) | 83.632 | 19.612 | 0.208 | 100.000 |
| SC/ST (%) | 29.757 | 15.920 | 0.000 | 91.977 |
| Inter-quartile range/mean pce | 0.473 | 0.104 | 0.202 | 0.954 |
| Polarization (FW measure) | 0.115 | 0.030 | 0.051 | 0.237 |

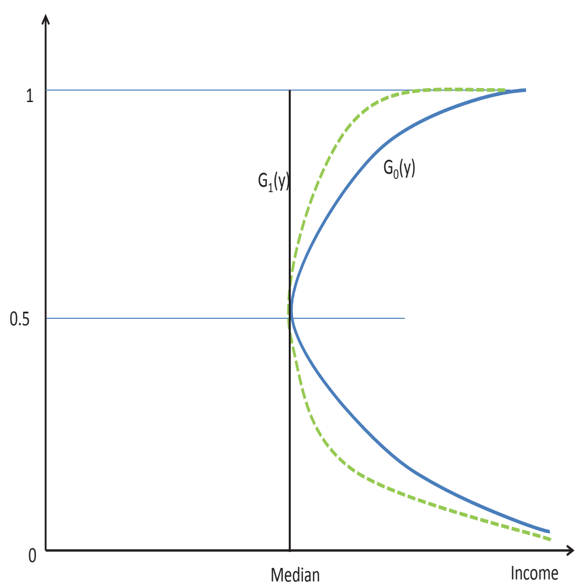
Table 2: Descriptive Statistics (2004-05). *Notes:* The information on electoral outcomes is from the Election Commission of India statistical reports. The national elections were held in 1991-92, 1996, 1998 and 1999. The data on the consumer expenditure and other demographic characteristics is from the NSS 61st round which was conducted during 2004-05.



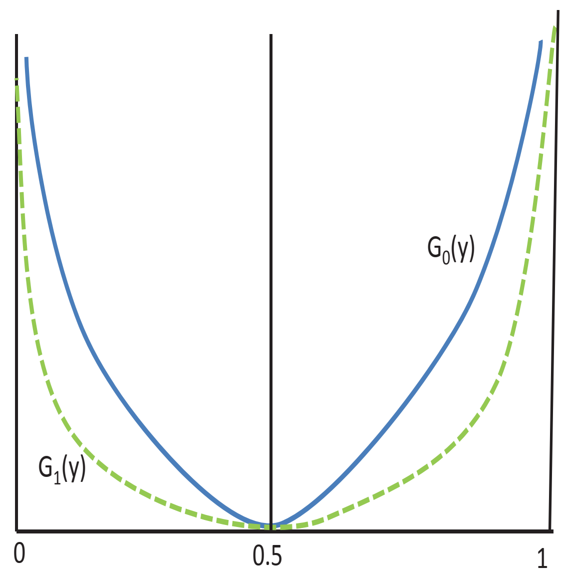

(a) Panel A




(b) Panel B




(c) Panel C




(d) Panel D

Figure 2: Foster-Wolfson "Squeeze". Panel A shows the shift in the income distribution. Panel B shows the distributions once they are median normalized. In Panel C the image has been reflected on the axis of the median. Panel D shows the Polarization Curves as in Foster-Wolfson (2009)

| | [1] | [2] | [3] | [4] | [5] |
|---|---|---|---|---|---|
| Average margin | 0.123*** | 0.106** | 0.101** | 0.110** | 0.108** |
| | (0.044) | (0.045) | (0.043) | (0.044) | (0.043) |
| Population | | 6.783 | 6.208 | 6.329 | 6.067 |
| | | (8.366) | (8.255) | (8.141) | (8.110) |
| Rural percent | | -0.172*** | -0.171*** | -0.162*** | -0.150** |
| | | (0.059) | (0.058) | (0.059) | (0.059) |
| Hindu percent | | | 0.023 | 0.033 | 0.031 |
| | | | (0.054) | (0.055) | (0.054) |
| SC/ST percent | | | -0.082* | -0.094** | -0.093** |
| | | | (0.042) | (0.044) | (0.042) |
| Headcount (poverty) | | | | 0.041 | |
| | | | | (0.025) | |
| Income gap (poverty) | | | | | 0.216*** |
| | | | | | (0.078) |
| Observations | 358 | 358 | 358 | 358 | 358 |
| Adjusted $R^2$ | 0.338 | 0.366 | 0.378 | 0.385 | 0.399 |

Table 3: Linear panel regression. *Notes:* Dependent variable is the Gini coefficient. Average margin is constructed using data from the previous 3-4 general elections. All regressions contain district fixed effects and time dummies. Robust standard errors clustered by district. *significant at 10% **significant at 5% ***significant at 1%.

| | [1] | [2] | [3] | [4] | [5] |
|---|---|---|---|---|---|
| Average vote share of winner | 0.138** | 0.122** | 0.119** | 0.127** | 0.132** |
| | (0.060) | (0.062) | (0.059) | (0.061) | (0.059) |
| Population | | -0.194 | 0.487 | 0.550 | 0.494 |
| | | (5.281) | (5.294) | (5.239) | (5.238) |
| Rural percent | | -0.162*** | -0.145*** | -0.142*** | -0.136*** |
| | | (0.048) | (0.047) | (0.048) | (0.048) |
| Hindu percent | | | -0.007 | 0.002 | -0.001 |
| | | | (0.051) | (0.053) | (0.052) |
| SC/ST percent | | | -0.082** | -0.090** | -0.090** |
| | | | (0.040) | (0.042) | (0.040) |
| Headcount (poverty) | | | | 0.030 | |
| | | | | (0.024) | |
| Income gap (poverty) | | | | | 0.182** |
| | | | | | (0.079) |
| Observations | 358 | 358 | 358 | 358 | 358 |
| Adjusted $R^2$ | 0.311 | 0.351 | 0.366 | 0.369 | 0.381 |

Table 4: Linear panel regression. *Notes:* Dependent variable is the Gini coefficient. Average vote share of winning party is constructed using data from the previous 3-4 general elections. All regressions contain district fixed effects and time dummies. Robust standard errors clustered by district. *significant at 10% **significant at 5% ***significant at 1%.

| | [1] | [2] | [3] | [4] | [5] |
|---|---|---|---|---|---|
| Average margin | 0.005*** | 0.004*** | 0.005*** | 0.005*** | 0.005*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Population | | 0.254 | 0.249 | 0.256 | 0.242 |
| | | (0.315) | (0.315) | (0.307) | (0.306) |
| Rural percent | | -0.005** | -0.005* | -0.004* | -0.004 |
| | | (0.002) | (0.003) | (0.003) | (0.003) |
| Hindu percent | | | -0.001 | 0.000 | -0.000 |
| | | | (0.002) | (0.002) | (0.002) |
| SC/ST percent | | | -0.000 | -0.001 | -0.001 |
| | | | (0.002) | (0.002) | (0.002) |
| Headcount (poverty) | | | | 0.002** | |
| | | | | (0.001) | |
| Income gap (poverty) | | | | | 0.012*** |
| | | | | | (0.003) |
| Observations | 358 | 358 | 358 | 358 | 358 |
| Adjusted $R^2$ | 0.265 | 0.280 | 0.277 | 0.294 | 0.322 |

Table 5: Linear panel regression. *Notes:* Dependent variable is the log of the inter-quartile range normalized by the mean pce. Average margin is constructed using data from the previous 3-4 general elections. All regressions contain district fixed effects and time dummies. Robust standard errors clustered by district. *significant at 10% **significant at 5% ***significant at 1%.

| | [1] | [2] | [3] | [4] | [5] |
|---|---|---|---|---|---|
| Average margin | 0.006*** | 0.005*** | 0.005*** | 0.006*** | 0.006*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Population | | 0.267 | 0.256 | 0.261 | 0.250 |
| | | (0.361) | (0.360) | (0.352) | (0.351) |
| Rural percent | | -0.006* | -0.006* | -0.005* | -0.005 |
| | | (0.003) | (0.003) | (0.003) | (0.003) |
| Hindu percent | | | 0.001 | 0.001 | 0.001 |
| | | | (0.003) | (0.003) | (0.003) |
| SC/ST percent | | | -0.002 | -0.002 | -0.002 |
| | | | (0.002) | (0.002) | (0.002) |
| Headcount (poverty) | | | | 0.002 | |
| | | | | (0.002) | |
| Income gap (poverty) | | | | | 0.009** |
| | | | | | (0.004) |
| Observations | 358 | 358 | 358 | 358 | 358 |
| Adjusted $R^2$ | 0.197 | 0.213 | 0.212 | 0.217 | 0.233 |

Table 6: Linear panel regression. *Notes:* Dependent variable is the Foster-Wolfson measure of Polarization. Average margin is constructed using the previous 3-4 general elections. All regressions contain district fixed effects and time dummies. Robust standard errors clustered by district. *significant at 10%**significant at 5% ***significant at 1%.

| | [1] | [2] | [3] | [4] | [5] |
|---|---|---|---|---|---|
| Margin in last election | 0.054** | 0.038 | 0.043* | 0.047* | 0.048* |
| | (0.023) | (0.026) | (0.026) | (0.026) | (0.025) |
| Population | | 0.890 | 1.659 | 1.796 | 1.793 |
| | | (5.310) | (5.379) | (5.359) | (5.343) |
| Rural percent | | -0.161*** | -0.142*** | -0.139*** | -0.132*** |
| | | (0.051) | (0.050) | (0.051) | (0.051) |
| Hindu percent | | | 0.001 | 0.010 | 0.007 |
| | | | (0.051) | (0.053) | (0.052) |
| SC/ST percent | | | -0.091** | -0.099** | -0.100** |
| | | | (0.043) | (0.045) | (0.043) |
| Headcount (poverty) | | | | 0.027 | |
| | | | | (0.024) | |
| Income gap (poverty) | | | | | 0.173** |
| | | | | | (0.078) |
| Observations | 358 | 358 | 358 | 358 | 358 |
| Adjusted $R^2$ | 0.304 | 0.341 | 0.359 | 0.361 | 0.372 |

Table 7: Linear panel regression. *Notes:* Dependent variable is the Gini coefficient. All regressions contain district fixed effects and time dummies. Robust standard errors clustered by district. *significant at 10% **significant at 5% ***significant at 1%.

| | [1] | [2] | [3] | [4] | [5] |
|---|---|---|---|---|---|
| Vote share of winner last election | 0.091*** | 0.069* | 0.073** | 0.077** | 0.079** |
| | (0.034) | (0.037) | (0.037) | (0.037) | (0.036) |
| Population | | 1.242 | 1.994 | 2.137 | 2.146 |
| | | (5.354) | (5.404) | (5.375) | (5.360) |
| Rural percent | | -0.160*** | -0.142*** | -0.140*** | -0.133*** |
| | | (0.049) | (0.049) | (0.050) | (0.050) |
| Hindu percent | | | -0.001 | 0.008 | 0.005 |
| | | | (0.052) | (0.053) | (0.053) |
| SC/ST percent | | | -0.089** | -0.096** | -0.097** |
| | | | (0.043) | (0.044) | (0.043) |
| Headcount (poverty) | | | | 0.027 | |
| | | | | (0.023) | |
| Income gap (poverty) | | | | | 0.171** |
| | | | | | (0.079) |
| Observations | 358 | 358 | 358 | 358 | 358 |
| Adjusted $R^2$ | 0.307 | 0.344 | 0.361 | 0.363 | 0.374 |

Table 8: Linear panel regression. *Notes:* Dependent variable is the Gini coefficient. All regressions contain district fixed effects and time dummies. Robust standard errors clustered by district. *significant at 10% **significant at 5% ***significant at 1%.

| | [1] | [2] | [3] | [4] | [5] |
|---|---|---|---|---|---|
| Swing Congress | -1.930*** | -1.673*** | -1.664*** | -1.667*** | -1.650*** |
| | (0.631) | (0.615) | (0.629) | (0.622) | (0.612) |
| Population | | 11.320 | 10.568 | 10.827 | 10.589 |
| | | (8.529) | (8.310) | (8.208) | (8.167) |
| Rural percent | | -0.161*** | -0.164*** | -0.158*** | -0.146** |
| | | (0.059) | (0.058) | (0.059) | (0.059) |
| Hindu percent | | | 0.049 | 0.058 | 0.058 |
| | | | (0.056) | (0.057) | (0.056) |
| SC/ST percent | | | -0.086** | -0.095** | -0.097** |
| | | | (0.043) | (0.046) | (0.044) |
| Headcount (poverty) | | | | 0.031 | |
| | | | | (0.025) | |
| Income gap (poverty) | | | | | 0.197** |
| | | | | | (0.081) |
| Observations | 358 | 358 | 358 | 358 | 358 |
| Adjusted $R^2$ | 0.331 | 0.361 | 0.375 | 0.378 | 0.392 |

Table 9: Linear panel regression. *Notes:* Dependent variable is the Gini coefficient. All regressions contain district fixed effects and time dummies. Robust standard errors clustered by district. *significant at 10% **significant at 5% ***significant at 1%.

| | [1] | [2] | [3] | [4] | [5] |
|---|---|---|---|---|---|
| Swing congress | -0.081*** | -0.073** | -0.074** | -0.074** | -0.073** |
| | (0.029) | (0.029) | (0.030) | (0.030) | (0.030) |
| Population | | 0.486 | 0.471 | 0.480 | 0.472 |
| | | (0.376) | (0.370) | (0.363) | (0.361) |
| Rural percent | | -0.005* | -0.005* | -0.005* | -0.005 |
| | | (0.003) | (0.003) | (0.003) | (0.003) |
| Hindu percent | | | 0.002 | 0.002 | 0.002 |
| | | | (0.003) | (0.003) | (0.003) |
| SC/ST percent | | | -0.002 | -0.002 | -0.002 |
| | | | (0.002) | (0.002) | (0.002) |
| Headcount (poverty) | | | | 0.001 | |
| | | | | (0.001) | |
| Income gap (poverty) | | | | | 0.008** |
| | | | | | (0.004) |
| Observations | 358 | 358 | 358 | 358 | 358 |
| Adjusted $R^2$ | 0.175 | 0.195 | 0.197 | 0.198 | 0.213 |

Table 10: Linear panel regression. *Notes:* Dependent variable is the Foster-Wolfson measure of Polarization. All regressions contain district fixed effects and time dummies. Robust standard errors clustered by district. *significant at 10% **significant at 5% ***significant at 1%.

| | [1] | [2] | [3] | [4] | [5] | [6] |
|---|---|---|---|---|---|---|
| Average margin | 0.256*** | 0.230*** | 0.206*** | | | |
| | (0.039) | (0.038) | (0.040) | | | |
| Average vote share of winner | | | | 0.336*** | 0.303*** | 0.270*** |
| | | | | (0.043) | (0.046) | (0.048) |
| Population | -3.427 | -2.591 | -2.313 | -0.188 | -0.172 | -0.228 |
| | (6.864) | (6.481) | (6.442) | (7.019) | (6.661) | (6.560) |
| Rural percent | -0.112** | -0.111** | -0.109** | -0.107** | -0.107** | -0.109** |
| | (0.051) | (0.049) | (0.049) | (0.046) | (0.045) | (0.045) |
| Hindu percent | 0.013 | 0.020 | 0.009 | -0.005 | 0.012 | 0.005 |
| | (0.079) | (0.077) | (0.076) | (0.072) | (0.070) | (0.068) |
| SC/ST percent | -0.068 | -0.088 | -0.088* | -0.076 | -0.096** | -0.096** |
| | (0.056) | (0.055) | (0.053) | (0.047) | (0.048) | (0.046) |
| Headcount (poverty) | | 0.053** | | | 0.049* | |
| | | (0.025) | | | (0.027) | |
| Income gap (poverty) | | | 0.247*** | | | 0.254*** |
| | | | (0.085) | | | (0.091) |
| Overall fit, $P > chi^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Observations | 358 | 358 | 358 | 358 | 358 | 358 |

Table 11: Panel regression, linear two–step GMM. *Notes:* Dependent variable is the Gini coefficient. All regressions contain district fixed effects and time dummies. Robust standard errors corrected according to Windmeijer (2005). *significant at 10% **significant at 5% ***significant at 1%.

| | [1] | [2] | [3] | [4] | [5] | [6] |
|---|---|---|---|---|---|---|
| Average margin | 0.006*** | 0.006*** | 0.005*** | | | |
| | (0.001) | (0.001) | (0.002) | | | |
| Average vote share of winner | | | | 0.008*** | 0.008*** | 0.007*** |
| | | | | (0.002) | (0.002) | (0.002) |
| Population | 0.281 | 0.316 | 0.331 | 0.263 | 0.289 | 0.285 |
| | (0.292) | (0.293) | (0.294) | (0.297) | (0.300) | (0.300) |
| Rural percent | -0.005** | -0.006** | -0.006** | -0.006** | -0.006** | -0.006** |
| | (0.002) | (0.002) | (0.003) | (0.003) | (0.002) | (0.003) |
| Hindu percent | -0.004 | -0.004 | -0.004 | -0.004 | -0.004 | -0.003 |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| SC/ST percent | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Headcount (poverty) | | 0.001 | | | 0.001 | |
| | | (0.001) | | | (0.001) | |
| Income gap (poverty) | | | 0.004 | | | 0.005 |
| | | | (0.004) | | | (0.004) |
| Overall fit, $P > chi^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Observations | 358 | 358 | 358 | 358 | 358 | 358 |

Table 12: Panel regression, linear two–step GMM. *Notes:* Dependent variable is the Foster-Wolfson measure of Polarization. All regressions contain district fixed effects and time dummies. Robust standard errors corrected according to Windmeijer (2005). *significant at 10% **significant at 5% ***significant at 1%.