Title
Essays in Development Economics

Author
Chakravarti, Ritadhi

Publication Date
2017

Peer reviewed|Thesis/dissertation

Essays in Development Economics

by

Ritadhi Chakravarti

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Alain de Janvry, Chair
Professor Ernesto Dal Bo
Aprajit Mahajan

Spring 2017

Essays in Development Economics

Abstract

Essays in Development Economics

by

Ritadhi Chakravarti

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Alain de Janvry, Chair

The relationship between democracy and inequality and the role of democratic participation in fostering redistribution has been a critical question of interest to economists over time. My dissertation focuses on a key aspect of democratic participation - the ability of citizens to form political parties and contest elections - and empirically identifies the relationship between the distribution of political power, and the distribution of public resources. To this effect, I study the economic impacts emanating from the political mobilization of historically underprivileged and socio-economically disadvantaged citizen groups in India - the low caste and indigenous communities - which has occurred over the past three decades. While comprising over half the nation's population, citizens from these groups have faced institutionalized discrimination through the hierarchical caste system, severely restricting their access to education, public institutions and basic services. This process of political mobilization by low caste citizens resulted in the formation of powerful caste-based political parties which contested elections in an attempt to capture political power through the democratic process, and subsequently, use their nascent state power to redirect public resources to their communities. The core chapters of this dissertation empirically document the economic and social impacts emanating from this change in the distribution of political power through democratic processes, resulting in the political empowerment of marginalized citizen groups.

The first chapter of my dissertation identifies the economic impacts of this process of political mobilization. Using the outcome of close elections between parties exclusively representing low caste interests and other mainstream parties as the source of exogenous variation, the paper shows that the marginal legislator from low caste parties increases the share of state expenditures allocated towards targeted and untargeted welfare schemes benefitting low caste citizens by 2 and 1 percent. In a departure from existing studies in the literature which have exploited the system of mandated representation of low caste politicians in India, my paper shows that the mandated representation of low caste politicians has no impact on welfare transfers for low caste citizens when unaccompanied by a strong ex-ante party commitment in favour of low caste welfare. This paper also highlights the trade-offs emanating from the redistributive spending choices of fiscally constrained legislators by identifying the

negative impact of the marginal legislator from low caste parties on public investments in road construction and electricity generation. Using tests of heterogeneity, the paper argues that the redistributive spending preferences of legislators from low caste parties is one of the mechanisms explaining the negative causal impact between the electoral success of these parties and regional inflows of private investment.

Low caste citizens in India have often been victims of targeted violence by social elites in an attempt to preserve the existing social hierarchy and consolidate economic rents. This forms the motivation for the second chapter of my dissertation, identifying the causal impact of low caste parties on the incidence of targeted crimes against low caste citizens. The empirical results from this paper show that the marginal legislator from low caste parties reduces the incidence of violent crimes against low caste citizens by 4-14 percent and contribute significantly to the protection of civil rights. The paper also identifies the mechanisms explaining the reduced form impact - first, through higher rates of arrest, prosecution and diligence with which law enforcement agencies pursue such cases of targeted violence; second, through changes in the attitudes of low-caste citizens towards state institutions which affect the enforcement of criminal law.

Finally, the third chapter of my dissertation focuses on a major question of interest in the field of development economics and economic growth - the role of agricultural growth in the process of structural transformation. Using data from a panel of 273 Indian districts over a twenty-five year period and rainfall shocks as an instrument for agricultural growth, I present causal evidence that rainfall-induced agricultural growth has a significant impact on the structural transformation process by increasing the share of rural workers employed in the manufacturing sector. My results show that the non-farm employment gains from higher agricultural growth is concentrated amongst unskilled workers with relatively low levels of education, signalling the poverty alleviating aspect of agricultural growth. By studying the differential impact of agricultural productivity across regions with initially low levels of agricultural productivity and urbanization, I also document the potential of agricultural growth to facilitate convergence across regions in terms of rural employment diversification.

Acknowledgments

The saga of my dissertation would not have been possible without the wisdom and support of many. I make a humble attempt to recognize those who made this journey possible.

My father, Professor Ranabir Chakravarti, who stood out as a role model through my life and provided me with glimpses from a very tender age into the world of academia. I doubt if I would chosen this path had it not been for him. A fantastic teacher, in addition to an incisive and productive researcher - he remains to me, the ideal academic, and a source of inspiration for the years to come.

My three advisors - Professors Alan de Janvry, Ernesto Dal Bo, and Aprajit Mahajan, who labouriously shared their thoughts regarding my academic work, and continuously supported me through this process. Ernesto - I wish I had spent more time discussing my ideas with you; it is hard to put in words how much I learnt from our discussions.

Abhay Aneja - study partner, colleague, co-author, friend and guide; enough said, I just don't know what I would have done without you!

Pratichi, for reminding me to corroborate sources before reaching any conclusion and forever be doubting of 'facts'; for putting up with my tantrums through year five, and providing in exchange unconditional love, support and care.

Ryan Olver, who sat with me for five years in various corners of the International House, reiterating a thousand times that 'it's not that bad' and teaching me first year trade theory.

The late Bangladeshi author, Humayun Ahmed, for those stunning novels on human relations which provided me with an escape route into alternate reality when the going got tough.

Professor Pranab Bardhan, who humoured me a countless number of times and shared his views on my (often half-baked) ideas.

Professors Noam Yuchtman, Reed Walker and Rui de Figueiredo for their very careful reading of the drafts of the second chapter.

Paulette Cha, for bearing with my handwriting and helping me through micro theory and public finance; for repeatedly checking in on me over the past five years and sending in her good wishes before every major assignment.

Professor Debashree Dattaray, Professor Debalina Chakravarty and Reajul Alam for all those afternoons and evenings spent in the International House. Was any semester more enjoyable than Fall 2013?

Seth Garz and Francesca Jensenius, for sharing invaluable data, numerous suggestions, encouragement and advice.

Professor Elisabeth Sadoulet, for guiding me through the first three years of my PhD life and teaching me causal identification.

Across the Univeristy of California-Berkeley campus, participants in the Development Workshop (Department of Agricultural and Resource Economics), the Development Lunch (Department of Economics), Political Economy Research Lunch (Department of Economics), and the Department Seminar (Department of Agricultural and Resource Economics) for their thoughtful feedback on my presentations.

Nagraj Rao, for arm-twisting me to apply to the department of Agricultural and Resource Economics at the University of California-Berkeley, even offering to refund the application fee in the (most likely) event of my application being rejected.

Ashesh Prasann and Yue Hou for sharing their statements of purpose six years ago and carefully reading mine.

Sajad Zafranchi, Varun Kanade and Ashley De Mello, for their hours of company during the first two years and providing a patient ear to all my chatter on Indian history and politics.

Yasir Khan, for all his words of wisdom, and for dragging me back into my PhD life whenever I would seek refuge in the world of Indian Railways, cricket and Bengali literature.

Rahul Verma, for befriending me on my first day at UC Berkeley and exposing me further to the world of electoral politics.

Professor Jeff Perloff, for the timely warning, offered in the most understanding terms, that I would not make it through the prelims. It was a thoroughly needed wake-up call!

Christopher Whaley, for guiding me through micro-theory.

David Schonholzer and Gavin McCormick, for constantly reminding in the corridors of the International House to 'be patient', and reassuring me that 'I will be fine!'

My mother, Dr. Tutul Chakravarti, for being a constant skeptic and critic - it becomes annoying as you get older, but it does keep you grounded!

Manraj Gill, for helping me take a break from my PhD through high-quality fast bowling.

Vishwanath Bulusu, Karanpal Singh and Jin Mingjao, for that rip-roaring first year road trip where they implicitly trusted my driving.

The kitchen, cleaning and maintenance staff at International House, for keeping me well-fed and healthy during my five years at Berkeley.

And finally, Rashmi Ekka, who shared her life's story and in the process, taught me to understand the lives of India's *bahujan* from a completely different perspective. For fostering in me the values of empathy, and providing unflinching support through the worst of times, constantly reminding me to trust my abilities and push ahead. Rashmi, much of this dissertation is a reflection of your efforts - I would not have been here without you.

Notice of Name Change

This is to notify that I had decided to change my name from ‘Ritadhi Chakravarti’ to ‘Subrata Kumar Ritadhi’ on September 30, 2016. As I have not had the opportunity to officially effect this name change in a court of law, the PhD dissertation is being submitted under my present official name, ‘Ritadhi Chakravarti’.

Contents

List of Figures

List of Tables

Chapter 1

Identity Politics, Targeted Redistribution and Private Investment: Evidence from India's Silent Revolution

1.1 Introduction

The role of political institutions in reducing inequality remains a fundamental question in economics (Acemoglu and Robinson, 2008; Bardhan, 2016). Classical economic theory predicts that increased democratic participation amongst citizens would lead to higher levels of redistribution through an improved representation of the median voter's preferences (Meltzer and Richard, 1981). However, in a review of existing empirical works, Acemoglu et al. (2013) note the absence of a robust correlation between democracy and redistribution. Explaining the lack of such a relationship, Acemoglu et al. (2013) provide a theoretical framework whereby elite capture of institutions through lobbying or control of political parties restrict the redistributive gains attained from political reforms aiming to enhance democratic participation by citizens in public policy. To overcome this, Acemoglu et al. (2013) propose that reforms to political institutions guaranteeing representation to select citizens (or *de jure* political power), need to be complemented with broader changes in the distribution of political power in society (or *de facto* political power) to realize the full redistributive impact of such reforms. Exploiting a unique setting in India concerning the political representation of economically deprived ethnic minorities, this paper empirically tests this proposition through a comparison of the redistributive gains for marginalized citizens obtained through the respective channels of *de jure* and *de facto* political power.

Over the past century, the inclusion of marginalized citizens and women in public institutions has been a critical challenge for societies across the world. As noted by Duflo (2012), significant progress has been made in this direction over the past two decades with respect to women's representation in policy-making through the adoption of quotas in political institutions, guaranteeing a threshold level of representation for women. In the Indian context, there exists an analogous system of electoral quotas to ensure the inclusion of economically

deprived low caste ethnic minorities - namely *Dalits* (Scheduled Castes or SCs) and *Adivasis* (Scheduled Tribes or STs) in federal and state legislatures.[1] The constitutionally mandated electoral quotas sets aside or 'reserves' a pre-determined share of electoral constituencies from which only low caste ethnic minority candidates can contest elections. Consequently, the institution of electoral reservations mechanically ensures the presence of a fixed share of low caste candidates across state (federal) legislatures - proportional to their regional (national) population shares - altering in the process the *de jure* distribution of political power in favour of low caste citizens. The welfare gains obtained through these electoral quotas for low caste citizens (and women) has been the subject of numerous studies in the literature with researchers obtaining divergent results.[2]

In a departure from the existing literature focusing exclusively on the welfare gains obtained through changes in *de jure* political power, the present paper empirically identifies the economic impacts emanating from changes in *de facto* political power, stemming from the political mobilization of low caste citizens, led by the *Dalits* and the Other Backward Castes (OBCs). This process of political mobilization, initiated in the mid-1980s, resulted in the formation of political parties on the basis of caste identity, which contested elections to challenge the existing political elite who had dominated India's polity since independence and were drawn mostly from the privileged upper castes (Jaffrelot, 2003). The primary goal of these caste-based parties was to capture political power through electoral competition and subsequently, redirect state resources towards low caste groups and increase the representation of low caste citizens in public institutions. A key policy tool for these parties to achieve the latter objective was the use of affirmative action quotas in public enterprises, leading us to term them as 'Affirmative Action' (AA) parties for the remainder of the paper.[3]

As described by Jaffrelot (2003), the electoral success of caste based 'AA' parties between 1990 and 2010 resulted in a shift in *de facto* political power from upper caste elites to low caste citizens, over and beyond what was achieved through the institution of electoral quotas. We exploit the electoral success of AA parties to identify the welfare impacts for low caste citizens emanating from this shift in *de facto* political power, relative to that obtained through electoral quotas. Through this comparison, the paper also addresses a key debate focusing on the impact of a legislator's ascriptive identity on public policy. Using the system of electoral reservations to control for the caste identity of individual legislators, the paper provides causal evidence that the mandated representation of politicians from low caste backgrounds translates into economic benefits for low caste citizens only when low caste politicians belong to political parties with a strong policy preference for low caste welfare. While providing empirical evidence supporting the hypothesis forwarded by Acemoglu and

[1] The quotas were introduced after the nation's independence in 1947 in federal and state legislatures. Since 1993, under the 73rd and 74th amendments, these quotas have been expanded to local government and also includes provisions for women. The electoral quotas are in addition to an aggressive program of affirmative action in public employment and higher education for these same ethnic minority groups

[2] Initial papers studying the impact of mandated political representation of ethnic minorities such as Besley et al. (2005); Chin and Prakash, (2011) Clots-Figueras (2011); Chattopadhyay and Duflo (2004); and Pande (2003) report welfare gains from this policy for low caste citizens. However, more recent papers such as Dunning and Nilekani (2013), Jensenius (2015), and partially, Bardhan et al. (2010) do not find any welfare gains on comparable policy outcomes.

[3] These parties have been categorized as 'soft Left' political parties by Besley and Burgess (2000).

Robinson (2008) and Acemoglu et al. (2013), this finding highlights the critical role played by the policy objectives of political parties in determining policy outcomes, consistent with citizen candidate models predicting full policy divergence by competing political parties from the preferences of the median voter (Lee et al., 2004).

For causal identification, the paper extends a regression discontinuity design based on the outcome of close elections between AA and non-AA parties. This generates quasi-random regional variation in the number of legislators elected from parties exclusively representing low caste interests and by extension, the quasi-random variation in *de facto* political power. The paper uses this quasi-random variation to causally identify the economic impacts resulting from the election of an additional AA party legislator, both for low caste populations, and the broader regional economy. The economic relevance of this change in *de facto* political power in favour of low caste citizens is augmented due to the extensive socio-economic deprivation faced by low caste groups in India. This is discussed extensively by Kijima (2006) who provides evidence on the differences in returns to education and consumer expenditure between low caste SC/ST and non-SC/ST households. These findings are supported by Thorat et al. (2009) who show that a disproportionately higher share of low caste SC households are located below the poverty line, have lower land holdings, and are under-represented in both public and private sector employment.

The key empirical results of the paper show that the electoral success of caste-based AA parties increases the share of public resources allocated towards low caste households [4] with the marginal AA party legislator being associated with a causal increase in the share of state expenditures allocated towards both targeted and untargeted low caste welfare programs. In monetary terms, the impact of the marginal AA party legislator is equivalent to a 0.6 percent increase in the monthly per capita consumption for an urban low caste household.[5] The marginal legislator however impacts public expenditures only when AA parties have won at least 30 percent of the elections in the state for that electoral cycle, underlining that a threshold level of political power is necessary before individual legislators can significantly influence social welfare allocations in favour of marginalized citizens.

To compare the respective impacts of *de jure* and *de facto* political power on welfare expenditures, the paper focuses solely on elections contested in constitutiences reserved for low caste candidates through the electoral quotas. This implicitly controls for the caste identity[6] of elected politicians and identifies how welfare expenditures are affected through variations in the legislator's party affiliation. If changes to *de jure* political power and the caste identity of legislators - guaranteed through electoral quotas - are sufficient to influence public policy, we would expect no difference in welfare expenditures between legislators elected from reserved constituencies across AA and non-AA parties. The empirical results show that relative to low caste legislators from mainstream national parties, low caste legislators from AA parties allocate a significantly higher share of state expenditures towards targeted welfare transfers for low caste citizens, consistent with Acemoglu and Robinson (2008) and Acemoglu et al. (2013).

[4] To estimate economic impacts, we include all three of *Dalits*, *Adivasis* and OBCs in the category of low caste citizens.

[5] This share would be doubled if the reference point was a rural low caste household.

[6] Caste identity here refers more to the the *varna* than the *jati*.

In addition to the state expenditure results, the paper also documents the positive impact of a change in *de facto* political power for low caste citizens on public employment. The marginal AA party legislator increases the number of low caste workers employed in the public sector by 6 percent with the effect being driven by urban, secondary educated low caste workers. These results show that the economic effects of AA parties' electoral success extend beyond budgetary allocations and enhances the representation of marginalized citizens in public institutions.

Finally, through an analysis of the spending preferences of AA party legislators, the paper documents the tradeoffs arising from the electoral success of political parties with strong preferences for welfare spending in settings where legislators are fiscally constrained. We show that the restriction on state legislators in India to impose direct taxes or borrow from private markets lead AA party legislators to engage in a redistribution of public resources away from non-exclusive public investments in physical infrastructure, in favour of select welfare schemes benefitting low caste citizens. Specificially, the marginal AA party legislator has no impact on aggregate state expenditures, revenues, and deficits, but is associated with a 3 percent reduction in the share of public expenditures allocated towards electricity generation and road construction. As physical infrastructures are strongly correlated with firm performance in both cross-country and firm-level studies, we identify the impact of AA parties on private investment. The reduced form results shows that the marginal legislator from AA parties affects private investment along both the extensive and the intensive margins with the number of factories in operation and state-level capital formation in the manufacturing sector declining by 1 and 3 percent in response to a quasi-random increase in AA party representation. However, this negative impact is observed only when AA parties have also won over 30 percent of the elections in the state and is concentrated amongst industries which are most reliant on physical infrastructure - namely industries which are capital intensive and have high levels of fuel consumption. These heterogeneous effects offer suggestive evidence that the redistributive expenditure choices made AA party legislators can serve as a potential mechanism to explain their corresponding negative impact on manufacturing investment, indicating that higher public expenditures for underprivileged citizens in this context comes at a cost to private investment.

Within the literature, this paper's primary contribution lies in providing causal evidence supporting the theoretical predictions of Acemoglu and Robinson (2008) and Acemoglu et al. (2013) and establishing that political representation for marginalized citizens through changes in *de jure* political power needs to be complemented with changes in *de facto* political power to impact public resource allocation for disadvantaged populations. In this regard, this paper, to the best of our knowledge is the first to study the interaction between political institutions and the policy preferences of political parties, and causally identify their respective impacts on economic outcomes for marginalized citizens. In the process, the paper also contributes to the growing literature studying the economic impacts of political representation for social and ethnic minorities. Contrary to existing studies which focus on how policy outcomes are affected by the ascriptive identity of individual legislators, this paper identifies the welfare gains obtained through the electoral success of political parties with a specific policy mandate for disadvantaged citizens. While Dunning and Nilekani (2013) provide some initial insight that party affiliations might be dampening the impact of electoral quotas on economic outcomes, their paper interprets the results through a model of political clientelism

and not party agendas mediating politicians' spending choices.

By documenting the impact of political party policy preferences on budgetary allocations, this paper also adds to the broader literature studying the impact of political parties on economic outcomes. Ferreira and Gyourko (2009) for instance show that the party affiliations of mayors in U.S. cities have no effect on their spending preferences or on broader economic outcomes. A recent study by Beland (2015) shows that Democratic party governors in the U.S. are causally associated with a reduction in the earnings gap between blacks and whites but does not discuss the channels through which this might be occuring. Using the same identification strategy, Beland and Oloomi (2016) however find no effect of party affiliation on the spending preferences for U.S. governors. By identifying a significant impact of caste-based AA parties on public spending and the public hiring of low caste workers, this paper provides alternative evidence that party agendas can significantly impact both public expenditure choices and household economic outcomes. Importantly, the paper looks at the impact of party platforms on state spending which is an explicit function of the budgetary choices made by legislators, permitting a clearly delineation of the mechanism through which party platforms map into economic outcomes.

Finally, by studying the potential costs on private investment arising from the redistributive spending preferences of AA party legislators, this paper also provides empirical evidence in support of existing theoretical models which predict that redistributive spending by politicians can lead to economic inefficiences (Persson and Tabellini, 2000; Lizzeri and Persico, 2001). Lizzeri and Persico (2001) for instance present a theoretical framework where the reallocation of public expenditures from investments yielding long-run returns towards pork barrel spending results in the under-provision of the long-term public investment from its socially optimal level. This paper, by linking the expenditure choices of AA party legislators with the subsequent decline in private investment, shows how redistributive spending can affect broader parameters of the regional economy.

The remainder of the paper is organized as follows: Section 2 presents a conceptual framework and discusses the potential channels through which AA party legislators can affect economic outcomes; Section 3 contains a brief overview of the electoral process in India and a description of the AA parties; Section 4 describes in detail our empirical strategy centered around close contests between AA and non-AA parties and Section 5 performs a series of empirical tests to validate our identification strategy; Section 6 presents the key results of the paper which we validate with a number of robustness checks in Section 7. Section 8 offers some concluding comments.

1.2 Conceptual Framework: Electoral Success of Political Parties and Economic Outcomes

1.2.1 State Legislators and Public Expenditures

The key agent of interest in this paper are state legislators belonging to caste-based AA parties who have a strong policy preference for improving welfare outcomes for economically disadvantaged low caste populations. Under India's federal polity, state governments are responsible for law and order, education, public health, nutrition, social welfare, intra-state

physical infrastructure, as well as allocations to local governments (Clots-Figueras, 2011). This makes state legislators a vital stakeholder in the overall development process. A fundamental responsibility of state legislators in India is to determine public spending undertaken by the state government. Every year, legislators across state legislatures vote to determine various allocations to be made by the state government across different expenditure categories for the upcoming year. The annual state budget is prepared by the incumbent government and presented in the legislature in the first quarter of each calendar year. Subsequently, after a few weeks of debate, the budget is put to vote and adopted only if a simple majority of legislators vote in its favour. During the period of debate, individual legislators also have the opportunity to table additional allocations which are adopted through majority voting (Clots-Figueras, 2011).

Based on the above institutional framework and as per Rehavi (2008), spending allocations, Y, on any budgetary category k, in a N-member legislature in state s and year y can be mathematically expressed as:

$$Y_{st}^{k} = \phi_{st}^{k} + \sum_{j=1}^{N} \pi_{jst} y_{jst}^{k} \tag{1.1}$$

In equation (1.1), ϕ is a state-year specific factor determining allocations to spending category Y, irrespective of the preferences of individual legislators. The individual spending preference of each legislator on category k is represented by y. π is the weight accorded to each legislator's desired level of spending and can be considered to be the share of legislators who support the individual legislator's preferred level of spending, y^k and is a function of the bargaining power and legislative support available to each individual legislator. In the simplest framework, the budgetary process can be considered equivalent to a two-stage game where legislators can initially approach the state's finance department, responsible for framing the budget, with their preferred levels of spending on each category k. Alternatively, they can present their spending proposals during the discussion of the budget. In either event, the final outcome is conditional on the legislative support accorded to the individual legislator's desired level of spending by the remaining legislature. Accordingly, based on (1.1) and the majority voting rule, the final spending level on category k will be close to a legislator's preferred level if π is sufficiently large, signifying a sufficient degree of support from the remaining legislators.

Applying this framework to compare the respective impacts of *de jure* and *de facto* political power with respect to the institution of electoral reservations, the first observation is that while reservations guarantees the election of a fixed share of low caste legislators, this share is below 30 percent in the majority of states and moreover, is typically split across two or more political parties. As shown by Jaffrelot (2003), very few low caste legislators belonging to traditional mainstream parties were elected from non-reserved constitutencies, implying that prior to the electoral success of AA parties, low caste legislators were never a numerical majority in legislatures. Resultantly, unless there is a convergence in the spending preferences between low and high caste legislators, equation (1.1) suggests that electoral reservations by themselves would be insufficient to shift public resource allocation in favour of low caste citizens.

Moreover, a key distinction between traditional mainstream parties and caste-based AA

parties is that the leadership positions of the former are dominated by high caste citizens. In such a situation, even if the majority of low caste legislators elected through the reservation system belong to a mainstream party which has obtained a legislative majority, their preferred level of budgetary allocations might not be realized if other party legislators and the party leadership assign a low weight to the individual preferences of low caste legislators during the process of budgetary allocations. This possibility is supported through the findings of Jensenius (2015), who shows through interviews of politicians and bureaucrats that the quality of low caste politicians elected through electoral reservations have been consistently questioned, with the majority of peers and bureaucrats labeling these politicians as 'inefficient' and 'weak' (Jensenius, 2015). This suggests that the bargaining power of legislators elected from reserved constituencies would be low in a traditional mainstream party and even when the mainstream party enjoys a legislative majority, final budgetary allocations on Y^k would diverge from low caste legislators' preferred allocation, y^k, unless the party's policy preferences on category k are sufficiently aligned with the preferences of low caste legislators.

On the contrary, a caste-based AA party by virtue of its social composition and overarching policy preference of exclusively represening the interests of low-caste citizens can overcome the hindrances mentioned above. First, as both the majority of legislators and party leaders in such parties hail from lower caste origins, it is conceivable that a higher weight would be assigned to legislator preferences for increasing budgetary allocations for low caste citizens. This is also consistent with the political stance of AA parties which have a strong welfare agenda directed towards low caste citizens and rely heavily on targeted public expenditures to achieve their political objectives (Jaffrelot, 2003). Moreover, if the concerned AA party enjoys a legislative majority, higher spending allocations proposed by AA party legislators on issues pertaining to the welfare of low caste populations would have a higher propensity of receiving legislative support - in other words, π would be sufficiently high, resulting in the convergence of Y^k and y^k. This explanation would be consistent with Acemoglu and Robinson (2008) and Acemoglu et al. (2013) that changes in the distribution of political power would be necessary to attain welfare gains for marginalized citizens.

Section 1.6.1 shows that the empirical evidence is consistent with this conceptual framework. Section 1.6.1 shows that the marginal legislator from AA parties significantly impacts public expenditures only when a threshold level of legislators are also elected from AA parties; section 1.6.1 shows that amongst legislators elected through electoral quotas, only those belonging to AA parties have a positive impact on targeted welfare spending for low caste citizens, supporting the contention that little weight is placed on the policy preferences of low caste legislators in traditional mainstream parties.

Finally, while our empirical results show that the electoral success of AA parties lead to higher welfare expenditure for low caste populations, it is worth noting that voter demographics and politician quality have the potential to impede this process. This is discussed in detail by Jensenius (2015) who show through politician surveys that low caste politicians are often hindered by public expectations and electoral incentives from redistributing towards their own groups. For instance, some low caste legislators report avoiding targeted welfare transfers as they do not wish to be viewed as working solely for their own communities. Other low caste legislators report that they have to depend on electoral support from non low caste populations, which precludes their ability to target state resources towards their

respective communities (Jensenius, 2015).

1.2.2 State Legislators' Impact on Public Hiring: Channels of Influence

In addition to formulating public policy and voting on legislation in the state legislature, state legislators in India are also responsible for both initiating and supervising developmental activities in their own constitutencies. This is borne out through field surveys by Jensenius (2015) who reports that state legislators allocate a large share of their time to constituency visits where they oversee local development and are influential in recommending citizens to employment opportunities and local contracts. This is in addition to lobbying the state government for higher allocation of public goods to their respective constitutencies. A key leveraging tool permitting elected politicians to influence executive decisions lies in their ability to significantly influence the careers of state bureaucrats by recommending them for promotions and transfers. As shown by Iyer and Mani (2012), elected state-level politicians have considerable control over the future postings and career advancement of bureaucrats and attempt to use punitive transfers to undesirable locations or departments if bureaucrats do not comply with their recommendations. In this regard, we identify the impact of AA party legislators on the public hiring of low caste workers, which has been a key political agenda for AA parties (Jaffrelot, 2003) and a major channel through which politicians in general transfer public resources to select beneficiaries (Enikolopov, 2014).

We hypothesize two channels through which legislators from AA parties can facilitate this outcome. The first is a legislative channel whereby AA parties can use their numerical strength in the legislature to enact new laws, either pertaining to a general expansion of the public sector with provisions for preferential hiring of low caste workers, or an expansion of existing affirmative action quotas. If the former approach is undertaken, we would expect to observe some positive spillovers for high caste workers, unless all the new positions are filled by low caste workers. If the second approach is preferred and existing quotas for low caste workers are increased while holding the size of the public sector constant, it would be equivalent to a direct reallocation of employment and public salaries from high to low caste workers. Both the above approaches require legislative action and thereby, would be conditional on whether AA parties have the necessary legislative support in the state legislature, in a manner similar to that described in Section 1.2.1.

A second channel is the supervisory channel, emanating from the ability of legislators to influence the career trajectories of bureaucrats. The presence of affirmative action quotas by themselves may not guarantee improved employment prospects for its designated beneficiaries as a number of positions set aside for low caste candidates are often left unfilled. For instance, in 2010, a leading Indian daily, The Hindu, reported that only 7 percent of federal public sector positions were occupied by individuals belonging to the OBC category, despite the stipulated share being 27 percent. Similarly, positions for SCs and STs also remain unfilled and their represenation in public services is often well below the shares stipulated by law. The reason for this is multi-fold - while some positions are left vacant due to a lack of qualified candidates, bureaucratic inefficiency and indifference is also a major cause for these unfilled vacancies. Under such circumstances, legislators from parties with a policy

preference for increased representation of low caste workers in public institutions can play a critical role in ensuring the implementation of existing policies through improved supervision of the bureaucracy. Importantly, influencing the implementation of existing policies is a function of effort invested by the legislator through lobbying of the concerned bureaucrat and should be independent of the legislative support enjoyed by a political party in the state legislature. In this regard, if AA parties affect public employment through the supervisory chanel, we would expect the marginal legislator from AA parties to have an impact on public employment for low caste workers, irrespective of the overall electoral success of AA parties in the state and their ability to enact new legislation.

1.2.3 Potential Costs from Higher Welfare Expenditures: Impact on Private Investment

If AA party legislators choose to use the public exchequer to funnel state resources to low caste citizens, the net impact of such targeted spending on the regional economy would depend on the source of funds used to finance such expenditures. If legislators are able to impose taxes or run deficits, the higher spending on low caste citizens can come through an expansion in public expenditures.[7] The Indian Constitution however assigns limited fiscal powers and borrowing abilities to state legislators. State legislators are incapable of imposing direct taxes and rely solely on indirect consumption taxes[8] for their own source of revenue.[9] The inability of state legislators to levy direct taxes is compounded through the inability of states to borrow from private marketes. State deficits have to be financed by public sector banks or the federal government and as the reduction of fiscal deficit was a key agenda of the economic reforms undertaken since 1991, there has been considerable pressure on state governments since to contain their deficits. The sum of these regulations imposes a signficant budgetary constraint on state politicians' budget, making them incapable of financing higher spending through revenues or deficits.

Intuitively, this suggests that higher welfare spending for low caste citizens by AA party legislators would be implemented through a redistribution of public resources. This paper's interest is to ascertain whether it occurs through a redistribution of public expenditures from non-exclusive public goods to targeted welfare spending. In this regard, we empirically identify the impact of AA party legislators on major expenditure categories to identify reallocations in spending priorities. Of particular interest within redistributive spending is the impact of AA party legislators on investments in physical and social infrastructure which are critical to the medium and long-term health of an economy.

Economic theory predicts that the reallocation of public resources from productive investments to pork-barrel spending can lead to economic inefficiencies (Lizzeri and Persico, 2001; Persson and Tabellini, 2000 etc.). The large literature on investment climate and firm performance also highlight the critical role played by physical and social infrastructure in

[7] The imposition of higher taxes or a higher level of fiscal deficit would have a distortionary impact on the economy but that would affect economic outcomes through channels other than public expenditures.

[8] These are chiefly the sales taxes/value added tax and excise duties.

[9] For instance, in 2013, only 50 percent of state revenues across all states were generated through own tax revenues.

encouraging investment and business activity.[10] Dethier et al. (2010) discuss how an improved investment climate increases returns to existing investments and subsequently, these higher returns attract future investments. In their extensive review of the existing literature, Dethier et al. (2010) discuss a number of papers which have detected a positive relationship between infrastructure investment, particularly reliable power supply, and economic performance, using both country and firm-level data. Specifically, using panel data across 88 countries over 40 years, Calderon et al. (2011) estimate the output elasticity of infrastructure capital to be 0.1, distinguishing in the process infrastructure capital from other forms of physical capital. The paper also finds little heterogeneity in the output elasticity of infrastructure capital across countries on the basis of countries' per capita incomes or initial endowments of infrastructure capital. These results using country-level data are consistent with those obtained using firm-level data. Dollar et al. (2005) show using firm-level surveys across India, Pakistan, Bangladesh and China that the reliable supply of electricity has a significant impact on firm output, capital formation and profits.

In light of the robust findings of these papers and others discussed by Deither et al. (2010), we examine whether any reallocation of public expenditures by resource constrained AA party legislators from non-exclusive to targeted public goods impact regional inflows of private investment. Using capital formation in the manufacturing sector as a measure for private investment, we empirically test for differential impacts of the marginal AA party legislator across the legislative strength of AA parties and industries' reliance on physical infrastructure to link the reduced form impact of AA party legislators on manufacturing investment with legislators' redistributive spending preferences.

1.3 Background on India's Electoral System and AA Parties

1.3.1 Background on Elections to State Legislative Assemblies

There are three tiers of government in India - federal, state and local, all of which are democratically elected in a model akin to the parliamentary system. As this paper examines the economic impacts of state legislators, we focus exclusively on state-level elections. Each state legislature has a number of electoral constituencies, proportional to the state's population. There remains considerable variation across states with regard to the number of electoral constituencies with the largest (and most populous) state in the sample (Uttar Pradesh) having a total of 403 constituencies, and the smallest state (Uttarakhand) having 70 constituencies. Elections are conducted by the Election Commission of India which is a constitutional body, unaffiliated with any political party and independent of the incumbent government, either at the state or the federal level. The results of each election is based on the first-past-the-post principle with the party securing the maximum number of votes

[10] See for instance Aschauer (1989); Aterido et al. (2007); Bastos and Nasir (2004); Calderon et al. (2011); Dollar et al. (2005); Escribano and Guasch (2005); Escribano et al. (2010).

within an electoral constituency being declared the winner.[11]

The candidate associated with the winning party is subsequently elected to the state legislature and is responsible for representing that constituency for the term of the legislature (typically 5 years). During matters of voting on legislations, each legislator has a single vote, regardless of the constituency from which they are elected. With regard to government formation, the party which wins at least half the electoral constituencies in a state is invited to form the state government.[12]

1.3.2 Brief Overview of AA Parties

The political mobilization of historically underprivileged low caste groups started primarily in the mid-1980s, under the auspices of two political parties, namely the Janata Dal (JD) and the Bahujan Samaj Party (BSP). The latter catered exclusively to the Scheduled Castes (SC) or *Dalits* while the former mainly represented a broader coalition of underprivileged groups the Other Backward Castes (OBC) - drawn mostly from the lowest strata of the fourfold caste hierarchy. While the two parties appealed to two distinct communities, there was an inherent convergence in their political goals: namely to capture political power through electoral politics and subsequently, increase the representation of low caste individuals in public institutions, and engage in the targeted redistribution of public resources to low caste groups. There were two key features common to both these parties which distinguished them from other mainstream parties: first, the majority of the leadership of these parties were comprised of individuals hailing from the lower castes; and second, a high proportion of electoral candidates from these parties hailed from low caste backgrounds.[13]

The nascent political party, JD, firmly established itself in the Indian political scenario in the aftermath of the federal elections of 1989, when it was able to lead a coalition government at the federal level. Even though the government itself was short-lived and remained in power for ten months, it was successful in enacting legislation which set aside 27 percent of all federal public sector positions for the OBC community [14]. Apart from its short stint at the federal level, the JD also managed to secure electoral majorities and form governments in multiple states through the 1990s, particularly in the populous north Indian states of Uttar Pradesh and Bihar. During this period though, the JD also splintered into multiple constituents across states, giving rise to powerful regional parties which wielded considerable clout at the state level. Between 1995 and 2012, all the major regional off-shoots of the erstwhile

[11] The pluralist framework of Indian democracy allows for multiple political parties contesting the election in each electoral constituency. Individuals unaffiliated to any political party may also contest the elections as independent candidates.

[12] In the event that no single party has won at least 50 percent of the electoral constituencies, a coalition government can be formed with multiple parties.

[13] For instance, the JD committed that 60 percent of its candidates to the federal elections would be from low caste background (Jaffrelot, 2003).

[14] This was on the basis of the recommendations of the Mandal Commission. The Mandal Commission was established in 1977 by the federal government to determine caste groups eligible for affirmative action, and also quantify the level of affirmative action required to adequately represent lower castes in public institutions. Although the commission submitted its recommendations in 1980, the subsequent federal governments led by the Congress Party refused to act upon the proposed recommendations.

JD managed to enjoy at least one full five year stint in power across four major states - one located in southern India (Karnataka), one in eastern India (Orissa) and the remaining two in north India (Uttar Pradesh and Bihar). Similarly, the BSP, after two short-lived attempts at governance, successfully obtained a majority in 2007 and ruled Uttar Pradesh, Indias most populous state with a sixth of the country's population, till 2012.

Aside from winning the requisite number of elections to form state governments, the JD and its regional constituents have maintained over the past two decades a considerable electoral presence in at least five major states. For instance, in the 2000-2010 period, the two major fragments of the JD[15] in Bihar have consistently polled between 35 and 45 percent of the popular vote. Likewise, in Uttar Pradesh, the BSP and the principal JD offshoot[16] received 30 percent of the votes in 1993, and their combined vote share steadily increased to over 50 percent by 2002 and have stayed at that level for the next decade. The stability in vote shares of the AA parties underline the presence of a core support base which can provide them with the necessary bargaining power in the state legislature to influence public policy, even when they are part of the opposition benches.

Along with the regional off-shoots of the JD, we also include the Left parties and the DMK based in the southern state of Tamil Nadu within the ambit of AA parties. These parties have been in existence for longer than either the JD or the BSP and have achieved electoral successes since 1967 in the states of Tamil Nadu, Kerala and West Bengal. While the Left have typically preferred to frame policy debates around class as opposed to caste, the strong correlation between caste and income results in lower castes forming a major support base for the Left parties. The DMK on the other hand was formed after a split in the Justice Party in 1946, which was essentially a 'rationalist' anti-upper caste party, formed in the 1920s to challenge the hegemony of the upper castes in the southern state of Tamil Nadu. Both the Left and the DMK[17] have consistently formed electoral alliances with the BSP and the JD off-shoots, and supported them on legislative issues within state and federal legislatures.[18] When in power, both the Left parties and the DMK have also implemented policies targeted to benefit low caste populations. The DMK for instance championed affirmative action policies in Tamil Nadu, leading to a large percentage of public sector jobs being reserved for low caste citizens; the Left parties have successfully engaged in far reaching land reforms comprising of titling sharecroppers and providing land ownership to landless labourers, both of whom were drawn mostly from low caste populations. However, as these parties are not strictly formed along caste identities, we undertake a robustness check in Section 7 to ascertain that our core results are not sensitive to this broader classification of AA parties.

[15] These are the Rashtriya Janata Dal (RJD) and the Janata Dal(United) - JD(U) respectively.

[16] This is the Samajwadi Party (SP).

[17] The DMK as a party split in 1972 leading to the formation of the All India Anna Dravida Munnetra Kazhagam (AIADMK). In this paper, we do not classify the AIADMK as an AA party as the core leadership of this party has been dominated by upper castes and the party had a distinctly toned down 'anti-upper caste' rhetoric after splitting from its parent organization.

[18] The Left parties and the DMK for instance supported the National Front federal government led by the JD in 1990; they also were part of the federal United Front government which was a coalition government formed mainly by the splinter groups associated with the JD. In Bihar and Uttar Pradesh, the Left parties have typically allied with the JD and its splinter groups - namely the RJD and the SP

1.4 Empirical Strategy and Data

1.4.1 Defining the Source of Exogenous Variation for AA Party Success

The primary goal of this paper is to identify the impact of change in *de facto* political power in favour of low caste citizens due to the electoral success of AA parties, on regional and household economic outcomes, measured at the state (district) level. To this effect, we aim to estimate an equation of the form:

$$Y_{rt} = \alpha_r + \delta_t + \theta_{st} + \beta PCAA_{rt} + \gamma X_{rt} + \epsilon_{rt} \tag{1.2}$$

In (1), Y represents the outcome of interest, measured in some geographical region r (state or district) and time period t. Time-invariant region-level determinants of the outcome are captured through the region fixed-effect α and shocks affecting all regions in time period t are controlled through fixed effect δ. For district-level data, we also include a state-time fixed effect denoted by θ. $PCAA$ denotes the share of elections won by AA parties in the region during the time period of interest while X includes a set of time-varying region specific controls.

Specification (1) estimated using OLS would provide biased estimates of β due to unobservables correlated with both the electoral success of AA parties in a region, and the relevant outcomes of interest, with the direction of the bias being ambiguous.[19] To overcome this endogeneity concern, our paper constructs a region-specific AA party representation shock for each electoral period, based on the outcome of close elections between AA and non-AA parties. Critically, our measure of the AA party representation shock hinges on the assumption that the outcome of close elections between AA and non-AA parties can be deemed as 'good as random'.

The challenge in constructing a region-specific electoral shock is compounded in our case due to a mismatch in the levels at which the elections are conducted, and the levels at which economic outcomes are measured. Unlike most other regression discontinuity designs (RDD), the constituency level electoral outcomes do not map directly into the administrative unit at which the economic outcomes are measured (state or district). Thereby, we are unable to use a sharp RDD which have characterized the majority of studies in this literature. As described in Section 1.3.1, there are multiple electoral constituencies located within a region, necessitating an aggregation of constituency level results to the appropriate region. Based on the aggregation of electoral results at the constituency level, the exogenous variation in the number of AA party legislators for any arbitrary region r in time period t is expressed as:

[19] For instance, we would expect the support for AA parties to be higher in regions with a higher proportion of low castes voters. However, a higher proportion of low caste voters would also mechanically increase their representation in public institutions, thereby imposing an upward bias in our estimated impact of AA party success on economic outcomes. Similarly, states such as Bihar and Uttar Pradesh, where AA parties have been successful have historically been states with low rates of economic growth, human capital formation and a poor investment climate. This can bias the impact of AA party success on private investment inflows downwards, overstating the negative impact AA parties might have on state-level investment inflows.

$$RAA_{rt} = \sum_{c=1}^{N} \mathbb{I}(AACW_{rct}) - 0.5 * \sum_{c=1}^{N} \mathbb{I}(AACE_{rct}) \tag{1.3}$$

In (2), N denotes the total number of electoral constituencies located in region r. $AACW$ is an indicator variable equaling 1 if an AA party won a close election against a non-AA party in constituency c, while $AACE$ is an indicator variable equaling 1 if there was a close election involving an AA party and a non-AA party in constituency c. Conditional on the outcome of close elections being drawn from a stochastic process and each contestant in a close election having an equal chance of winning, we scale the second term in Equation (2) by 0.5. The first term in Equation (2) captures the actual number of AA party legislators elected in the region during a given electoral cycle while the second term denotes the expected number of AA party legislators who should have been elected on the basis of the number of close elections contested by AA parties against non-AA parties. The difference provides us with the 'unexpected' variation in the number of AA party legislators elected from the region in the given electoral period. By construction therefore, we would expect this measure to be centred around 0.

It is worth noting that our measure of exogenous AA party representation does not take into account two types of electoral contests - those involving two AA parties in a close race, and those involving two non-AA parties in a close race. Conceptually, this makes the construction of RAA similar to an instrumental variables framework involving 'compliers', 'non-compliers', and 'switchers'. In the event of a close election involving two AA parties, regardless of which AA party wins the race, we would have an AA party legislator representing the constituency and the situation is akin to that of 'compliers'. On the contrary, with two non-AA parties contesting a close election, there would be a non-AA party legislator representing the constituency irrespective of the results of the election and this is similar to the case of 'non-compliers'. In contrast, when an AA party contests against a non-AA party in a close election, with the AA (non-AA) party emerging victorious, there is a quasi-random switch from a non-AA (AA), to an AA (non-AA) party legislator representing the constituency and the paper relies solely on such elections to identify the impact of AA party legislators on economic outcomes.

Subsequently, the core empirical specification becomes:

$$Y_{rt} = \alpha_r + \delta_t + \theta_{st} + \beta RAA_{rt} + \gamma X_{rt} + \epsilon_{rt} \tag{1.4}$$

Conditional on the total number of close elections contested by AA parties, β in Equation (1.4) identifies the causal effect of a 'positive AA representation shock', occuring due to an additional close election won by an AA party against a non-AA party. As the total number of close contests occuring in a region is likely to be endogenous to regional economic outcomes, it is included as a covariate in X in each specification.

Constitutionally, elections to state legislative assemblies are held every 5 years and the timing of elections is state-specific, with approximately 6 states facing elections in any given year. In this respect, RAA is invariant for a given state (district) within each electoral cycle, and the main empirical specifications control for the number of years which has lapsed since the previous election in the states.

Finally, for the purposes of this paper, an election is determined to be 'close' if the difference in the margin of victory between an AA and a non-AA party is less than 5 percent of the total votes cast in the election. We choose the 5 percent threshold as it is the smallest threshold which provides us in expectation with at least 1 close election involving AA parties for our sample districts. To verify that our results are not sensitive to this choice, we undertake in Section 1.7.1 robustness checks where we reduce the threshold to 4 and 2 percent and demonstrate that our core results do not change substantially upon alterations of the 5 percent cutoff for close elections.

1.4.2 Data

This paper uses data from four sources: the electoral data is obtained from the Election Commission of India. The data on individual employment outcomes is obtained from the National Sample Survey's household surveys. The data for state expenditures is obtained from the Reserve Bank of India (RBI) while the data on state-level private investment is obtained from the Annual Survey of Industries' (ASI) state-level data on manufacturing plants. Additional covariates are sourced from the decennial Census of India and the Handbook of Statistics on the Indian Economy hosted by the Reserve Bank of India.

The electoral data provided by the Election Commission of India contains information on every election to state legislative assemblies across all states. We use a rich sample covering over 22,000 elections between 1987 and 2011 across 19 major Indian states. Out of these 22,000 elections, nearly three-fourths, or about 16,000 elections, involved at least one AA party. For the elections involving at least one AA party, 2,450 - or about 15 percent - are deemed to be close elections at the 5 percent margin, involving one AA and one non-AA party.[20] Out of these 2,450 close elections involving an AA and a non-AA party at the 5 percent margin, AA parties won 1,214 close elections - or about 49.5 percent - and lost 1,236 close elections, providing preliminary support to the contention that the outcome of close elections between AA and non-AA parties can be considered to be equivalent to a coin-flip. Tables 1 and 4 of Section 11.1 (Appendix) present the summary statistics for the electoral variables at the state and district levels. On average, there are 200 elections per state during any electoral cycle, with 25 elections deemed close at the 5 percent margin involving an AA and a non-AA party (Section 11.1, Table 1). Within a district, there are on average 12 elections to the state legislative assembly out of which 1 is a close election at the 5 percent margin between an AA and a non-AA party (Section 11.1, Table 4). Importantly, at both the state and the district-level, the quasi-random variation in regional AA party representation is not significantly different from 0.

To identify the spending preferences of AA party legislators, we use data from the Reserve Bank of India's annual publication titled 'State Finances: A Study of Budgets'. This contains annual state government revenues and expenditures for major spending categories across all states and includes both revenue and capital expenditures. For the purposes of this paper, we combine for each category revenue and capital expenditures to obtain total

[20] An additional 389 'close' elections within the 5 percent bandwidth were contested between two AA parties while two non-AA parties contested an additional 2,160 elections within the 5 percent margin. Table 11.2 in the Appendix presents some descriptive statistics comparing these three types of constituencies.

expenditures for that category in each year. As the expenditure amounts are measured in current rupees, we convert all annual values to 2012 rupees using an imputed inflation index.[21] The data on public expenditures is available from 1990 and as the rise of AA parties began in the late 1980s, we are able to identify the impact of AA party legislators on public expenditures over a two decade interval between 1990 and 2010, across 16 (19[22]) major states covering over 90 percent of the nation's population. The summary statistics for per capita expenditures and their respective shares in total expenditures for six major expenditure categories are presented in Table 2 of Section 11.1 (Appendix).

To identify the impact of AA party legislators on district-level employment, we rely on employment data from the nationally representative household sample surveys conducted by the National Sample Survey Organization (NSSO). The NSSO engages in extensive surveys (subsequently referred to as 'NSS') once every five years to gauge employment and consumption behaviour amongst individuals and households, covering over 1,00,000 households across all states and districts in India. In addition to identifying detailed sectors of employment, the NSS also records the educational qualifications of respondents and a host of other demographic factors such as geographic location, age and household size. Since 1999 (round 55), the NSS began recording the type of institution in which respondents were employed - public or private - from which we construct our measure of public employment. Moreover, since 1999, the NSS also started classifying the OBCs as a separate category whereas previously, they had only identified SCs and STs amongst the low castes. This permits us to construct district-level estimates of public sector employment across 4 survey rounds - 55, 61, 66 and 68[23] - and at least three electoral cycles for every district. The aggregation of individual responses to the district is warranted as our source of exogenous variation is at the district-level and we weigh individual responses using the NSS assigned household weights [24]. As district boundaries have changed in this period a number of times, we aggregate our data based on 1966 district boundaries which provide us with a consistent panel of 271 districts across 19 states and 4 NSS survey rounds. Based on the NSS data and consistent with the findings of Kijima (2006), Table 5 in Section 11.1 (Appendix) documents the significant socio-economic deprivation faced by individuals from low-caste backgrounds. Low caste individuals have a higher propensity to reside in rural areas, face twice the level of poverty relative to high caste individuals, and have signficantly lower levels of educational attainment, even in urban areas. They are also much less likely to be employed in public sector positions in comparison to high caste individuals.

To determine if the electoral success of AA parties affect state-level private investment, we use data from the Annual Survey of Industries (ASI). The ASI presents key industrial statistics for all registered[25] manufacturing plants across India. The sampling unit for the ASI

[21] The lack of a readily available annual inflation index for each state forces us to impute the inflation rate from the growth of current and constant net state domestic product (NSDP). Specifically, we calculate our estimated rate of inflation as the difference between NSDP growth measured in current rupees and constant rupees. We use this imputed inflation index to convert all the remaining monetary measures to 2012 rupees.

[22] In 2000, three states, namely Uttar Pradesh, Madhya Pradesh and Bihar, were bifurcated to form three new states: Uttarakhand, Chhattisgarh and Jharkhand.

[23] The corresponding years are 1999-00, 2004-05, 2009-10 and 2011-12.

[24] Household weights are the inverse of the sampling probability for each household

[25] Units which employ over 20 workers (10 if using electricity) have to register themselves with the state

is the 'factory' - equivalent to a manufacturing unit - and the variables covered include fixed capital, gross capital formation, value of inputs and output, net value added, aggregate fuel consumption, number of workers employed and wages paid to workers. The ASI's annual publications aggregates the factory-level data across all industries in the state and also provides state-level estimates disaggregated at the 2-digit industry level[26]. We extract this data to construct state-level estimates of fixed capital and gross capital formation for each 2-digit industry which forms our two key measures of firm investment. The annual monetary values are converted to 2012 values using the state-specific imputed inflation index discussed above. The summary statistics for the key outcome variables are presented in Table 3 of Section 11.1 (Appendix).

1.5 Empirically Validating that the Outcome of Close Elections is Exogenous

1.5.1 Constituency Level Covariate Balance

Prior to presenting the reduced form results, we empirically validate our claim that the outcome of close elections involving AA parties at the 5 percent margin is indeed 'as good as random', and the region-level AA party representation shock is exogenous to both state and district-level observables. As electoral outcomes are determined at the constituency level, we first demonstrate graphically that the outcome of close elections are uncorrelated across constituency level observables, in an approach similar to Meyersson (2014).

At the constituency level, the running variable of interest is the victory margin of AA parties. In the simplest case where an election has two candidates, one belonging to an AA party and the other to a non-AA party, the victory margin is the difference in vote share between the AA and the non-AA party with a positive (negative) victory margin signifying an AA party victory (defeat). With multiple candidates from AA and non-AA parties, we define AA victory margin - $AAVM$ - as the following: in the event of an AA party victory in constituency c and election year y;

$$AAVM_{cy} = WAAVS_{cy} - \max\left(NAAVScy\right) \tag{1.5}$$

In Equation 1.5, $AAVM$ denotes the AA party victory margin in constituency c and election year y. The vote share of the winning AA party is represented by $WAAVS$ and we subtract from it the maximum vote share received amongst all non-AA parties contesting the election. Analogously, in the event that a non-AA party wins the election, we define $AAVM$ as:

$$AAVM_{cy} = WNAAVS_{cy} - \max\left(AAVScy\right) \tag{1.6}$$

where $WNAAVS$ is the vote share received by the winning non-AA party and we subtract from it the maximum vote share received amongst all AA parties which contest the election

government under the provisions of the Factory Act of 1947.

[26] This is based on the industry codes provided by the National Industrial Classification.

in constituency c, year y.

Based on this formulation for AA party victory margin, we present a graphical representation in Figure 1.1 of the McCrary test which tests for a discontinuity in the running variable - $AAVM$ - at the cutoff point 0, where the outcome of a constituency level election changes discontinuously from AA party defeat to AA party victory. We present results from a total of 16,277 constituency level elections, held across 19 states in the 1987-2011 period. The horizontal axis is divided into 500 bins between AA party victory margins of -0.3 and 0.3 with the red vertical line at 0 representing the cutoff point. The grey shaded circles represent the number of observations in each bin corresponding to the respective AA party victory margin; the red and orange lines display the fitted values from a local second order polynomial regression with the black lines representing 95 percent confidence intervals. Visually, the plot shows no evidence of any discontinuity at the cutoff and the 95 percent confidence intervals also overlap, validating that there is no selective sorting of electoral outcomes at the cutoff point.

Next, we perform covariate balance checks across the support of AA party victory margin, similar to Meyersson (2014). At the constituency level, there are 8 covariates - namely AA party vote share; number of registered voters (natural log); voter turnout; share of constituencies reserved for low caste SC/ST candidates; number of candidates; share of male winners; share of winners from SC/ST communities; and age of the winner. For each covariate, we calculate the unconditional mean of the covariate in each of the 500 bins of AA victory margin between -0.3 and 0.3. Figure 1.2 presents the graphical plots from this excercise. Akin to Figure 1.1, AA party victory margin is plotted on the horizontal axis as the running variable with the red vertical line representing the cutoff point at 0. The grey shaded circles are the unconditional means of the covariates in each bin of AA party victory margin; the coloured lines are the fitted values from a second order local polynomial regression estimated on each side of the cutoff while the black lines denote the 95 percent confidence intervals. For each of the 8 covariates displayed in the figure, we cannot visually identify any break or discontinuity at the cutoff point of 0 and the confidence intervals also overlap in every plot, confirming covariate balance across the cutoff for AA party victory margin.

1.5.2 District and State-Level Covariate Balance

As the paper's identification strategy hinges on the number of 'unexpected' legisators elected from AA parties in a state (district), we first present a graphical comparison between expected close wins and actual close wins for AA parties in the state (district) at the 5 percent threshold for close elections. As laid out in Equation (2), if each contestant has an equal chance of winning a close election, we would expect AA parties to win half the number of close elections they contest in a state (district). The state-level results from this exercise are presented in Figure 1.3 and the district-level results are presented in Figure 1.4. In both the figures, the horizontal axis plots the expected number of AA party victories, based on the number of close contests between AA and non-AA parties in the state (district). The vertical axis records the actual number of AA party victories in close contestss and we restrict the sample in Figure 1.4 to districts with at least 1 close election between an AA and a non-AA party. The green line is the 45 degree line where expected number of close wins equal actual

close wins while the red dashed line plots the linear relationship between expected and actual victories through close elections. For both figures, the linear trend line is very close to the 45 degree line and the majority of the points are evenly spread around the 45 degree line. This provides visual evidence that AA parties have an equal likelihood of winning close elections contested against a non-AA party at the 5 percent margin.

Formally, we validate the exogeneity of the quasi-random variation in regional AA party representation by individually regressing our independent variable of interest, RAA, on state (district) covariates. At the district-level, we perform our test on 8 political covariates sourced from data provided by the Election Commission of India, and 8 socio-economic covariates obtained from the household surveys conducted by the NSS. [27] At the state-level, the political covariates remain unchanged and we obtain from the decennial Census of India, and the Reserve Bank of India's Handbook of Statistics of the Indian Economy 8 socio-economic covariates.[28] In each specification, we control for the total number of close elections contested by AA parties as well as region (state or district) and time fixed effects, akin to the main specifications, and the standard errors are clustered at the level of district (state-electoral cycle).The state results are presented in Tables 3 and 4 while the district results are presented in Tables 1.3 and 1.4. At neither the state, nor the district level do we find any of the covariates to be significantly correlated with our measure of quasi-random variation in regional AA party representation.

1.6 Results

This section contains the key findings of the paper. We first document the redistributive aspect of AA parties by showing the causal impact of AA party legislators on public expenditures and public employment for low caste workers. Subsequently, we present the reduced form results documenting the causal impact of AA parties on manufacturing investment and finally, present empirical evidence consistent with the claim that a potential channel through which the electoral success of AA parties affect private investment is investors' response to the expenditure choices made by AA party legislators.

[27] The political covariates include total elections in a district; the share of elections reserved for minority SC/ST candidates; the number of contestants in each election; support for AA parties at the district-level; district-level voter turnout; the total number of close elections involving AA parties in the district; and the share of elections won by AA parties in the previous electoral cycle which provides us with a measure of historical support to AA parties in the district. The socio-economic covariates at the district level include the district's population; share of rural population in the district; share of low caste SC/ST/OBC population in the district; average household size; the share of households below the poverty line as measured through the headcount ratio; share of adults with at least secondary education and the labour force participation rate amongst working age adults.

[28] The state-level socio-economic covariates are: logged state population; share of low caste SC/ST population in the state; share of literates; share of workers; share of urban households; state gender ratio; state population density; and logged net state domestic product.

1.6.1 AA Party Success and State Expenditures

This section presents results identifying the effect of a change in *de facto* political power for low caste citizens on public spending. In this regard, we identify the impact of a quasi-random increase in AA party representation on the share of state expenditures allocated across six expenditure heads: namely targeted low caste welfare; untargeted low caste welfare in the form of social security transfers and rural development[29]; road construction and investments in power generation; health and education spending; agriculture and irrigation expenditures; and pensions and administrative expenses. Collectively, these six categories account for almost 85 percent of state government expenditures across our sample.

The first category measures public expenditures exclusively targeted towards low caste citizens whereby recipients must belong to a low-caste background to be a beneficiary of any spending under this category.[30] Expenditures under this category typically take the form of scholarships to low caste students, construction of housing for low caste populations and also the targeted provision of public goods exclusively to low caste households. This is also the key outcome variable selected by Pande (2003) to study the direct impact of mandated representation of low caste politicians on low caste citizens. While this remains the most direct measure of state benefits accruing to low caste households, we use expenditures on rural development and social welfare programs as a measure of untargeted welfare spending for low caste citizens. This is motivated by the fact that a disproportionately high share of low caste households are located in rural areas and almost one out of every three low caste households fall below the official poverty line. Low caste households therefore would have a larger propensity to benefit from overall rural development and social protection schemes such as expansions in the public distribution system providing subsidized food grains to poor citizens.

If AA party legislators implement their party agenda to transfer state resources to low caste citizens, we would expect their impact to be channelized along the two above discussed expenditure categories. In addition, to test whether AA party legislators transfer resources to low caste citizens through an expansion in public expenditures or redistributive spending, we identify the causal impact of AA party legislators on health and education; road and capital investments in power generation; agriculture and irrigation; and administrative salaries and pensions. The empirical specification used to identify the impact of AA party legislators on state expenditures is:

$$Y_{sy} = \alpha_s + \delta_y + \beta RAA_{sy} + \gamma X_{sy} + \epsilon_{sy} \tag{1.7}$$

The dependent variable measures the share of state expenditures allocated to each expenditure category. The unit of observation in Equation (1.7) is state-year with s representing the state and y the year. To facilitate the interpretation of the coefficients, each expenditure share is multiplied by 100. β in Equation (1.7) measures the causal impact of a positive AA party representation shock on the percentage point increase (decrease) in a given expenditure category. As the unit of observation is the state, the source of identification is state-level AA

[29] This is inclusive of grants made to rural local government institutions (Panchayati Raj Institutions) and expenditures under the category of nutrition

[30] This includes OBCs, *Dalits* and *Adivasis*.

party representation shocks, necessitating an aggregation of RAA to the state-level. Time-varying state-level covariates are contained in X[31] while α and δ denote state and year fixed effects. The standard errors are clustered at the level of state-electoral cycle. The selection of this level of clustering as opposed to the more traditional state level is motivated due to the relatively small number of states (19) in India which creates a concern about having too few clusters (Clots-Figureas, 2011). We verify in Section 1.7.4 that the results are not sensitive to this choice of clustering.

Reduced Form Results: AA Party Representation Shock and State Expenditures

The reduced form results identifying the impact of AA party legislators on state government expenditures is shown in Table 1.5. Panel A reports the results with covariates while Panel B excludes all covariates with the exception of total close elections contested by AA parties and state and year fixed effects. The inclusion of covariates do not influence the coefficients in Panel A and the coefficients in Panel A and Panel B are statistically indistinguishable from one another. In this regard, all subsequent state-level expenditure results include covariates.

Columns (1) and (2) of Panel A shows that the marginal AA party legislator has a positive and statistically significant impact on the share of targeted and untargeted welfare expenditures towards low caste citizens. At the mean expenditure shares, the coefficients imply that the marginal AA party legislator increases the share of state expenditure allocated towards targeted low caste welfare schemes by 2 percent and untargeted low caste welfare schemes by a little over 1 percent.[32].

To have a better understanding of the relative magnitude of the marginal legislator, we convert the percentage amounts to monetary terms and benchmark it against the per capita monthly consumption of low caste households. In 2012 values, the average per capita expenditure for targeted and untargeted low caste welfare was 162 rupees and 668 rupees respectively. Conditional on AA parties having no impact on aggregate public spending (verified in Table 1.6), a 2 and 1.3 percent increase in each expenditure category amounts to a combined per capita increase of 12 rupees in state expenditures allocated towards low caste welfare. The NSS' household surveys reports the average urban low caste household's monthly per capita expenditure in 2011-12 to be 2,000 rupees. Thus, the monetary impact of the marginal AA party legislator's spending choices can be interpreted to be equivalent to a 0.6 percent increase in the monthly per capita consumption of an urban low caste household.[33].

Column (3) of Table 1.5 shows that the marginal AA party legislator has a significant negative impact on the share of state expenditures allocated towards physical infrastructure - namely road construction and investment in electricity generation. At the mean of the dependent variable, the coefficient in column (3) amounts to a 3 percent reduction in the

[31] This includes both political and economic covariates such as the total number of close elections contested at the state-level by AA parties; state-level political competition; state-level voter turnout; the share of constituencies reserved for minority candidates in the state; constant net state domestic product; and the state-level share of literates, share of workers and rate of urbanization.

[32] 0.0617/3.104 in column (1) and 0.1499/13.399 respectively.

[33] If the benchmark is the per capita monthly consumption of rural low caste households, the corresponding increase would be 1.2 percent

share of state expenditures allocated to physical infrastructure.[34] In monetary terms, the average per capita expenditure on roads and power investments in 2012 rupees was 441 rupees, with a 3 percent decline amounting to a net reduction of 14 rupees per capita (if aggregate state expenditures remain unchanged) - almost equivalent to the 12 rupees per capita increase in low caste welfare spending. This result suggests that the marginal AA party legislator is redistributing state resources from non-exclusive public goods such as roads and electricity to targetable public goods such as transfer payments or social protection for select beneficiaries.

The last three columns (columns 4-6) of Table 1.5 inform us that AA party legislators have no significant impact on health and education or agriculture and irrigation spending but increase the share of state expenditures allocated towards pensions and administrative salaries by 1 percent. This would be consistent with the positive impact of AA party legislators on public hiring of low caste workers (shown in Section 6.2) if the hiring was concentrated in administrative positions. As one primary research question of this paper is to identify the redistributive spending preferences of AA party legislators, we restrict our attention to the share of targeted and untargeted low caste transfers and public spending on physical infrastructures for the remainder of the paper.

Finally, we verify in Table 1.6 that AA party legislators have no significant impact on state expenditures, revenues or deficits. This confirms that the increase in welfare spending undertaken by AA party legislators is neither financed through higher revenues, nor deficits. This is consistent with the institutional structure of Indian states whereby states have limited fiscal powers and are constrained in their ability to borrow from private markets.

When Does the Marginal Legislator Matter? Differential Impact of AA Party Legislators by Legislative Strength of AA Party

The results discussed in Section 1.6.1 identifies the average effect of the marginal AA party legislator on various state expenditure categories. However, given that the median state legislature comprises of 200 legislators, it warrants the question as to how influential an additional representative would be in such a large pool of individuals. To answer this question, we identify heterogeneous impacts of the marginal AA party legislator by the legislative strength of AA parties in the state legislature. As the average state in our sample witnesses 30 percent of elections being won by AA parties, we identify the differential impact of the marginal AA party legislator when AA parties have won more than 30 percent of the elections to the state legislature. Specifically, we estimate the following equation:

$$Y_{sy} = \alpha_s + \delta_y + \beta_1 RAA_{sy} + \beta_2 \mathbb{I}(ShAAWin > 0.3)_{sy} * RAA_{sy} + \beta_3 \mathbb{I}(ShAAWin > 0.3)_{sy} + \gamma X_{sy} + \epsilon_{sy} \quad (1.8)$$

In Equation (1.8), $\mathbb{I}(ShAAWin > 0.3)$ is a categorical variable equaling 1 if AA parties have won more than 30 percent of the elections in the state for the given electoral cycle. The results in Table 1.7 show that marginal AA party legislator is influential only when

[34] Evaluated as -0.2708/8.422

more than 30 percent of the legislators in the state legislature hail from AA parties. In each instance, β_1 is statistically insignificant and even though the interaction term also fails to attain statisitcal significance, the sum of β_1 and β_2 in each case is significant at the 5 percent margin, signifying that the marginal AA party legislator has a causal impact on state expenditures only if AA parties control at least 30 percent of the state legislature. This result is consistent with the functioning of legislatures which have majority voting rules, as well as Acemoglu and Robinson (2008) and Acemoglu et al.'s (2013) prediction that a change in the distribution of political power is necessary for higher levels of redistribution towards disadvantaged populations. As observed in Table 1.7, an additional electoral victory for an AA party affects public policy only when AA parties have also won a significant share of elections in the state and have the necessary legislative strength to support the policy preferences of individual legislators.

Mandated Representation or Party Agenda? Comparing the Impact of Low Caste Politicians from AA Parties and Mainstream Parties

The reduced form results in Section 1.6.1 provides causal evidence that changes *de facto* political power for disadvantaged citizens can enhance targeted redistribution. In this section, we exploit the system of electoral reservations to compare the respective impacts of changes to *de jure* and *de facto* political power for low caste populations on targeted welfare expenditures for low caste citizens. This comparison also permits us to determine the key factor driving public resource allocation for disadvantaged populations: the socio-ethnic identity of individual legislators or a broader policy preference amongst political parties for targeted welfare spending.

Constitutionally, since independence, every state has a certain number of 'reserved' constituencies from which only candidates belonging to low caste backgrounds can contest elections.[35] The specific number of reserved constituencies are proportional to the fraction of low caste citizens in the state with the average share of reserved constituencies equaling 23 percent across all major states. As only low caste candidates can contest an election from a reserved constituency, mechanically, it ensures that the elected representative from a reserved constituency will be a low caste legislator, irrespective of party affiliation. We exploit this feature to compare the impact of low caste legislators on targeted welfare transfers for low caste citizens across AA parties and two mainstream national parties - the right-wing BJP and the centrist Congress party.[36] If electoral reservations - guaranteeing the representation of a fixed share of low caste politicians - is the key driver of redistributive spending, should

[35] Low caste herein refers to the Scheduled Castes and the Scheduled Tribes only and does not include the Other Backward Castes

[36] Regarding the choice of parties for the comparison, the BJP and the Congress are two of India's largest and richest political parties, functioning across the country. The BJP, dominated by upper caste leaders and drawing its support mostly from citizens belonging to the upper and middle castes lie on the opposite spectrum to AA parties. The Congress on the other hand is India's centrist 'catch-all' party which prior to the AA parties was the premier party for low caste populations, making it the primary competitor for AA parties in the 1980s. However, despite its broad appeal to the citizenry, the majority of the Congress leadership and legislators hails from the upper castes, which drove the opposition of AA parties to the Congress (Jaffrelot, 2003).

expect to see no difference in the impact of legislators elected from reserved constituencies on targeted transfers to low caste citizens across the three parties.

The results are presented in Table 1.8. The empirical specification is identical to Equation (1.7). The only variance is that the AA party representation shock is now estimated seperately across close elections occuring in reserved and non-reserved constituences between AA and non-AA parties. For each party, the first column identifies the causal impact of the marginal legislator elected from reserved constituencies and the second column identifies the impact of the marginal legislator elected from non-reserved constituencies. The dependent variable in each instance is the share of state expenditures allocated exclusively to low caste welfare. From Table 1.8, we see that regardless of the type of constituency from which they are elected, the marginal AA party legislator (columns 1 and 2) has a positive and significant impact on the share of state expenditures allocated to targeted low caste welfare. In contrast, neither the marginal BJP legislator (columns 3 and 4), nor the marginal Congress legislator (columns 5 and 6) has any significant impact on the share of state expenditures targeted towards low caste citizens.

In terms of relative magnitudes, we are unable to reject the equality of the coefficients for the marginal Congress legislator and the marginal AA party legislator across reserved and non-reserved constituencies. However, the coefficient for the marginal Congress legislator from reserved constituencies (column (5)) is less than half the impact of the marginal AA party legislator from reserved constituencies (column (1)) and the former is also not statistically significant. When comparing the marginal BJP legislator against the marginal AA party legislator, we are able to reject the equality of the coefficients across both reserved and non-reserved constituencies. Moreoever, we are also able to reject the hypothesis that the marginal BJP legislator elected from a reserved constituency has the same impact on the share of targeted spending for low caste citizens as the marginal AA party legislator elected from a non-reserved constituency (comparison of column (2) and column (3)).

The results in Table 1.8 provide two key insights. First, it shows that redistributive spending for disadvantaged populations is impacted only when changes in *de facto* political power complement changes to *de jure* political power. This is consistent with Acemoglu and Robinson (2008) and Acemoglu et al.'s hypothesis (2013). The results also inform that mandated political representation for disadvantaged citizens would increase redistribution only when accompanied by political parties with a strong policy preference for redistribution towards disadvantaged populations. In the absence of such a party mandate, elite capture of political parties, combined with majority voting rules, can subdue the impact of mandated political representation of minority politicians on welfare expenditures.

1.6.2 AA Party Representation Shock and Public Employment for Low Caste Workers: Reduced Form Estimates

Section 1.6.1 establishes that changes to *de facto* political power favouring disadvantaged citizens results in targeted redistribution. This is through the most direct process of legislators influencing public expenditures. In this section, we now identify whether there are corresponding impacts on public sector employment which legislators can influence through legislative actions as well as supervision of the bureaucracy.

AA Party Representation Schok and District-Level Public Employment

We test the causal impact of a district-level AA party representation shock on public sector employment using the following specification:

$$Y_{dsr} = \alpha_d + \delta_r + \theta_{sr} + \beta RAA_{dsr} + \gamma X_{dsr} + \epsilon_{dsr} \tag{1.9}$$

The outcome of interest is the natural log of the total number of working aged[37] individuals (subsequently referred to as 'workers') employed in the public sector for district d, located in state s, and measured during NSS survey round r. We choose the total number of jobs as our outcome of interest as it forms a direct measure of the quantum of benefits provided by the marginal AA party legislator. [38] A log transformation is warranted due to the rightward skew in the distribution of this variable. The log transformation also permits us to interpret the estimated coefficient of interest, β, in Equation (1.9) as the percent change in public sector employment in response to an additional close election won by an AA party. α, δ and θ correspond to district, survey round and state-survey round fixed effects. The inclusion of state-survey round fixed effects permits us to control for any state specific trends in AA party support during a given electoral period which might be correlated with district public hiring. RAA is the district-level AA party representation shock, based on the electoral cycle preceding survey round r.

Table 1.9 presents aggregated results showing the impact of an AA party representation shock on high and low caste workers across both rural and urban locations. Panel A contains the results for low caste workers while panel B contains results for high caste workers. Each specification controls for the total number of close elections contested between AA and non-AA parties in the district, and the logged population of working-aged individuals. The results in columns (4)-(6) also control for a number of political and socio-economic covariates.[39]. District, survey round and state-round fixed effects are also included and the specifications are weighted by the number of households in the district. Robust standard errors, clustered at the district level are presented in the parentheses.

Focusing on low caste workers in Panel A, we first note that the results are stable to the inclusion of covariates. In this regard, we focus on the coefficients in columns (4)-(6) which include covariates. Column (4) of Panel A in Table (1.9) show that the marginal AA party legislator causes a 5 percent increase in the number of low caste workers employed in public enterprises. Column (5) of Panel A shows that this effect is generated primarily through urban workers low caste workers - the marginal AA party legislator in the district increases the public hiring of urban low caste workers by 11 percent and the coefficient is highly significant. There is no corresponding effect on rural low caste workers with the coefficient

[37] This includes people aged between 19 and 65.

[38] An alternative outcome of interest would be the share of working aged individuals from low caste households employed in the public sector but this share can be rising over time solely due to demographic features such as a reduction in the number of low caste working age individuals. Regardless, we show in the Appendix that our results are unchanged if we use the share of individuals as a dependent variable.

[39] The set of controls include district-level voter turnout; respective voteshares of AA, INC and BJP; the average number of candidates contesting; years elapsed since the past election; district labour force participation rate; share of population who have completed at least secondary education; average age of workers; and the rate of urbanization

being both small in magnitude and statistically insignificant. In terms of gross magnitude, at the mean of the dependent variable, the marginal AA party legislator causes the hiring of nearly 4,000 urban low caste workers in public enterprises in the district . Finally, Panel B of Table (1.9) shows that the marginal AA party legislator has no impact on public hiring of high caste workers in either rural or urban locations.

AA Party Representation Shock and Public Employment: Differential Effects by AA Party Legislative Strength

In Section 1.2.2, we outlined that AA party legislators can affect public hiring of low caste workers through the passage of new legislation or enhanced implementation of existing affirmative action policies. We distinguish between these two channels by testing the differential impact of the marginal AA party legislator across the legislative strength of AA parties in the state legislature. If the public hiring of low caste workers is through the legislative channel and an expansion of affirmative action quotas, we would expect the marginal AA party legislator to have an impact only when AA parties have won a threshold level of elections to the state legislature. As the results in Section 1.6.1 show that the marginal AA party legislator impacts public expenditures only when AA parties have won at least 30 percent of the elections in the state, I test differential effects of the marginal legislator on public hiring across this cutoff. The specification tested is the following:

$$Y_{dsr} = \alpha_d + \delta_r + \theta_{sr} + \beta_1 RAA_{dsr} + \beta_2 \mathbb{I}(ShAAWin > 0.3)_{sy} * RAA_{dsr} + \gamma X_{dsr} + \epsilon_{dsr} \quad (1.10)$$

In equation (1.10), $\mathbb{I}(ShAAWin > 0.3)$ is defined as in equation (1.8). If AA party legislators affect public hiring of low caste workers solely through the supervision of the bureaucracy to implement existing laws, we would expect $\beta_1 > 0$ and $\beta_2 = 0$. If public hiring is determined through the legislative channel, we would expect $\beta_1 = 0$, $\beta_2 > 0$ and $\beta_1 + \beta_2 > 0$.

The results in Table 1.10 suggests that AA party legislators impact public employment for low caste workers primarily through the legislative channel. The coefficient on β_1 is positive for urban low caste workers and all low caste workers but imprecisely estimated. The point estimate is also half the magnitude of the average effect of the marginal AA party legislator on low caste public hiring across all districts (Panel A, Table 1.9). The interaction term, β_2 is positive in each instance, albeit not statistically significant and the sum of the coefficients are jointly significant at the 10 percent level for all low caste workers and at the 1 percent level for urban low caste workers.[40]. Thus, while we cannot rule out the role of the supervisory channel, particularly with regard to the hiring of urban low caste workers, the marginal AA party legislator has the most impact on the public hiring of low caste workers when AA parties have also won more than 30 percent of the elections in the state. This result again underlines the necessity of a substantial shift in political power before individual legislators can affect policy outcomes.

[40] For rural low caste workers, the sum of the coefficients has a p-value of 0.2

1.6.3 AA Party Representation Shock and Private Investment: Reduced Form Results

The reduced form results in Section 1.6.1 provide evidence on the redistributive spending preferences of AA party legislators in favour of welfare transfers benefiting low caste citizens and away from non-exclusive public investments in physical infrastructure. As infrastructure development is a critical component for long term growth, particularly through its impact on firm performance and entrepreneurial activity, we identify the impact of AA parties on state-level investment inflows.

We use two measures of private investment in this paper: fixed capital and capital formation in the manufacturing sector, reported by the ASI at the 2-digit industry level, aggregated across all manufacturing units operating in the state for each 2-digit industry. As per the ASI, gross capital formation is the sum of fixed capital formation and all other physical assets while fixed capital is the depreciated value of all fixed assets held by a manufacturing unit.[41] These two variables can be considered to be a measure of capital investments along the intensive margin for each industry. To measure capital investment along the extensive margin, we use the number of factories operating in the state for each 2-digit industry category. To identify whether AA party legislators affect returns to capital investments, we use net value added which is the difference between total outputs and total inputs in a year, net of depreciation. Finally, we also test the impact of AA party legislators on the hiring of workers to confirm that our results are not identifying a spurious relationship occuring due to the substitutability of capital and labour.

The specification used to identify the impact of AA party legislators on private investment is the following:

$$ln(Y_{isy}) = \alpha_s + \delta_y + \phi_{iy} + \beta RAA_{sy} + \lambda \mathbf{X}_{sy} + \epsilon_{isy} \tag{1.11}$$

In equation (1.11), Y is the industrial outcome of interest for the 2-digit industry i, located in state s and year y. In addition to state and year fixed effects, α and δ, we also include industry-year fixed effects ϕ to control for annual industry specific shocks.[42] The unit of observation is industry-state-year and the standard errors are clustered at the level of state-electoral cyle which is the level at which the treatment varies.

The reduced form results are presented in Table 1.11 with Panel A showing the results with state-level covariates and Panel B showing the results without covariates. The inclusion of covariates do not affect the magnitude or precision of the reduced form coefficients and all subsequent specifications include covariates. The reduced form coefficients in Panel A identify a negative relationship between an AA party representation shock and capital investments in the manufacturing sector along both the intensive and the extensive margins. Based on the reduced form coefficients, the marginal AA party legislator is associated with a 3 and 2 percent reduction in fixed capital (column 2) and gross capital formation (column 3). This is accompanied by a decline in manufacturing investment along the extensive margin

[41] Fixed assets include land, buildings, machinery and transport equipment owned by the manufacturing unit.

[42] The industry-year fixed effects also accounts for changes in the classification of 2-digit industries over time by incorporating a separate intercept for each 2-digit industry group for every year.

- the marginal AA party legislator reduces the number of factories operating by 1 percent, although in the absence of firm-level panel data, one cannot infer whether this is due to firms relocating operations to a different state or exiting the market. Consistent with a reduction in capital investment, gross output (column 4) also declines by 2 percent. Returns on capital investments are also lower in the presence of AA party legislators with the marginal legislator associated with a 3 percent decline in net value added (Column (5)). Column (6) informs us that this decrease in capital investment is not due to AA party legislators influencing firms in any way to hire more workers, resulting in a subsequent substitution of capital with labour - the marginal AA party legislator is associated with a 2 percent decline in the number of workers hired. Collectively, the results in Table 1.11 document the causal negative impact of AA party legislators on manufacturing investment, value added, output and the number of factories in operation. In the subsequent section, we provide empirical evidence consistent with the explanation that this negative impact stems from investors' response to the redistributive expenditure choices made by AA party legislators.

1.6.4 Establishing the Mechanism: Public Expenditure Choices of AA Party Legislators and Private Investment

In this section, we establish that a potential channel through which AA party legislators could be affecting private investments is through their expenditure choices. We contend that the reductions in infrastructure investment undertaken by AA party legislators dampens the prevailing investment climate and causes investors to reduce capital investments in states where AA parties are successful. We provide empirical evidence to establish this mechanism by identifying the heterogenous impact of AA party legislators on private investment and state expenditures across the legislative strength of AA parties, industry fuel consumption, and industry capital intensity. At the outset though, it is important to emphasize that this is not the only mechanism which can explain the negative impact of AA parties on manufacturing investment documented in Section 1.6.3. While the empirical evidence provided is consistent with our claim, in the absence of additional firm level evidence explaining investors' investment decisions, the results in this section should be interpreted as suggestive and not conclusive.

Differential Impact of AA Party Legislators on State Expenditures and Private Investment by Legislative Strength of AA Parties

The results in section 1.6.1 established that the marginal AA party legislator has a significant impact on public expenditures only when AA parties have won at least 30 percent of the elections to the state legislature. In this regard, if investors are responding to the redistributive spending choices made by AA party legislators, we would also expect the reductions in investment to be concentrated in states where AA parties have a corresponding presence in the state legislature. To ascertain if this is indeed the case, we test for the heterogeneous impacts of the marginal AA party legislator on manufacturing investment by the legislative strength of AA parties.

$$ln(Y_{isy}) = \alpha_s + \delta_y + \phi_{iy} + \beta_1 RAA_{sy} + \beta_2 \mathbb{I}(ShAAWin > 0.3)_{sy} * RAA_{sy} + \lambda \mathbf{X}_{sy} + \epsilon_{isy} \quad (1.12)$$

In equation (1.12), $\mathbb{I}(ShAAWin > 0.3)$ is defined as in equation (1.8), indicating the states in which AA parties have won more than 30 percent of the elections in any given electoral cycle. β_1 estimates the impact of the marginal AA party legislator on private investment when AA parties have won less than 30 percent of the elections in the state while β_2 estimates the additional impact of the marginal legislator when AA parties have won more than 30 percent of the elections in the state.

Table 1.12 presents the results from estimating this specification. The coefficient on β_2 is negative in each instance and statistically significant in two of the specifications.[43] The sum of the coefficients are also jointly significant for three out of the five dependent variables - namely number of factories (column 1), fixed capital (column 2) and output (column 4). Importantly, the coefficient on β_1 is in each instance positive, and not statistically significant, confirming that the marginal AA party legislator has no impact on private investment and hiring when AA parties do not have a critical level of support in the state legislature. This is consistent with the results in Section 1.6.1 where the marginal AA party legislator has an impact on public expenditures only when AA parties have won more than 30 percent of the elections in the state.

Differential Impact of AA Party Legislators on Private Investment by Industry Fuel Consumption

Section 1.6.1 established that the marginal AA party legislator has a significant negative impact on the share of state expenditures allocated to physical infrastructure. If this reduction in infrastructural spending drives investors' decisions to reduce manufacturing investment, we would expect the reduction in investment to be concentrated in states where the manufacturing sector is more reliant on physical infrastructure. In this regard, we test for the heterogeneous impact of AA party legislators on manufacturing investment across industries with high and low fuel consumption. Fuel consumption, as reported by the ASI includes the consumption of both electricity and fossil fuels. We normalize aggregate fuel consumption for each industry in the state by the total number of factories in operation. This provides us with the fuel consumption per factory for each industry-state-year combination. We would expect industries having higher values of fuel consumption per factory to be more reliant on electricity and roads and thereby, more responsive to reductions in public investments in electricity generation and road construction. We use an interaction term to test this hypothesis and estimate the following equation:

$$ln(Y_{isy}) = \alpha_s + \delta_y + \phi_i y + \beta_1 RAA_{sy} + \beta_2 HighFuel_{isy} * RAA_{sy} + \gamma_2 X_{sy} + \epsilon_{isy} \quad (1.13)$$

In Equation (1.13), *HighFuel* is a categorical variable equaling 1 if the aggregate fuel consumption per factory for any industry i, in state s and year y exceeds the median fuel

[43] For those which are not significant at the conventional levels of statistical significance, the p-values are never greater than 0.25.

consumption per factory across all industries in year y. The results from this estimation, shown in Table 1.13 are consistent with the our proposed hypothesis. The coefficient on the interaction term, β_2 is negative and statistically significant in each specification while the coefficient on β_1 is not significant in any of the specifications. The sum of the coefficients are also jointly significant for each outcome, confirming that the marginal AA party legislator is associated with a causal reduction in private investment only for industries having a high level of fuel consumption per factory.

Differential Impact of AA Party Legislators on Private Investment by Industry Capital Intensity

The results in the previous section established that the negative impact of AA party legislators on private investment is restricted to firms with high levels of fuel consumption. As we would expect firms with high capital intensity to also have higher levels of fuel consumption, our final test of the mechanism is a test of heterogenous effects by the capital intensity of firms. Capital intensity is defined as the ratio of fixed capital and workers hired. We define industries with high capital intensity if the ratio of fixed capital and workers hired for that industry exceeds the median fixed capital to worker ratio across all states for that year. To this extent, the specification tested is identical to equation (1.13) with the exception that instead of *HighFuel*, the dummy variable used is *CapInt* denoting high capital intensity. The results in Table 1.14 confirm this hypothesis and are very similar to those obtained in Table 1.13. The interaction term is negative and statistically significant in 4 out of 5 specifications. β_1 is not statisitcally significant in either specification and the sum of β_1 and β_2 is statistically significant in each specification, signifying that the marginal AA party legislator has a negative impact on private investment only for industries which have high capital intensity.

1.7 Robustness Checks

In this section, we present robustness checks to verify that the results discussed in Section 6 are stable to changes in the definition of close elections, alternate definitions of AA parties, alternate levels of clustering, restrictions in sample size and the exclusion of population weights. Two of our robustness checks are specific to the district-level results while one robustness check is only for the state-level results. In this respect, we first show the results from robustness checks which are common to both the district and the state-level specifications, before presenting the robustness checks which apply exclusively to the state (district) level results.

1.7.1 Altering the Threshold to Determine Close Elections

All the results discussed in Section 6 are based on the threshold where a 5 percent difference in vote share between the winner and the runners-up in an election constitute a 'close election'. A natural concern therefore is whether our results are sensitive to this definition of close elections. We thereby test the validity of our results at 2 alternate and narrower margins of close elections: namely 4 and 2 percent. The results are presented in Tables 1.15, 1.16, and

1.17. The first three columns in each table shows the results when the threshold for close elections is 4 percent; the last three columns shows the results when the threshold is further lowered to 2 percent.

At the 4 percent threshold for close elections, the results are very much unchanged. At the state level, the marginal AA party legislator continues to have a significant positive impact on the share of expenditures allocated to targeted and untargeted low caste welfare, accompanied by a negative and significant impact on physical infrastructure spending (Table 1.15). The causal impact of AA parties on manufacturing investment is negative but the precision of the coefficients are dampened but the signs remain unaltered (Table (1.17), Panel A). We are also able to detect a positive and significant impact of the marginal AA party legislator on public sector employment for urban low caste workers (Table 1.16) but no corresponding impact on rural low caste workers.

At the 2 percent margin for close elections, the state-level results on gross capital formation and fixed capital remain negative and statistically significant while the remaining coefficients remain negative and significant at the 15 percent level (Table 1.17, Panel B). For state expenditures, we are able to detect a negative and statistically significant impact of the marginal AA party legislator on roads and power investments while the coefficients for targeted and untargeted low caste welfare are positive and statistically indistinguishable from those obtained with a 5 percent threshold for close elections, albeit imprecisely estimated (Table 1.15, Columns (4) and (5)). At the district level, the reduced form coefficient estimating the impact of AA parties on public sector employment for urban low caste workers (Table 1.16, column (5)) remains positive and is almost significant at the 10 percent margin.

1.7.2 Robustness Check: Verifying that Results are not Driven by Individual State

The primary source of exogenous variation in the paper is quasi-random shocks to the number of AA party legislators elected to the state legislative assembly in each electoral cycle. As the majority of AA parties are regional parties with a powerful presence in their respective states but little influence outside, it leads to a concern whether the results are being driven by the impact of a particular AA party in a single state. To verify this isn't the case, we re-estimate our results for the key outcome variables by dropping one state at a time. The results from this exercise is presented as coefficient plots in Figures 1.5, 1.6 and 1.7. The 19 point estimates are very similar in all the coefficient plots for each of the public expenditure, public employment and private investment specifications, confirming that the core results are not being driven by the impact of a specific AA party operating within a single state.

1.7.3 Robustness Check: Restricting the Classification of AA Parties to Strictly Caste Based Parties

In our discussion of AA parties in Section 1.3.2, we had argued for the inclusion of the Left parties and the DMK within the 'AA' category due to the respective similarities in their political goals, the overlap in their target voters, and the electoral support provided by the latter parties to the exclusively caste-based AA parties. However, both the DMK

and the Left parties were formed much before the remaining AA parties and thereby, have had more experienced legislators at their disposal. The DMK and the Left parties have also formed governments in three states on multiple occasions between 1967 and 1987, providing the parties with greater administrative experience. All these factors can make legislators belonging to the Left parties and the DMK more effective in implementing their preferred policies, vis-a-vis the relatively more inexperienced legislators from AA parties. If this is true, the positive impact of AA party representation could be driven solely due to these two party groups, which represent low caste interests, but did not emerge out of the upsurge in political mobilization amongst low caste populations in the 1980s. This would negate the paper's argument that the increased welfare allocations for low caste citizens are driven by a change in *de facto* political power for low caste populations through the process of electoral competition. To verify this is not the case, we re-run our specifications after excluding the states of West Bengal, Kerala and Tamil Nadu, in which the Left and the DMK are the strongest.[44]

The results are presented in Tables 1.18, 1.19, and 1.20. On the whole, the results are qualitatively unchanged after restricting the sample. At the state level, the restriction in sample size affects the precision of some of our coefficients. Nonetheless, we still identify a positive and statistically significant impact of AA parties on the share of state expenditures allocated exclusively towards low caste citizens; and a negative and statistically significant impact on the share of expenditures allocated towards roads and power investments. The district-level results also report a positive and statistically significant impact of AA parties on public employment for urban low caste workers.

1.7.4 Robustness Check: Clustering at the State Level

As our quasi-random treatment effect varies with every electoral cycle in the state, we had argued for the clustering of the standard errors for the state-level results at the level of state-electoral cycle.[45] To validate that the precision of the results are not an artifact of this choice of clustering, we re-estimate the specifications after clustering the standard errors at the traditional state-level. From Tables 1.21 and 1.22, we verify that this is not the case. The marginal AA party legislator has a statistically significant impact on targeted low caste spending as well as road construction and investments in electricity generation (Table 1.21) while the coefficient for untargeted low caste welfare spending is almost significant at the 15 percent level (p-value of 0.159). The precision of the results estimating the causal impact of AA parties on manufacturing investment remains unchanged upon clustering at the state level.

[44] The DMK's political presence is strictly restricted to the southern state of Tamil Nadu while over 80 percent of the state assembly races won by the Left are located in West Bengal and Kerala.

[45] The small number of states (19) in our sample is a second reason to cluster the standard errors at the level of state-electoral cycle.

1.7.5 Restricting the Sample to States Where AA Parties Never Formed Coalitions with Non-AA Parties

The multi-party nature of Indian democracy, combined with a first-past-the-post method of determining electoral outcomes means that political parties often form alliances with each other to strengthen their electoral prospects. While some alliances are formed between political parties on ideological grounds, such as the Left and the JD, some alliances are also born out of political opportunism. While political alliances between multiple AA (non-AA) parties do not pose any threats to our identification, political alliances between AA and non-AA parties can serve as a potential confound to the identification strategy. For instance, a pre-poll political alliance between an AA and a non-AA party can result in a smaller number of close elections between AA and non-AA parties, potentially biasing our estimated impact of AA parties.[46] In our sample, AA parties and non-AA parties have formed significant pre-poll alliances in three states - Orissa, Bihar and Jharkhand - between 1998 and 2010, affecting approximately a fifth of our districts. Aside from that, there have been minor coalitions formed between AA and non-AA parties in Tamil Nadu in 2006 and 2011 respectively. To counter this potential confound, we present the impact of AA party legislators on low caste public employment from a sample restricted to three states - West Bengal, Kerala and Uttar Pradesh - where AA parties have never formed alliances with non-AA parties, while retaining a major electoral presence in each state.[47] The results are presented in Table 1.23. The sample size is greatly reduced[48] but we are still able to detect a large and statistically significant positive impact of the marginal AA party legislator on public employment for urban low caste workers.

1.7.6 Restricting the Sample to Districts with At Least 1 Close Election

In Section 2.4.1, we had illustrated that roughly a sixth of the elections in our sample can be classified as close elections involving an AA and a non-AA party with the difference in vote shares between the winner and the loser being less than 5 percent. At the district level, this implies that in a third of our sample districts, there are no close elections between AA and non-AA parties, and our measure of district-level AA representation shock, RAA is by default 0.[49] To ensure that this is not affecting our results, we restrict our sample of districts to those in which there is at least 1 close election between AA and non-AA parties across

[46] Political alliance between an AA and a non-AA party can constrain the AA party's ability to target state resources to low caste populations. The fact that we find evidence of public resource targeting by AA parties towards low castes despite these alliances (major alliances has been in Bihar and Orissa where the ruling AA party had a coalition with the BJP for over 10 years in both states.) suggests that our estimates might be understating the true impact of AA parties.

[47] In Uttar Pradesh, the BSP, which is a key AA party twice formed coalitions with the BJP - a key non-AA party. However, both the coalitions were *post-poll* coalitions and also, did not last for more than 1 year in each instance.

[48] We do not undertake this robustness check at the state-level due to the reduced sample.

[49] Due to the occurrence of at least 1 close election between an AA and a non-AA party at the 5 percent level, our state-level results are not subject to any concern in this regard

the 4 NSS survey rounds and re-estimate our specifications. The results are presented in Table 1.24 and are almost identical to those obtained with the full sample. The impact of the marginal AA party legislator on public employment for urban low caste workers remains positive and statistically significant.

1.7.7 Robustness Check: Excluding Population Weights from Regressions

The district level employment results in Section 1.6.2 are population weighted. The rationale for population weighting is that the NSS' district-level sample size is proportional to the district's population and the NSS uses different sampling frames for urban and rural populations. In this section, we verify that the core results are not being driven by our choice of populations weights. Table 1.25 shows the results in the absence of population weights - while the precision of the coefficients is slightly affected, the coefficient remains stable in magnitude to the exclusion of these weights.

1.8 Conclusion

This paper uses a unique political setting from India to study the economic impacts of increased political representation for marginalized citizens through the process of political mobilization and electoral competition. The existence of institutional reforms designed to increase the political representation of economically deprived low caste citizens through electoral quotas permits us to contrast the welfare gains emanating from these two distinct modes of political representation for disadvantaged poplations. The core empirical results show that a change in the structure of political power, achieved through the electoral success of caste-based AA parties, is neessary to complement the institution of electoral quotas for the latter to impact public resource allocation for low caste citizens. The paper therefore provides causal evidence in support of Acemoglu and Robinson (2008) and Acemoglu et al's (2013) prediction that changes in *de facto* political power need to be accompanied with changes in *de jure* political power to realize the latter's full redistributive potential. Relatedly, by showing that the mandated political representation of ethnic minority politicians have an impact on public resource allocation only when the politicians belong to political parties with an explicit policy preference for minority welfare, the paper provides alternative evidence to the body literature studying the impact of a legislator's ascriptive identity on economic outcomes.

By highlighting the criticality of party platforms in facilitating public resource allocation towards low caste populations, the paper also provides evidence that the policy preferences of political parties can have a significant influence on economic outcomes, both for select population groups, and the regional economy. Fereira and Gyourko (2009) who find no impact of party affiliation on local economic outcomes contend that cities might not be the appropriate geographical unit to study the impact of political parties on economic outcomes. This paper identifies the impact of party platforms based on the legislative actions of state-level politicians committed to their party mandates, suggesting that an appropriate level to study the policy impact of political parties could be the state, where the legislative actions

of elected politicians have a much wider bearing on the economy.

Finally, by linking the redistributive spending preferences of elected representatives with the negative impact on regional private investment, the paper showcases the possible economic tradeoffs stemming from redistributive spending policies undertaken by fiscally constrained legislators. It needs to be stressed however that increased welfare spending for disadvantaged populations need not come at a net cost to the regional economy if governments have the necessary state capacity to effectively finance the enhanced levels of social expenditures. In cases such as India, where lower levels of governments have limited taxation powers, advances in the administration and collection of indirect taxes, along with improvements in the efficacy of public investments can reduce the potential distortions associated with higher spending on social protection schemes for vulnerable populations.

1.9 Figures

Figure 1.1: McCrary Test for Discontinuity of AA Victory Margin

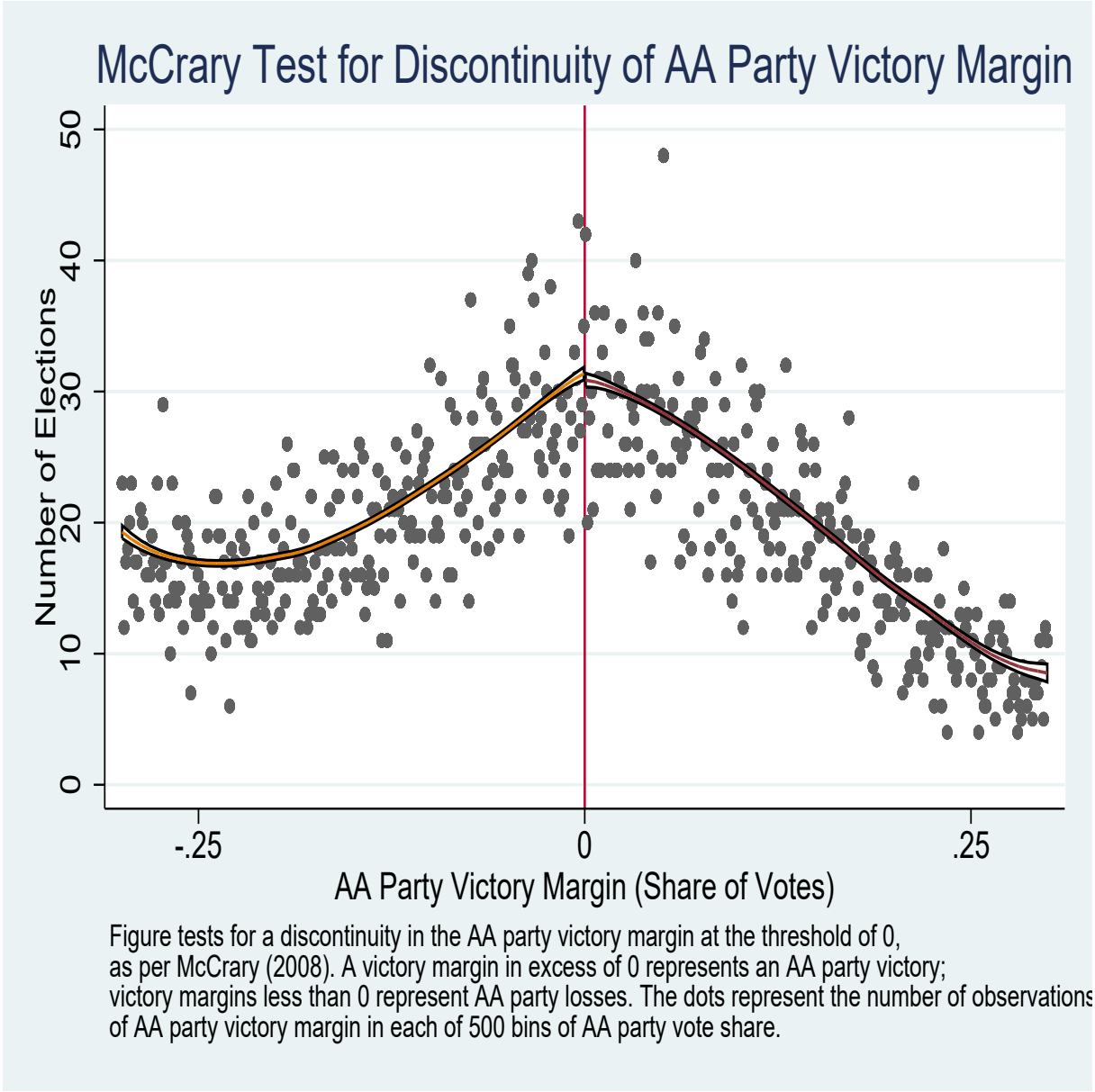
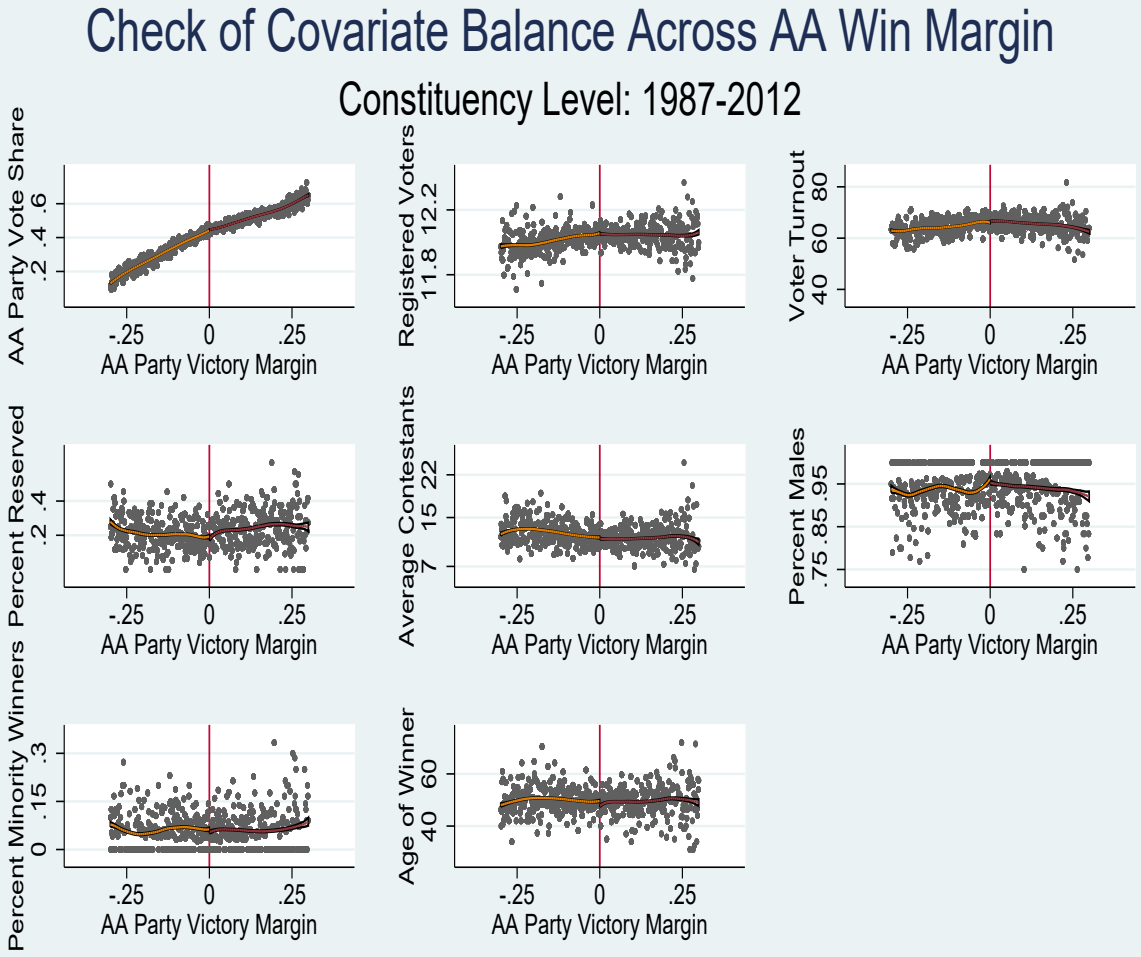


Figure 1.2: Balance Across Constituency Level Covariates



Covariate balance checks at the constituency level. The following covariates are tested: AA party vote share; number of registered voters; voter turnout; share of constituencies reserved for SC/ST candidates; number of candidates contesting; share of male winners; share of winners from SC/ST community; share of male winners. Each dot corresponds to unconditional means across 500 bins of AA party vote share. The solid lines represent 95 percent confidence intervals.

Figure 1.3: Predicted AA Party Close Wins vs Actual AA Party Close Wins, State

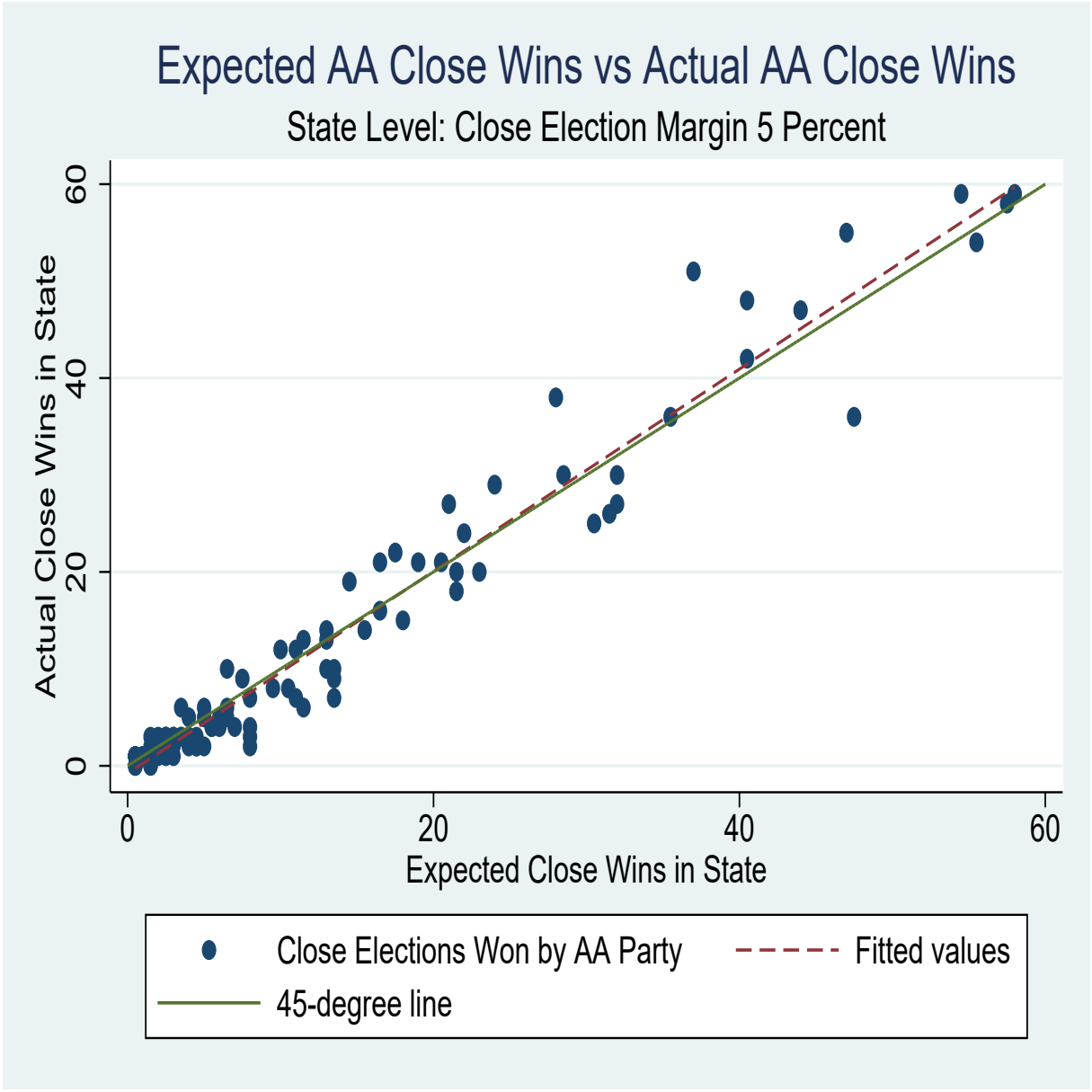


Figure 1.4: Predicted AA Party Close Wins vs Actual AA Party Close Wins, District

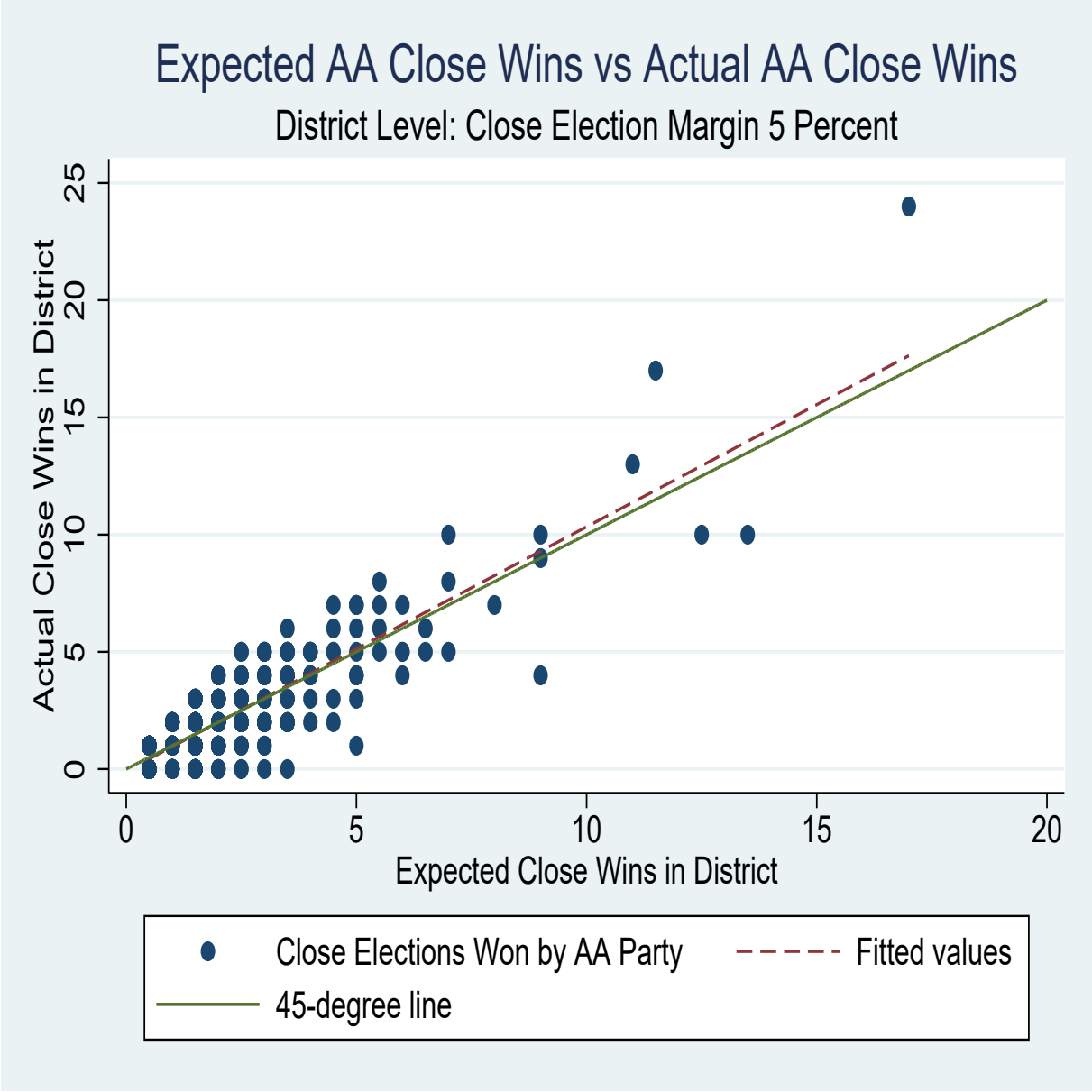


Figure 1.5: Robustness of Public Expenditure Results to Dropping Individual States

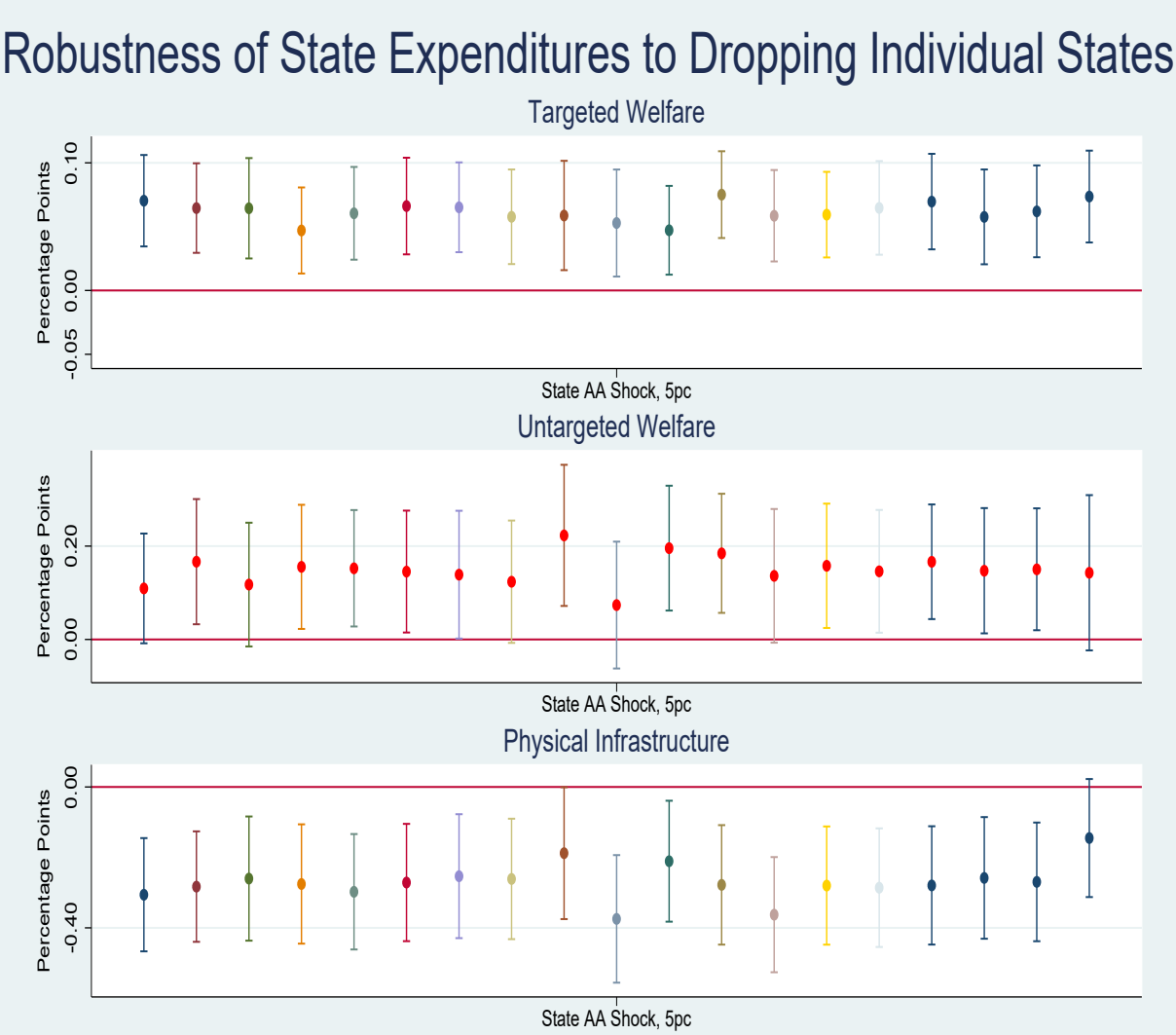


Figure 1.6: Robustness of Private Employment Results to Dropping Individual States

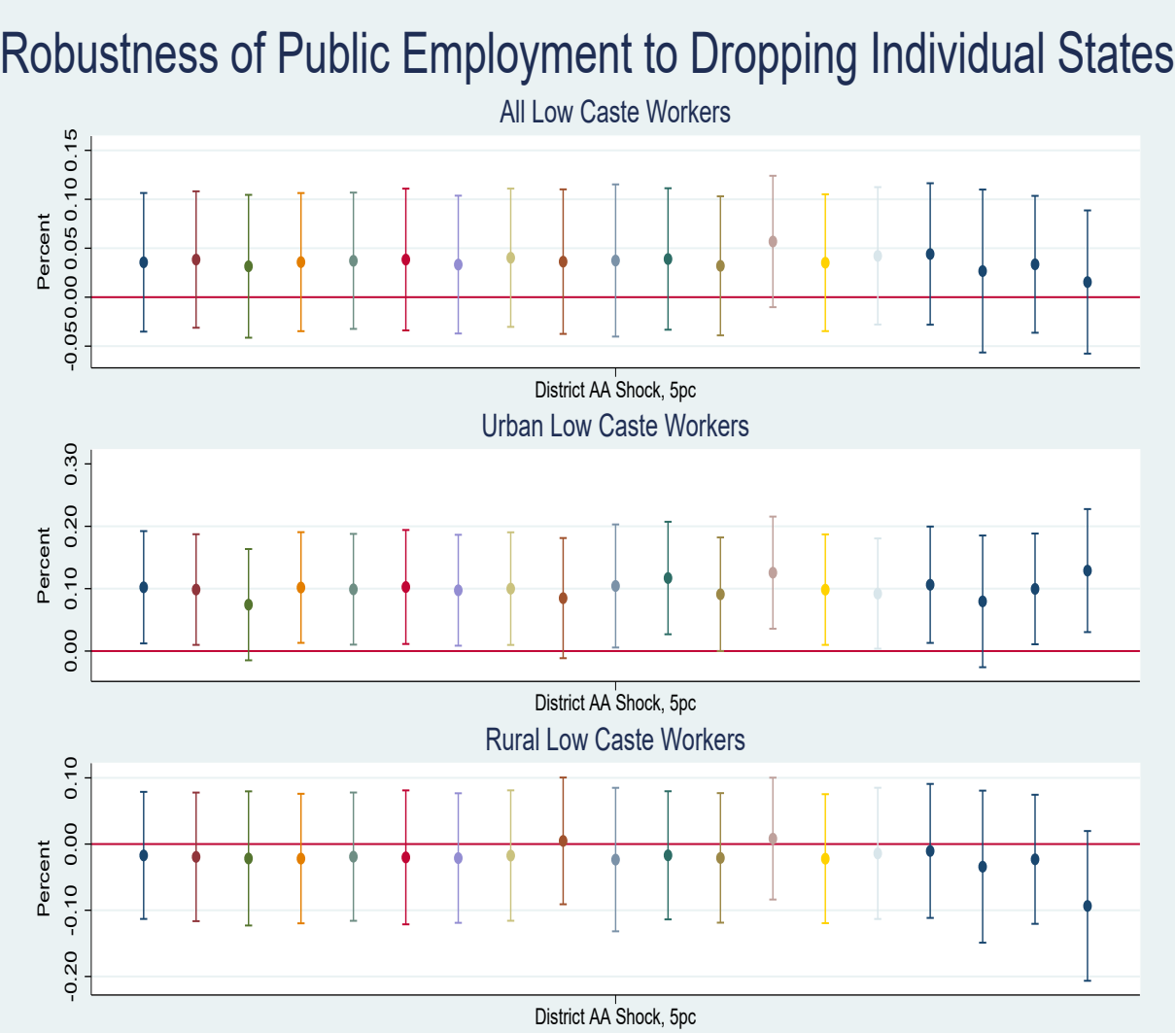
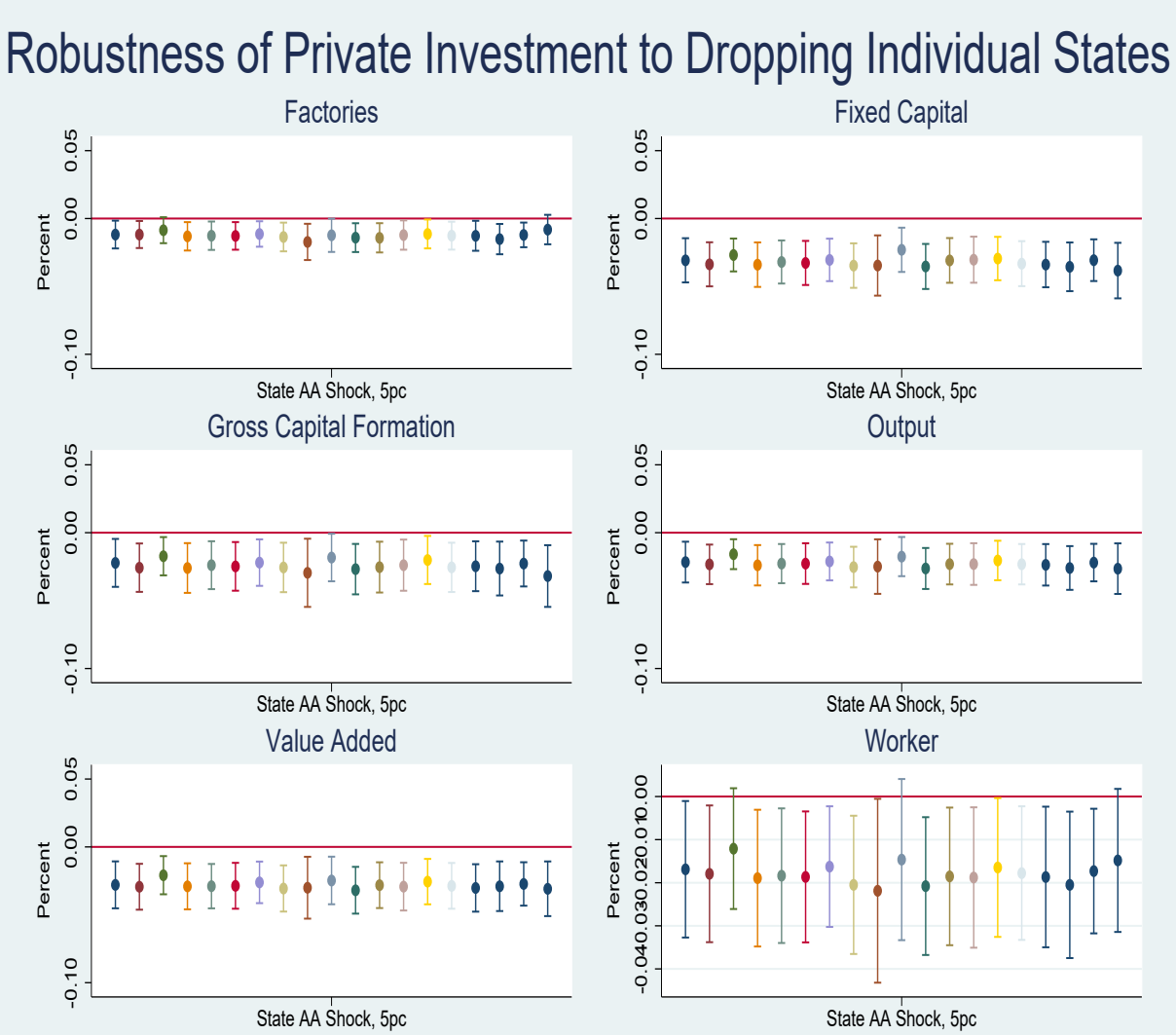


Figure 1.7: Robustness of Private Investment Results to Dropping Individual States



1.10 Tables

1.10.1 Main Results

Table 1.1: Verifying Political Covariates Do Not Predict State-Level AA Representation Shock at 5 Percent Margin

	(1) AA Shock	(2) AA Shock	(3) AA Shock	(4) AA Shock	(5) AA Shock	(6) AA Shock	(7) AA Shock
Total Elections, State	-.0182						
	(.0163)						
No. of Reserved Constituencies		-.0358					
		(.0309)					
Average Effective Number of Parties			-.1372				
			(.4018)				
State AA Vote Share				.3427			
				(.6138)			
State Voter Turnout					-1.3264		
					(11.0620)		
Percent AA Wins, Lag 1						2.3496	
						(2.4217)	
Total Close Elections, 5 pc	-.0665	-.0652	-.0298**	-.0698	-.0642	-.1396***	-.0658
	(.0510)	(.0511)	(.0127)	(.0539)	(.0568)	(.0504)	(.0509)
Observations	106	106	106	106	106	87	106
R^2	.4968	.4966	.4670	.4942	.4922	.6300	.4921

Standard errors in parentheses (clustered at the state-electoral cycle level). The dependent variable in each specification is the difference between the total close elections won by AA parties in the state (5 percent margin), and half the close elections contested by AA parties in the state. All regressions control state and election year fixed effects, in addition to the total number of close elections contested by AA parties.

Table 1.2: Verifying Demographic Covariates Do Not Predict State-Level AA Representation Shock

	(1) AA Shock	(2) AA Shock	(3) AA Shock	(4) AA Shock	(5) AA Shock	(6) AA Shock	(7) AA Shock	(8) AA Shock
State Population (log)	-2.6437							
	(5.9119)							
Percent Population SC/ST		-12.1549						
		(43.6627)						
Percent Literate			9.8666					
			(31.5986)					
Percent Workers				-39.2192				
				(43.0443)				
Percent Urban					12.6114			
					(29.7530)			
Gender Ratio						-37.4498		
						(77.4185)		
Population Density							-.0043	
							(.0078)	
State Domestic Product (log)								-.9605
								(2.5541)
Observations	74	74	74	74	74	74	74	84
R^2	.6937	.6932	.6942	.7000	.6962	.6944	.6997	.6320

Standard errors in parentheses (clustered at the state-electoral cycle level). The dependent variable in each specification is the difference between the total close elections (5 percent margin) won by AA parties in the state, and half the close elections contested by AA parties in the state. All regressions control for state and year fixed effects, in addition to the total number of close elections contested by AA parties.

Table 1.3: Detetermining Exogeneity of District-Level AA Party Representation Shock

	(1) AA Shock	(2) AA Shock	(3) AA Shock	(4) AA Shock	(5) AA Shock	(6) AA Shock	(7) AA Shock
Total Elections in District	-.0036						
	(.0032)						
Percent Reserved Elections		-.0132					
		(.0788)					
Average Contestants			-.0070				
			(.0045)				
District AA Vote Share				.0947			
				(.2113)			
District Voter Turnout					.0032		
					(.2253)		
Percent of Elections Won by AA Parties, Lag 1						-.0558	
						(.1186)	
No. Close Elections, 5 pct	.0314	.0249	.0258	.0216	.0250	.0281	.0250
	(.0305)	(.0284)	(.0277)	(.0351)	(.0278)	(.0321)	(.0280)
Observations	1155	1155	1155	1155	1155	1141	1155
R^2	.0113	.0097	.0123	.0103	.0097	.0105	.0097

Standard errors in parentheses (clustered at the district level). The dependent variable in each specification is the number of close elections won by AA parties in the district at the 5 percent margin, less half the number of close elections contested by AA parties against non-AA parties. The unit of observation is district-survey round. Estimation controls for the total number of close contests.

Table 1.4: Detetermining Exogeneity of District-Level AA Party Representation Shock

	(1) AA Shock	(2) AA Shock	(3) AA Shock	(4) AA Shock	(5) AA Shock	(6) AA Shock	(7) AA Shock	(8) AA Shock
Population (Log)	-.0490							
	(.0361)							
Percent Urban		.1463						
		(.1211)						
Percent Low Caste			-.0527					
			(.1172)					
Age				-.0048				
				(.0176)				
Household Size					.0274			
					(.0270)			
Literacy						.0259		
						(.1252)		
Household Consumption (Log)							-.0448	
							(.0397)	
Headcount Ratio								-.0376
								(.1103)
Observations	1058	1058	1058	1058	1058	1058	1058	1058
R^2	.0136	.0122	.0113	.0113	.0120	.0112	.0121	.0113

Standard errors in parentheses (clustered at the district level). The dependent variable in each specification is the number of close elections won (5 percent margin) by AA parties in the district, less half the number of close elections contested by AA parties against non-AA parties. The unit of observation is district-survey round. Estimation controls for the total number of close contests.

Table 1.5: AA Party Representation Shock and State Expenditures

			Panel A: With Covariates			
	(1) Low Caste Welfare	(2) Rural Dev and Soc. Welfare	(3) Roads and Power	(4) Health and Educ.	(5) Ag. and Irrig.	(6) Pensions and Admin.
State AA Shock, 5pc	.0617***	.1499*	-.2708***	-.0163	-.0564	.1431*
	(.0214)	(.0786)	(.1006)	(.0925)	(.0758)	(.0728)
Observations	363	363	363	363	363	363
R^2	.90	.77	.54	.80	.70	.82
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Variable Mean (Pct)	3.10	13.40	8.42	29.77	17.05	11.59
			Panel B: Without Covariates			
	(1) Low Caste Welfare	(2) Rural Dev and Soc. Welfare	(3) Roads and Power	(4) Health and Educ.	(5) Ag. and Irrig.	(6) Pensions and Admin.
State AA Shock, 5pc	.0393	.1564*	-.2081**	-.0160	-.0380	.1843**
	(.0265)	(.0807)	(.1046)	(.0885)	(.0759)	(.0887)
Observations	364	364	364	364	364	364
R^2	.86	.76	.46	.78	.68	.78
Controls	No	No	No	No	No	No
Dep. Variable Mean (Pct)	3.10	13.40	8.42	29.77	17.05	11.59

Standard errors in parentheses (clustered at the state-electoral cycle level). The dependent variable in each specification is the share of state expenditures for each state expenditure category. The unit of observation is state-year. Estimation controls for total number of close contests involving AA parties; state-level political competition; percent of constituencies reserved for SC/ST candidates; share of population who are literate; share of workers; rate of urbanization; as well as state and year fixed effects.

Table 1.6: AA Party Representation Shock and State Government Expenditures, Revenue and Deficit

	(1) Total Expenditures	(2) Per Capita Expenditures	(3) Total Revenues	(4) Per Capita Revenues	(5) Deficit Pct NSDP
State AA Shock, 5pc	-.0033	-.0050	-.0005	-.0021	-.0001
	(.0033)	(.0033)	(.0023)	(.0022)	(.0004)
Observations	326	326	326	326	326
R^2	.97	.95	.99	.98	.63
Controls	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	4.16e+06	5550.2929	4.38e+06	5885.3131	.0447

Standard errors in parentheses (clustered at the level of state). The unit of observation is state-year. Estimation controls for total number of close contests; state-level political competition; voteshares for AA parties; share of minorities; share of population who are literate; share of workers; rate of urbanization; constant net state domestic product (logged); as well as state and year fixed effects.

Table 1.7: Effect of AA Party Representation Shock on Public Expenditures - Differential Effects by AA Party Legislative Strength

	(1) Low Caste Spending	(2) Rural Dev and Social Welfare	(3) Roads and Power Investment
State AA Shock, 5pc	.0299	-.1660	-.2693
	(.0758)	(.3017)	(.2724)
Elections Won, 30pc*State AA Shock, 5pc	.0329	.3897	-.0125
	(.0800)	(.3148)	(.3106)
Observations	363	363	363
R^2	.90	.78	.54
Joint Significance - 5pc	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Dependent Variable Mean (Pct)	3.1037	13.3994	8.4225

Standard errors in parentheses (clustered at the state-electoral cycle level). The dependent variable in each specification is the share of each expenditure category in total state expenditures. The unit of observation is state-year round. Estimation controls for total number of close contests; state-level political competition; voteshares for AA parties and two major national parties; share of constituencies reserved for minority candidates; share of population who are literate; share of workers; rate of urbanization; constant net state domestic product; as well as state and year fixed effects.

Table 1.8: Comparing the Impact of Low Caste Politicians on Public Expenditures for Low Caste Citizens Across Party Affiliation

	(1) Low Caste Welfare	(2) Low Caste Welfare	(3) Low Caste Welfare	(4) Low Caste Welfare	(5) Low Caste Welfare	(6) Low Caste Welfare
State AA Shock Reserved , 5pc	.0698*					
	(.0355)					
State AA Shock Non-Reserved, 5pc		.0672**				
		(.0278)				
State BJP Shock, Reserved 5pc			-.1919***			
			(.0594)			
State BJP Shock, Non-Reserved 5pc				-.0081		
				(.0328)		
State Congress Shock, Reserved 5pc					.0247	
					(.0244)	
State Congress Shock, Non-Reserved 5pc						.0012
						(.0180)
Observations	363	363	363	363	363	363
R^2	.90	.90	.91	.90	.89	.89
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Pct)	3.10	3.10	3.10	3.10	3.10	3.10

Standard errors in parentheses (clustered at the state-electoral cycle level). The dependent variable in each specification is the share of state expenditures allocated towards low caste welfare. The unit of observation is state-year. Estimation controls for total number of close contests; state-level political competition; voteshares for AA parties and two major national parties; share of constituencies reserved for minority candidates; share of population who are literate; share of workers; rate of urbanization; constant net state domestic product; as well as state and year fixed effects.

Table 1.9: AA Party Representation Shock and District Level Public Employment

			Panel A: Low Caste Workers			
	(1) Public Jobs All	(2) Public Jobs Urban	(3) Public Jobs Rural	(4) Public Jobs All	(5) Public Jobs Urban	(6) Public Jobs Rural
District AA Shock, 5pc	.0466	.1158***	.0120	.0520*	.1101***	.0161
	(.0323)	(.0427)	(.0391)	(.0303)	(.0418)	(.0373)
Observations	1029	1016	1025	1024	1001	1010
R^2	.82	.78	.73	.84	.81	.75
Covariates	No	No	No	Yes	Yes	Yes
Dep. Variable Mean (Workers)	73834	35373	39049	73834	35373	39049
			Panel B: High Caste Workers			
	(1) Public All	(2) Public Urban	(3) Public Jobs Rural	(4) Public Jobs All	(5) Public Jobs Urban	(6) Public Jobs Rural
District AA Shock, 5pc	.0057	-.0047	.0490	-.0325	-.0372	.0302
	(.0542)	(.0628)	(.0670)	(.0482)	(.0547)	(.0629)
Observations	1027	1016	1016	1023	1001	1002
R^2	.78	.77	.68	.81	.81	.71
Covariates	No	No	No	Yes	Yes	Yes
Dep. Variable Mean (Workers)	43827	30251	13873	43827	30251	13873

Standard errors in parentheses (clustered at the district level). Panel A presents estimates for low caste workers; panel B presents estimates for high caste workers. The dependent variable in each specification is the logged number of low (high) caste workers employed in the public sector. The unit of observation is district-survey round. All estimation controls for the high caste working age population and total number of close contests. Estimations in columns (4)-(6) also control for district-level political competition; voteshares for AA and the two major national parties - BJP and INC; share of low caste and Muslim population; share of high caste workers with at least a secondary education; average age of high caste workers; gender ratio; rate of unemployment; and rate of urbanization. All specifications also control for district, survey round and state-survey round fixed effects. The results are weighted by the number of households in each district.

Table 1.10: AA Party Representation Shock and District Public Employment for Low Caste Workers - Differential Effects by Legislative Strength of AA Party

	(1) Public Jobs All	(2) Public Jobs Urban	(3) Public Jobs Rural
District AA Shock, 5pc	.0236	.0622	-.0909
	(.0405)	(.0489)	(.0696)
Elections Won, 30pc*District AA Shock, 5pc	.0395	.0671	.1478
	(.0604)	(.0795)	(.0906)
Observations	1024	1001	1010
R^2	.84	.81	.75
Joint Significance, 10pc	Yes	Yes	No
Covariates	Yes	Yes	Yes
Dependent Variable Mean (Workers)	73834	35373	39049

Standard errors in parentheses (clustered at the district level). The dependent variable in each specification is the logged number of low caste workers employed in the public sector. The unit of observation is district-survey round. All estimation controls for the low caste working age population and total number of close contests. Estimations in columns (4)-(6) also control for district-level political competition; voteshares for AA and the two major national parties - BJP and INC; share of low caste and Muslim population; share of low caste workers with at least a secondary education; average age of low caste workers; gender ratio; rate of unemployment; and rate of urbanization. All specifications also control for district, survey round and state-survey round fixed effects. The results are weighted by the number of households in each district.

Table 1.11: AA Party Representation Shock and Manufacturing Investment

			Panel A: With Covariates			
	(1)	(2)	(3)	(4)	(5)	(6)
	Factories	Fixed Capital	Capital Formation	Output	Value Added	Workers
State AA Shock, 5pc	-.0127**	-.0327***	-.0249**	-.0233***	-.0288***	-.0183*
	(.0061)	(.0097)	(.0108)	(.0088)	(.0100)	(.0093)
Observations	7158	7156	6457	7153	6998	7155
R^2	.76	.70	.63	.69	.64	.68
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	330	270333	61766	677943	114720	17320
			Panel B: Without Covariates			
	(1)	(2)	(3)	(4)	(5)	(6)
	Factories	Fixed Capital	Capital Formation	Output	Value Added	Workers
State AA Shock, 5pc	-.0113*	-.0304***	-.0221**	-.0208**	-.0256**	-.0175*
	(.0061)	(.0096)	(.0099)	(.0090)	(.0102)	(.0091)
Observations	7180	7178	6477	7175	7018	7177
R^2	.76	.70	.63	.69	.64	.67
Controls	No	No	No	No	No	No
Dependent Variable Mean	330	270333	61766	677943	114720	17320

Standard errors in parentheses (clustered at level of state-electoral cycle). The dependent variable in each specification is logged. Panel A presents results with covariates and Panel B presents results without covariates. Estimates in Panel A controls for the total number of close contests; rates of literacy, urbanization and share of workers; state-level voter turnout; and effective number of parties contesting. State, industry-year and year fixed effects are also included in each regression.

Table 1.12: AA Party Representation Shock and Manufacturing Industries - Differential Effects by Legislative Strength of AA Parties

	(1)	(2)	(3)	(4)	(5)	(6)
	Factories	Fixed Capital	Capital Formation	Output	Net Value Added	Workers
State AA Shock, 5pc	.0119	.0114	.0130	.0043	.0030	.0085
	(.0119)	(.0232)	(.0221)	(.0179)	(.0201)	(.0164)
Elections Won, 30pc*State AA Shock, 5pc	-.0291**	-.0452*	-.0328	-.0254	-.0350	-.0285
	(.0145)	(.0259)	(.0262)	(.0211)	(.0250)	(.0221)
Observations	7158	7156	6457	7153	6998	7155
R^2	.76	.70	.63	.69	.64	.68
Joint Significance - 10pc	Yes	Yes	No	Yes	Yes	No
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	330	270333	61766	677943	114720	17320

Standard errors in parentheses (clustered at the state-electoral level). The dependent variable in each specification is logged. Estimation controls for the total number of close contests; rate of literacy; state-level voter turnout; effective number of parties contesting; demographic characterstics such as urbanization, gender ratio, population density, and percent of workers. State, industry-year and year fixed effects are also included.

Table 1.13: AA Party Representation Shock and Manufacturing Industries - Differential Effects by Industry Fuel Consumption

	(1)	(2)	(3)	(4)	(5)	(6)
	Factories	Fixed Capital	Capital Formation	Output	Net Value Added	Workers
State AA Shock, 5pc	-.0002	-.0119	.0047	.0008	-.0025	-.0042
	(.0067)	(.0131)	(.0134)	(.0118)	(.0118)	(.0111)
High Fuel*AA Shock, 5pc	-.0267***	-.0360**	-.0553***	-.0445**	-.0489***	-.0257*
	(.0079)	(.0179)	(.0161)	(.0174)	(.0175)	(.0141)
Observations	7158	7156	6457	7153	6998	7155
R^2	.76	.74	.66	.72	.68	.70
Joint Significance - 10pc	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	330	270333	61766	677943	114720	17320

Standard errors in parentheses (clustered at the state-electoral level). The dependent variable in each specification is logged. Estimation controls for the total number of close contests; rate of literacy; state-level voter turnout; effective number of parties contesting; demographic characterstics such as urbanization, gender ratio, population density, and percent of workers. State, industry-year and year fixed effects are also included.

Table 1.14: AA Party Representation Shock and Manufacturing Industries - Differential Effects by Capital Intensive Industries

	(1) Factories	(2) Fixed Capital	(3) Capital Formation	(4) Output	(5) Net Value Added	(6) Workers
State AA Shock, 5pc	-.0050	-.0092	-.0006	-.0013	-.0021	-.0041
	(.0071)	(.0126)	(.0144)	(.0120)	(.0131)	(.0114)
High Capital*AA Shock, 5pc	-.0153*	-.0376**	-.0378**	-.0389**	-.0486***	-.0274**
	(.0090)	(.0154)	(.0177)	(.0152)	(.0163)	(.0124)
Observations	7158	7156	6457	7153	6998	7155
R^2	.76	.74	.66	.70	.66	.68
Joint Significance - 10pc	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	330	270333	61766	677943	114720	17320

Standard errors in parentheses (clustered at the state-electoral level). The dependent variable in each specification is logged. Estimation controls for the total number of close contests; rate of literacy; state-level voter turnout; effective number of parties contesting; demographic characterstics such as urbanization, gender ratio, population density, and percent of workers. State, industry-year and year fixed effects are also included.

1.10.2 Robustness Checks

Table 1.15: AA Party Representation Shock and State Government Expenditures - Altering Threshold for Close Elections

	(1) Low Caste Spending	(2) Rural Dev and Social Welfare	(3) Roads and Power	(4) Low Caste Spending	(5) Rural Dev and Social Welfare	(6) Roads and Power
State AA Shock, 4pc	.0640***	.1764**	-.3648***			
	(.0213)	(.0868)	(.1181)			
State AA Shock, 2pc				.0370	.1499	-.4094*
				(.0386)	(.1404)	(.2130)
Observations	363	363	363	363	363	363
R^2	.90	.77	.54	.89	.77	.53
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Pct)	3.10	13.40	8.42	3.10	13.40	8.42

Standard errors in parentheses (clustered at the state-electoral cycle level). The dependent variable in each specification is the share of each expenditure category in total state expenditures. The unit of observation is state-year. Estimation controls for total number of close contests; state-level political competition; voteshares for AA parties and two major national parties; share of constituencies reserved for minority candidates; share of population who are literate; share of workers; rate of urbanization; constant net state domestic product; as well as state and year fixed effects.

Table 1.16: AA Party Representation Shock and District Public Employment for Low Caste Workers - Alternate Definitions of Close Elections

	(1) Public Jobs Margin All	(2) Public Jobs Margin Urban	(3) Public Jobs Margin Rural	(4) Public Jobs Margin All	(5) Public Jobs Margin Urban	(6) Public Jobs Margin Rural
District AA Shock, 4pc	.0454	.1167*	.0109			
	(.0406)	(.0610)	(.0483)			
District AA Shock, 2pc				.0099	.1041	-.0284
				(.0467)	(.0728)	(.0715)
Observations	1024	1001	1010	1024	1001	1010
R^2	.84	.81	.75	.84	.81	.75
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	73834	35373	39049	73834	35373	39049

Standard errors in parentheses (clustered at the district level). The dependent variable in each specification is the logged number of low caste workers employed in the public sector. The unit of observation is district-survey round. All estimation controls for the low caste working age population and total number of close contests. Estimations also control for district-level political competition; voteshares for AA and the two major national parties - BJP and INC; share of low caste and Muslim population; share of low caste workers with at least a secondary education; average age of low caste workers; gender ratio; rate of unemployment; and rate of urbanization. All specifications also control for district, survey round and state-survey round fixed effects. The results are weighted by the number of households in each district.

Table 1.17: AA Party Representation Shock and Manufacturing Investment: Alternate Definitions of Close Elections

Panel A:
Close Election Threshold 4 Percent

	(1) Factories	(2) Fixed Capital	(3) Capital Formation	(4) Output	(5) Value Added	(6) Workers
State AA Shock, 4pc	-.0054	-.0195	-.0090	-.0128	-.0185	-.0068
	(.0083)	(.0124)	(.0131)	(.0111)	(.0133)	(.0128)
Observations	7158	7156	6457	7153	6998	7155
R^2	.76	.70	.63	.69	.64	.68
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Panel B:
Close Election Threshold: 2 Percent

	(1) Factories	(2) Fixed Capital	(3) Capital Formation	(4) Output	(5) Value Added	(6) Workers
State AA Shock, 2pc	-.0288*	-.0485*	-.0500*	-.0391	-.0548**	-.0353
	(.0171)	(.0267)	(.0289)	(.0250)	(.0268)	(.0250)
Observations	7158	7156	6457	7153	6998	7155
R^2	.76	.70	.63	.69	.64	.68
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses (clustered at the state electoral-cylce level). The dependent variable in each specification is logged. Estimation controls for the total number of close contests; rates of literacy, urbanization and the share of workers; state-level voter turnout; effective number of parties contesting. State, industry-year and year fixed effects are also included in each regression.

Table 1.18: AA Party Representation Shock and State Government Expenditures - Only Caste-Based AA Parties

	(1) Low Caste Spending	(2) Rural Dev and Social Welfare	(3) Roads and Power Investment
State AA Shock, 5pc	.0807***	.0630	-.2194*
	(.0270)	(.1070)	(.1256)
Observations	300	300	300
R^2	.91	.77	.57
Controls	Yes	Yes	Yes
Dependent Variable Mean (Pct)	3.10	13.40	8.42

Standard errors in parentheses (clustered at the state-electoral cycle level). The dependent variable in each specification is the share of each expenditure category in total state expenditures. The unit of observation is state-year. Estimation controls for total number of close contests; state-level political competition; voteshares for AA parties and two major national parties; share of constituencies reserved for minority candidates; share of population who are literate; share of workers; rate of urbanization; constant net state domestic product; as well as state and year fixed effects.

Table 1.19: AA Party Representation Shock and District Public Employment for Low Caste Workers - Only Caste-Based AA Parties

	(1) Public Jobs All	(2) Public Jobs Urban	(3) Public Jobs Rural
District AA Shock, 5pc	.0274	.1387**	-.0619
	(.0381)	(.0658)	(.0581)
Observations	900	878	886
R^2	.82	.81	.70
Covariates	Yes	Yes	Yes
Dependent Variable Mean	65521	33960	32125

Standard errors in parentheses (clustered at the district level). The dependent variable in each specification is the logged number of low caste workers employed in the public sector. The unit of observation is district-survey round. All estimation controls for the low caste working age population and total number of close contests. Estimations in columns (4)-(6) also control for district-level political competition; voteshares for AA and the two major national parties - BJP and INC; share of low caste and Muslim population; share of low caste workers with at least a secondary education; average age of low caste workers; gender ratio; rate of unemployment; and rate of urbanization. All specifications also control for district, survey round and state-survey round fixed effects. The results are weighted by the number of households in each district.

Table 1.20: AA Party Representation Shock and Manufacturing Industries: Only Caste-Based AA Parties

	(1)	(2)	(3)	(4)	(5)	(6)
	Factories	Fixed Capital	Capital Formation	Output	Net Value Added	Workers
State AA Shock, 5pc	.0003	-.0160	-.0095	-.0096	-.0169	.0026
	(.0076)	(.0131)	(.0139)	(.0114)	(.0114)	(.0090)
Observations	5872	5870	5290	5867	5723	5869
R^2	.77	.70	.65	.70	.65	.68
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	330	270333	61766	677943	114720	17320

Standard errors in parentheses (clustered at the state-electoral level). The dependent variable in each specification is logged. Estimation controls for the total number of close contests; rate of literacy; state-level voter turnout; effective number of parties contesting; demographic characterstics such as urbanization, gender ratio, population density, and percent of workers. State, industry-year and year fixed effects are also included.

Table 1.21: AA Party Representation Shock and State Government Expenditures - Clustering Results at State-Level

	(1) Low Caste Spending	(2) Rural Dev and Social Welfare	(3) Roads and Power Investment
State AA Shock, 5pc	.0617**	.1499	-.2708*
	(.0264)	(.1020)	(.1552)
Observations	363	363	363
R^2	.90	.77	.54
Controls	Yes	Yes	Yes
Dependent Variable Mean (Pct)	3.10	13.40	8.42

Standard errors in parentheses (clustered at the state level). The dependent variable in each specification is the share of each expenditure category in total state expenditures. The unit of observation is state-year. Estimation controls for total number of close contests; state-level political competition; voteshares for AA parties and two major national parties; share of constituencies reserved for minority candidates; share of population who are literate; share of workers; rate of urbanization; constant net state domestic product; as well as state and year fixed effects.

Table 1.22: AA Party Representation Shock and Manufacturing Industries: State-Level Clusters

	(1) Factories	(2) Fixed Capital	(3) Capital Formation	(4) Output	(5) Net Value Added	(6) Workers
State AA Shock, 5pc	-.0127*	-.0327***	-.0249**	-.0233***	-.0288***	-.0183**
	(.0062)	(.0077)	(.0088)	(.0070)	(.0072)	(.0069)
Observations	7158	7156	6457	7153	6998	7155
R^2	.76	.70	.63	.69	.64	.68
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	330	270333	61766	677943	114720	17320

Standard errors in parentheses (clustered at the state level). The dependent variable in each specification is logged. Estimation controls for the total number of close contests; rate of literacy; state-level voter turnout; effective number of parties contesting; demographic characterstics such as urbanization, gender ratio, population density, and percent of workers. State, industry-year and year fixed effects are also included.

Table 1.23: AA Party Representation Shock and District Public Employment for Low Caste Workers - States Where AA Parties Don't Form Coalitions

	(1) Public Jobs All	(2) Public Jobs Urban	(3) Public Jobs Rural
District AA Shock, 5pc	.0485	.0969*	.0458
	(.0383)	(.0578)	(.0429)
Observations	249	243	247
R^2	.81	.77	.78
Covariates	Yes	Yes	Yes
Dependent Variable Mean	54300	22797	31597

Standard errors in parentheses (clustered at the district level). The dependent variable in each specification is the logged number of low caste workers employed in the public sector. The unit of observation is district-survey round. All estimation controls for the low caste working age population and total number of close contests. Estimations in columns (4)-(6) also control for district-level political competition; voteshares for AA and the two major national parties - BJP and INC; share of low caste and Muslim population; share of low caste workers with at least a secondary education; average age of low caste workers; gender ratio; rate of unemployment; and rate of urbanization. All specifications also control for district, survey round and state-survey round fixed effects. The results are weighted by the number of households in each district.

Table 1.24: AA Party Representation Shock and District Public Employment for Low Caste Workers - Only Districts with Close Elections

	(1) Public Jobs All	(2) Public Jobs Urban	(3) Public Jobs Rural
District AA Shock, 5pc	.0469	.0998**	.0126
	(.0306)	(.0441)	(.0369)
Observations	750	737	742
R^2	.84	.82	.76
Covariates	Yes	Yes	Yes
Dependent Variable Mean	77230	35044	42959

Standard errors in parentheses (clustered at the district level). The dependent variable in each specification is the logged number of low caste workers employed in the public sector. The unit of observation is district-survey round. All estimation controls for the low caste working age population and total number of close contests. Estimations in columns (4)-(6) also control for district-level political competition; voteshares for AA and the two major national parties - BJP and INC; share of low caste and Muslim population; share of low caste workers with at least a secondary education; average age of low caste workers; gender ratio; rate of unemployment; and rate of urbanization. All specifications also control for district, survey round and state-survey round fixed effects. The results are weighted by the number of households in each district.

Table 1.25: AA Party Representation Shock and District Public Employment for Low Caste Workers - Unweighted Estimates

	(1) Public Jobs All	(2) Public Jobs Urban	(3) Public Jobs Rural
District AA Shock, 5pc	.0370	.0926*	-.0224
	(.0424)	(.0539)	(.0600)
Observations	1024	1001	1010
R^2	.78	.76	.69
Covariates	Yes	Yes	Yes
Dependent Variable Mean	44212	19669	24899

Standard errors in parentheses (clustered at the district level). The dependent variable in each specification is the logged number of low caste workers employed in the public sector. The unit of observation is district-survey round. All estimation controls for the low caste working age population and total number of close contests. Estimations in columns (4)-(6) also control for district-level political competition; voteshares for AA and the two major national parties - BJP and INC; share of low caste and Muslim population; share of low caste workers with at least a secondary education; average age of low caste workers; gender ratio; rate of unemployment; and rate of urbanization. All specifications also control for district, survey round and state-survey round fixed effects.

1.11 Appendix

1.11.1 Summary Statistics

Table A1.1: Summary Statistics of Key State-Level Independent Variables

Variable	Mean	Std. Dev.	Min.	Max.	N
State AA Shock, 5pc	-0.09	3.27	-11.5	14	398
Total Close Elections, 5pc	24.79	29.48	0	116	398
State AA Shock, 4pc	-0.11	2.47	-5.5	10.5	398
Total Close Elections, 4pc	19.89	24.09	0	97	398
State AA Shock, 2pc	0.05	1.32	-4	5.5	398
Total Close Elections, 2pc	9.91	12.8	0	55	398
State AA Party Vote Share	0.23	0.18	0.02	0.67	398
State Congress Vote Share	0.29	0.11	0.02	0.48	398
State BJP Vote Share	0.19	0.14	0	0.5	398
State Voter Turnout (Share)	0.65	0.09	0.23	0.85	398
Share of Reserved Constituencies	0.24	0.09	0.1	0.49	398
Effective Number of Parties	3.11	0.66	2.22	4.85	398
Share Literate	0.66	0.12	0.37	0.94	402
Share Workers	0.4	0.05	0.3	0.52	402
Share Urban	0.27	0.1	0.09	0.51	402

Table A1.2: Summary Statistics of Key Outcome Variables: State Expenditures

Variable	Mean	Std. Dev.	Min.	Max.	N
Targeted Low Caste Welfare (Percent)	3	2	0	13	402
Untargeted Low Caste Welfare (Percent)	13	5	3	35	402
Physical Infrastructure (Percent)	8	4	1	34	402
Health and Education (Percent)	30	6	17	44	402
Agriculture and Irrigation (Percent)	17	6	2	33	402
Pension and Wages (Percent)	12	5	1	32	402
Per Capita Targeted Low Caste Welfare (2012 Rupees)	149	128	7	837	402
Per Capita Untargeted Low Caste Welfare (2012 Rupees)	617	439	69	3719	402
Per Capita Physical Infrastructure (2012 Rupees)	405	295	23	3118	402
Per Capita Health and Education (2012 Rupees)	1376	679	120	6157	402
Per Capita Agriculture and Irrigation (2012 Rupees)	868	658	46	4358	402
Per Capita Pension and Wages (2012 Rupees)	555	393	8	3127	402

Table A1.3: Summary Statistics of Key Outcome Variables: Private Investment, 2 Digit Industry

Variable	Mean	Std. Dev.	Min.	Max.	N
Total Factories	330	601	0	8785	7183
Total Workers	17321	33241	0	492465	7183
Fixed Capital - 2012 Rupees	270359	670493	0	12614964	7183
Capital Formation - 2012 Rupees	61772	205173	-7753533	4703293	7183
Output - 2012 Rupees	678019	1455773	0	34494348	7183
Net Value Added - 2012 Rupees	114476	265278	-241537	5384645	7183

Table A1.4: Summary Statistics of District-Level Political Variables

Variable	Mean	Std. Dev.	Min.	Max.	N
District AA Shock, 5pc	-0.02	0.74	-5	7	1038
Total Close Elections, 5pc	1.76	4.23	0	102	1038
District AA Shock, 4pc	-0.01	0.61	-3.5	4.5	1038
Total Close Elections, 4pc	1.39	3.23	0	75	1038
District AA Shock, 2pc	0	0.41	-2	3	1038
Total Close Elections, 2pc	0.71	1.73	0	36	1038
District AA Party Vote Share	0.24	0.21	0	0.82	1038
District BJP Vote Share	0.23	0.15	0	0.6	1038
District INC Vote Share	0.27	0.14	0	0.62	1038
District Voter Turnout (Share)	0.65	0.1	0.36	0.89	1038
Effective Number of Parties	3.18	0.74	2.07	6.83	1038

Table A1.5: Summary Statistics of District-Level Socio-Economic Variables

Variable	Mean	Std. Dev.	Min.	Max.	N
Share Low Caste	0.69	0.19	0.12	1	1054
Share Low Caste, Rural	0.75	0.19	0.08	1	1050
Share Low Caste, Urban	0.58	0.2	0	1	1041
Share Public Sector	0.04	0.02	0	0.23	1054
Share Public Sector, Urban	0.07	0.04	0	0.38	1041
Share Public Sector, Rural	0.03	0.02	0	0.25	1050
Share Public Sector, Low Caste	0.03	0.02	0	0.25	1054
Share Public Sector, Low Caste Urban	0.07	0.05	0	0.57	1040
Share Public Sector, Low Caste Rural	0.02	0.02	0	0.26	1050
Share Public Sector, High Caste	0.06	0.04	0	0.41	1051
Share Public Sector, High Caste Rural	0.04	0.06	0	0.6	1040
Share Public Sector, High Caste Urban	0.1	0.07	0	0.67	1040
Share Public Sector, Low Caste, High Education	0.02	0.02	0	0.15	1054
Share Public Sector, High Caste, High Education	0.06	0.04	0	0.35	1051
Share of Low Caste, Secondary Educated	0.37	0.15	0.05	0.83	1054
Share of High Caste, Secondary Educated	0.66	0.15	0.08	1	1051
Headcount Ratio	0.25	0.15	0	0.81	1054
Headcount Ratio, Rural	0.26	0.16	0	0.89	1050
Headcount Ratio, Urban	0.23	0.15	0	0.76	1041
Headcount Ratio, Low Caste	0.28	0.16	0	0.94	1054
Headcount Ratio, High Caste	0.12	0.12	0	0.8	1051
District Urban Share	0.27	0.19	0	1	1054
District Share Female	0.5	0.03	0.4	0.66	1054

1.11.2 Additional Results - OLS Estimates

AA Party Representation Shock and Economic Outcomes: OLS Estimates

Table A1.6: AA Party Victories and State Government Expenditures

	(1) Low Caste Spending	(2) Rural Dev and Social Welfare	(3) Roads and Power Investment
Total AA Party Wins, State	.0032	.0155*	-.0178*
	(.0024)	(.0090)	(.0107)
Observations	363	363	363
R^2	.89	.77	.54
Controls	Yes	Yes	Yes
Dependent Variable Mean (Pct)	3.10	13.40	8.42

Standard errors in parentheses (clustered at the state-electoral cycle level). The dependent variable in each specification is the share of each expenditure category in total state expenditures. The unit of observation is state-year. Estimation controls for total number of close contests; state-level political competition; voteshares for AA parties and two major national parties; share of constituencies reserved for minority candidates; share of population who are literate; share of workers; rate of urbanization; constant net state domestic product; as well as state and year fixed effects. All regressions are weighted by the state's population.

Table A1.7: AA Party Electoral Success and Manufacturing Industries - OLS Results

	(1) Factories	(2) Fixed Capital	(3) Capital Formation	(4) Output	(5) Net Value Added	(6) Workers
AA Party Wins, State	-.0001	-.0015	-.0020**	-.0007	-.0012	-.0004
	(.0004)	(.0010)	(.0008)	(.0007)	(.0008)	(.0005)
Observations	7158	7156	6457	7153	6998	7155
R^2	.76	.70	.64	.69	.64	.68
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	330	270333	61766	677943	114720	17320

Standard errors in parentheses (clustered at the state-electoral level). The dependent variable in each specification is logged. Estimation controls for the total number of close contests; rate of literacy; state-level voter turnout; effective number of parties contesting; demographic characterstics such as urbanization, gender ratio, population density, and percent of workers. State, industry-year and year fixed effects are also included.

Table A1.8: AA Party Success and District Public Employment for Low Caste Workers - OLS Estimates

	(1) Public Jobs All	(2) Public Jobs Urban	(3) Public Jobs Rural	(4) Public Jobs All	(5) Public Jobs Urban	(6) Public Jobs Rural
Total AA Wins - District	.0039	.0231*	-.0006	.0023	.0263	.0042
	(.0106)	(.0130)	(.0154)	(.0105)	(.0168)	(.0157)
Observations	1029	1016	1025	1024	1001	1010
R^2	.82	.78	.73	.84	.81	.75
Covariates	No	No	No	Yes	Yes	Yes
Dependent Variable Mean (Workers)	73834	35373	39049	73834	35373	39049

Standard errors in parentheses (clustered at the district level). The dependent variable in each specification is the logged number of low caste workers employed in the public sector. The unit of observation is district-survey round. All estimation controls for the low caste working age population, total number of close contests and total number of electoral constituencies in the district. Estimations in columns (4)-(6) also control for district-level political competition; voteshares for AA and the two major national parties - BJP and INC; share of low caste and Muslim population; share of low caste workers with at least a secondary education; average age of low caste workers; gender ratio; rate of unemployment; and rate of urbanization. All specifications also control for district, survey round and state-survey round fixed effects. The results are weighted by the number of households in each district.

Population Weighted Estimates

Table A1.9: AA Party Representation Shock and State Government Expenditures - Estimates Weighted by State Population

	(1) Low Caste Spending	(2) Rural Dev and Social Welfare	(3) Roads and Power Investment
State AA Shock, 5pc	.0502**	.1140	-.3623***
	(.0215)	(.0846)	(.1065)
Observations	363	363	363
R^2	.89	.70	.48
Controls	Yes	Yes	Yes
Dependent Variable Mean (Pct)	3.10	13.40	8.42

Standard errors in parentheses (clustered at the state-electoral cycle level). The dependent variable in each specification is the share of each expenditure category in total state expenditures. The unit of observation is state-year. Estimation controls for total number of close contests; state-level political competition; voteshares for AA parties and two major national parties; share of constituencies reserved for minority candidates; share of population who are literate; share of workers; rate of urbanization; constant net state domestic product; as well as state and year fixed effects. All regressions are weighted by the state's population.

Table A1.10: AA Party Representation Shock and Manufacturing Industries - Population Weighted Estimates

	(1) Factories	(2) Fixed Capital	(3) Capital Formation	(4) Output	(5) Workers
State AA Shock, 5pc	-.0101**	-.0259***	-.0194**	-.0182**	-.0170**
	(.0049)	(.0089)	(.0089)	(.0074)	(.0075)
Observations	7158	7156	6457	7153	7155
R^2	.78	.74	.67	.73	.71
Controls	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	330	270333	61766	677943	17320

Standard errors in parentheses (clustered at the state-electoral level). The dependent variable in each specification is logged. Estimation controls for the total number of close contests; rate of literacy; state-level voter turnout; effective number of parties contesting; demographic characterstics such as urbanization, gender ratio, population density, and percent of workers. State, industry-year and year fixed effects are also included.

AA Party Legislators and Per Capita State Expenditures

Table A1.11: AA Party Representation Shock and Per Capita State Government Expenditures

	(1) Low Caste Spending	(2) Rural Dev and Social Welfare	(3) Roads and Power Investment
State AA Shock, 5pc	.0130*	.0109*	-.0284**
	(.0072)	(.0061)	(.0120)
Observations	363	363	363
R^2	.93	.92	.85
Controls	Yes	Yes	Yes
Dependent Variable Mean (Rupees)	149.42	616.96	404.91

Standard errors in parentheses (clustered at the state-electoral cycle level). The dependent variable in each specification is logged per capita expenditures for each category in 2012 rupees. The unit of observation is state-year. Estimation controls for total number of close contests; state-level political competition; voteshares for AA parties and two major national parties; share of constituencies reserved for minority candidates; share of population who are literate; share of workers; rate of urbanization; constant net state domestic product; total per capita expenditures in 2012 rupees; as well as state and year fixed effects.

1.11.3 Summary Statistics of Constituencies Based on Close Elections at 5 Percent Margin

Table A1.12: Summary Statistics of Constituency-Level Variables: Constituencies with non-AA vs AA Close Elections at 5 Percent Margin

Variable	Mean	Std. Dev.	Min.	Max.	N
AA Party Vote Share	0.436	0.09	0.142	0.704	2401
Congress Party Vote Share	0.217	0.178	0	0.516	2401
BJP Vote Share	0.129	0.134	0	0.485	2401
Number of Registered Voters	178043.614	48682.828	39638	594868	2401
Number of Contestants	11.66	7.274	2	53	2401
Effective Number of Parties Contesting	3.536	1.297	1.999	11.956	2401
Voter Turnout	66.407	13.323	4.046	96.127	2401
Elections with Male Winner	0.945	0.228	0	1	2401
Constituency Reserved for SC/ST	0.203	0.402	0	1	2401

Table A1.13: Summary Statistics of Constituency-Level Variables: Constituencies with AA vs AA Close Elections at 5 Percent Margin

Variable	Mean	Std. Dev.	Min.	Max.	N
AA Party Vote Share	0.703	0.148	0.385	0.987	386
Congress Party Vote Share	0.061	0.069	0	0.308	386
BJP Vote Share	0.109	0.107	0	0.334	386
Number of Registered Voters	227368.754	54156.263	82178	649763	386
Number of Contestants	13.51	5.736	3	42	386
Effective Number of Parties Contesting	4.074	1.127	2.096	10.761	386
Voter Turnout	54.078	9.165	32.589	85.883	386
Elections with Male Winner	0.930	0.255	0	1	386
Constituency Reserved for SC/ST	0.197	0.398	0	1	386

Table A1.14: Summary Statistics of Constituency-Level Variables: Constituencies with non-AA vs non-AA Close Elections at 5 Percent Margin

Variable	Mean	Std. Dev.	Min.	Max.	N
AA Party Vote Share	0.079	0.1	0	0.542	2702
Congress Party Vote Share	0.332	0.139	0	0.521	2702
BJP Vote Share	0.227	0.185	0	0.518	2702
Number of Registered Voters	159634.271	55801.117	20645	890784	2702
Number of Contestants	10.811	6.321	2	53	2702
Effective Number of Parties Contesting	3.436	1.302	1.997	13.204	2702
Voter Turnout	66.954	10.464	8.768	96.127	2702
Elections with Male Winner	0.938	0.241	0	1	2702
Constituency Reserved for SC/ST	0.229	0.421	0	1	2702

Chapter 2

Minority Representation and Protection from Racial Violence: The Impact of Caste-based Political Parties in India

"The majority not only makes the laws, but can break them as well." - Alexis de Tocqueville

2.1 Introduction

Proponents of liberal democracies have long argued that governance via majority rule is no more important than the protection of the basic rights of resident minorities.[1] Yet, even the most successful democracies have at times struggled to protect the civil, economic, and political rights of racial and ethnic minority groups. Targeted criminal violence is among the most egregious of the basic civil rights violations suffered by minorities.

Many democracies have struggled to protect minorities from this extreme form of majority oppression. In nineteenth and early-twentieth century America, racial violence took the form of white lynch mobs killing blacks throughout the South; nearly 4,000 "racial terror lynchings" took place between 1870 and 1950. Racial violence is also a problem today, though. A recent report suggests that more than half of EU member states witnessed an increase in racially-motivated crimes at the end of the last decade, and that "visible minorities, refugees, asylum seekers, and Roma appear to suffer the brunt of abusive treatment"(Nyiri and English 2007). Yet, despite the normative significance of racial violence in multi-ethnic democracies, there is virtually no empirical research on what factors are causally linked to a minority groups' right to be free from bodily harm.

In this paper, we provide the first causal evidence on one channel explaining the incidence

[1] To provide one example, James Madison warned on the eve signing the Constitution that the US must seek to protect both popular sovereignty and minority rights: "it is of great importance in a republic, not only to guard the society against the oppression of its rules, but to guard one part of the society against the injustice of the other part."

of violence targeted toward racial minorities. We assess whether political representation for the historically disenfranchised affects the incidence of discriminatory criminal violence they suffer, using India as a case study. Specifically, we study how minority representation affects crimes committed against India's historically disadvantaged Scheduled Castes (SCs) and Scheduled Tribes (STs) (i.e., former "untouchables" and marginalized indigenous tribes). These two ethnic groups are India's most disadvantaged minority subpopulations, having suffered centuries of economic and social exclusion by upper caste groups. Despite numerous attempts to remedy caste-based discrimination at the both the federal and state level, India still struggles to protect minority subpopulations from abuse on the basis of caste. According to national crime statistics, for example, around four low-caste women are raped every day by upper-caste criminals (Krishnan 2014). [2]

In principle, an increase in minority representation in government could lower or raise the incidence of racially-targeted crimes. Representation of historically marginalized minorities may alter government responsiveness to reports of discrimination or abuse, and may inspire greater minority confidence to report crimes to the police. These forces should deter potential offenders from committing crimes against minorities.[3] On the other hand, to the extent that political success of minority groups is perceived as a usurping of political power or an unacceptable shift in the socio-political status quo, power in the hands of minority politicians may engender resentment within racial/ethnic majority communities. Recent evidence in the United States, for example, suggests that racial animus toward minorities among whites increased during the 2008 election of President Barack Obama (Stephens-Davidowitz 2014; Mas and Moretti 2009). To the extent that this racial hostility translates to actions, one may observe an increase in crimes tied to ethnic/racial identity. Moreover, if minority politicians lack the political clout, capital, or skill to effectuate policy change, they may be unable to protect their constituents from criminal harms.

Whether minority representation improves state protection from criminal violence is thus an empirical question. To shed light on this relationship, we collect and analyze unique (and underexplored) statistics crimes for which the victim is an SC/ST individual, but the perpetrator is an upper caste individual. To assess the role of minority political representation, we examine the impact of caste-based political parties that emerged across India during the 1980s. Representatives from these parties had the explicit policy agendas of improving the socioeconomic livelihood of SC/ST constituents once in office. The empirical challenge we face is that locations that favor low-caste political parties may differ along many dimensions from locations that favor mainstream political parties, such as the BJP or Indian National Congress (INC). To overcome these challenges related to correlated unobservables or omitted variable bias, we use a variant of a regression-discontinuity design (RDD) framework. We exploit close elections between caste-party and traditional-party candidates, where the underlying assumption is that a constituency where the low caste candidate barely won is similar to a constituency where he/she barely lost on all unobserved characteristics that are correlated with the dependent variable.

[2] This is likely an underestimate since rapes are often under-reported across India due to stigma; the lack of a political voice also reduces reporting of crimes within SC/ST subpopulations.

[3] Although, any deterrence induced by changing willingness to report crimes may still appear to the researcher as a crime increase or null effect, depending on the size of the deterrent effect.

Traditional election RDD studies tie characteristics of winning candidates directly to the geographic units they represent in order to make causal statements about politicians or particular characteristics (such as race, gender, etc.). Because our outcomes of interest are measured and determined at a more aggregate level than the local-level elections we study, we exploit quasi-exogenous variation in the *total number* of close elections won by candidates from low-caste political parties. This allows us to assess whether states with larger-than-expected low-caste delegations experience a subsequent decrease or increase in crimes against persons belonging to low castes/tribes (or whether the reverse is true).

Our results suggest that (organized) political representation for ethnic minorities produces beneficial effects for these constituents. Political representation in local legislatures for low-caste groups produces a statistically significant and economically meaningful reduction in rates of violent crimes committed against these subpopulations. We observe decreases in the state-level rates of murder, rape, kidnapping, and assault–typically of about 4-14% for the marginal legislator. Moreover, placebo tests rule out that these results are artifacts of secular crime trends: low-caste political parties have no effect on overall crime rates, suggesting that politicians are privileging the protection of their constituents. Moreover, increasing political power in the hands of India's mainstream parties–the BJP and Congress (INC)–have no effect on caste-targeted crime levels.

We also consider the channels of influence through which the government representatives of minority subpopulations can affect targeted crimes. We show that minority representation may be affecting caste-targeted crime by changing the state's criminal justice institutions in a manner that prioritizes the protection of minority constituents. For example, states with greater low-caste representation arrest the perpetrators of targeted violence at higher rates, and prosecute these types of cases with greater vigor (less prosecutor withdrawal of caste-related criminal cases). We also identify changes in the personnel of law enforcement agencies in a manner consistent with increased protection minority subpopulations (more low-caste police officers). Finally, to rule out concerns of underreporting of crimes, we also document improving attitudes of low caste and indigenous tribal citizens towards state institutions–in particular towards politicians, police agencies, and courts.

Taken collectively, the results suggest that minority politicians can better protect the rights and dignity of minority citizens. Our results are thus broadly consistent with the deterrence of potential caste-crime perpetrators due to increased representation. Minority politicians can affect the enforcement priorities of government, as well as change attitudes of citizens in a manner consistent with reduced racially-motivated crime.

2.1.1 Contributions: Representation and Minority-targeted Crime in India

We contribute to two main streams of literature in undertaking this analysis. First, we contribute to recent work on the factors that affect crimes suffered by vulnerable populations. While caste-based discrimination has been explored in the context of labor markets and access to public resources,[4] the causes and consequences of crimes targeted toward SC and

[4] See Deshpande (2011) and Thorat and Newman (2010) for a review of relevant findings.

ST groups remains relatively under-studied. Anecdotal evidence suggests these groups suffer from high levels of violence, with the primary perpetrators being members of dominant castes (United Nations 2012).[5] Given the continued prevalence of this social problem, an understanding of the driving factors is warranted.

Existing work on the problem of low-caste criminal victimization has examined the association between crime levels and factors including the commonality of water sources with non-SC/ST groups (Bros and Couttenier 2015) and the ratio of expenditures dedicated to lower versus upper castes (Sharma 2015). Our study, however, is the first to provide evidence of a causal channel related to caste-based violence: substantive representation in government.[6]

Relatedly, we consider our findings relevant to a small group of studies that seeks to understand the general problem of ethnic/racial bias in criminal victimization. This literature primarily focuses on hate crime incidence in the United States. Studies suggest that economic factors such as wages, unemployment, and poverty, as well as social factors such as abuse rates and law enforcement activity, are correlated with identity-based crimes (Medoff 1999; Gale, Heath, and Ressler, 2002; Ryan and Leeson 2010). However, our study is the first to our knowledge that considers how minority representation in the legislature is causally related to violence against historically low-status racial or ethnic groups. In this regard, our study is most similar in spirit to Iyer, Mani, Mishra, and Topalova (2012), who explore how female political representation can impact the reporting of gender-related crimes. The authors of that study find that affirmative action for female political representatives results in a considerable increase in reported crimes against women. The authors argue that the observed increase in crimes against women is socially beneficial: the increase is due to better reporting of crime rather than to actual higher crime incidence. Our study differs in two major regards. First, we study benefits emanating from political party mobilization rather than mandated representation. Second, our analysis does not requires us to distinguish between reported and measured crimes: since reporting bias is likely positive, the measured declines in low caste crimes we report are likely underestimates of the impact of low caste representation.

Finally, we also contribute to the study of substantive political representation for disempowered caste and tribal populations in India. Our approach of studying low-caste political movements contrasts with many past studies of low-caste representation, which have focused on the impact of mandated set-asides for ethnic minorities in government (via India's "reservation system").[7] We are the first, however, to exploit the exogenous changes in organized caste-based representation provided to ethnic minorities arising from endogenous

[5] Upper castes are believed to commit such acts in order to preserve the existing social and economic order (**?**).

[6] Our study is also related to a small but growing economic literature on crime in India more generally. For instance, in one recent paper, Prasad (2012) finds that Indian economic liberalization reforms reduced levels of murder.

[7] Evidence suggests that this program of affirmative action in government has improved low-caste welfare modestly along dimensions such as poverty, redistribution of state resources, and public goods access (Besley, Pande, Rahman, and Rao 2004; Chin and Prakash 2011; Pande 2003). Evidence suggests that these groups continue to suffer from economic and social exclusion due to a combination of poor state protection and stigmatized identity (Deshpande 2011; Thorat and Newman 2010).

political mobilization in India. This focus distinguishes our paper from Iyer, Mani, Mishra, and Topalova (2012) in a fundamental way, as our paper focuses on nationwide endogenous political mobilization rather than Constitutionally mandated set-asides.

Having situated our study, the remainder of the paper will proceed as follows: we describe the context of our study, describing the formation of political parties dedicated to low caste political interests. We then discuss conceptually the various mechanisms through which minority political representation may affect minority crime rates. We then explain the methodology for the study, including descriptions of data sources and the empirical strategy. We then describe the main results, as well as results from a series of checks to test robustness, and finally, we will provide summary remarks to conclude.

2.2 Background: India's Caste System and the Rise of Low-caste Political Parties

India's population has for centuries been segmented socially (and *de facto* economically) into endogamous groups called *jatis*, or castes, which originated from ancient Hindu doctrine. At the bottom of the caste group ladder are the so-called "Untouchables"[8] or *Dalits*. The *Dalits* of the caste hierarchy have historically been relegated to daily tasks considered "impure" by the rest of the Indian society, such as cleaning streets and disposing of carcasses and human waste. On the same socioeconomic level as the *Dalits* are India's indigenous populations, or *Adivasis*. These tribal groups face large-scale exclusion from mainstream society due to geographical isolation, primitive agricultural practices, and distinctly non-mainstream customs.

In addition to exclusionary socioeconomic status, *Dalits* and *Adivasis*–the groups that comprise India's historically marginalized communities (HMCs)–have long suffered from other forms of overt discrimination at the hands of dominant caste elites. For decades, these subpopulations were forced to live in housing that was segregated from the rest of society, were denied access to schools, drinking water and other basic resources, and were unable to enter places of worship attended by upper castes.

Following independence from Great Britain, the nascent Indian state sought to remedy the centuries-old persecution of Dalits and Adivasis through a broad program of affirmative action. The Indian Constitution created the "reservation system," or a large-scale system of set-asides for low caste individuals within state legislatures and government jobs. "Scheduled Castes" (SC) and "Scheduled Tribes"(ST) are the Dalit and *Adivasi* groups, respectively, that are designated for federal support under the reservation system.[9] Both SCs and STs collectively constitute over a third of the government-defined HMCs that are eligible for

[8] There are four primary caste groups. At the top of the four-tiered caste hierarchy are the *Brahmins* (priests and teachers), followed by *Kshatriyas* (warriors and members of royalty). Next are the *Vaishyas* (traders, merchants and moneylenders), and finally the *Shudras* (menial labor and typically low-end jobs).

[9] In addition, there is a third category known as the "Other Backward Classes" (OBCs), to which reservations have been extended since the early 1990s. This group, comprised primarily of the "Shudra" *jatis*, while not burdened with the stigma of "untouchability," was considered socially and educationally "backward," and so worthy of government assistance through affirmative action.

benefits under the affirmative action schemes.[10]

A rich literature has emerged in the past decade on the economic effects of the India's system of government affirmative action. The seminal work in this areas comes from Pande (2003), who documents that increasing minority representation via minority set-asides leads to increased spending on ST welfare programs, but had liittle effect of SC welfare spending. This finding is potentially problematic, since SCs constitute a much larger share of India's minority population. Bardhan et al. (2010), using data from West Bengal, find that households in villages with an SC/ST local government leader tend to receive more benefits from the local government, particularly for housing and toilet construction and improvements.

One critical area of low-caste politics, however, remains relatively unexplored by development scholars in economics and political science: political party mobilization focused exclusively on the welfare of Scheduled Castes and Scheduled Tribes. Beginning in the late 1980s, political parties emerged within low-caste communities in response to both continued discrimination at the hands of forward-caste elites, as well as the perceived failure of the reservation system and mainstream political parties to meet the needs of low-caste communities (Jaffrelot 2006).[11] A large number of small, regional caste-based parties coalesced under the auspices of two broad national parties: the Janata Dal (JD) and the Bahujan Samaj Party (BSP). The goal of both national parties was to improve the socioeconomic status of OBC, SC, and ST communities beyond what had been achieved through the existing affirmative action policies.[12]

The BSP and JD (as well as related splinter parties) have had similar political goals over the last three decades: economic and social empowerment for India's most discriminated-against castes and indigenous groups. The central policy goal in this regard has been increased affirmative action for SC/STs within all public institutions.[13] These parties have fought aggressively for increased numbers of SC/STs in public sector jobs, as well as within higher education (Deshpande 2011). It is for this reason that the JD, BSP and their regional offshoots are referred to as "Affirmative Action parties" ("AA parties") (**?**).[14] Using historical sources to account for the regional fragmentation of the JD and the Left Parties over the past two decades, we classify a total of 46 political parties as "AA Parties" that represent the interests of low-caste constituents.[15]

To our knowledge, our study is the first paper to consider the impact of organized minority political parties in India on socio-economic outcomes of interest. As such, we

[10] SCs, STs and OBCs together constitute between 50-60 percent of India's national population.

[11] Dalits and OBCs were at the forefront of this movement.

[12] In the states of Tamil Nadu, Kerala and West Bengal, the political mobilization of socially and economically deprived groups started much earlier in the 1960s, often along class lines with greater emphasis being placed on large-scale economic deprivation than caste-based discrimination. However, a strong correlation between poverty and caste identity however meant that a number of party workers involved in the mobilization hailed from lower caste backgrounds.

[13] Chandra (2004) notes that in India, the platform of "social justice" is in fact coded language meant to signal support for all forms of affirmative action sought by the parties representing HMCs.

[14] Since the Left and the DMK parties have typically supported the JD and the BSP, and supported similar affirmative action policies for lower castes, we group them together in our classification of AA parties.

[15] A more detailed note on AA parties, their political goals, and their respective coalitions is presented in the Appendix.

also contribute to the existing literature on the importance of political parties. Causal estimates in this research area are both relatively new and useful given that empirical work on the effects of political parties have thus far provided ambiguous and sometimes conflicting results. Study in this area has thus far been primarily confined to understanding the effects of American political parties. In the US, Albouy (2011) finds that the party identity of US Congressional representatives shapes local spending priorities. In contrast, Ferreira and Gyourko (2009) find no effect of party identity on policy outcomes in US mayoral elections. Importantly, however, these studies focus on mainstream political parties in a rigidly two-party system, while our paper focuses for the first time on populist, identity-based political parties in a multi-party democratic setting.

2.3 Conceptual Framework: Potential Effects of Affirmative Action Party Representation on Incidence of Caste-targeted Crime

The causal impact of minority political representation on minority criminal victimization is theoretically ambiguous. We use a framework developed by Donohue and Levitt (2001) and further applied by Iyer, Mani, Mishra, and Topalova (2012) to consider the various points of the causal chain at which political representation can affect crimes committed against low-caste populations. A criminal incident implicates three main parties: the criminal, the victim, and the government justice system. The sequence of events is as follows: first, a potential criminal decides whether to commit a crime (Stage 1). After a criminal act occurs, the victim next decides whether to report the crime to state law enforcement authorities (Stage 2). Finally, representatives of the state (law enforcement officers, judges/prosecutors of the courts, etc.) must decide whether to record the crime, as well as how much effort to devote to investigating the crime and bringing the criminal to justice (Stage 3).

We work backwards to consider how politicians representing minority interests in particular could either increase or decrease the (measured or actual) incidence of crimes suffered by these communities. Starting at Stage 3, the government's decision process will influence the behavior of both victims and criminals. For example, elected officials with a explicit mandate to increase protection of disadvantaged low-caste communities may change how state institutions process alleged minority crimes. Importantly, state governments have significant power to exert over matters of crime control: most criminal justice functions–such as law enforcement, prosecution, sentencing, and punishment–are the responsibility of state governments (Unnithan 2013).

Given a state government's control over the state police force, minority politicians may increase the likelihood that police are more responsive to minority criminal victimization (i.e., by ensuring that the police investigate crimes more diligently).[16] While not elected, the police are subject to political forces given that criminal law and enforcement is under the control of state rather than the federal government. In addition, recent research demonstrates

[16] Senior police officers in India are recruited on a nationwide basis, and are then assigned to specific states by which they are controlled.

other ways that state-level politicians exercise significant control over public officials such as police officers. For example, Iyer and Mani (2012) show that elected representatives influence public bureaucrats' decision-making through their influence on career promotions and the geographic location of bureaucrat postings. Indeed, given that police systems are loosely connected at the federal level, state politicians can punish recalcitrant officers with punitive transfers across states.[17] Politicians can also utilize their bully pulpit to influence law enforcement priorities in a manner that affects crime. For example, police may improve the recording/processing of alleged crimes because state leaders have the ability to bring public attention to poor police behavior (such as sloppy handling of criminal cases) using sources such as news media outlets. Similarly, while the Constitution insulates the judiciary from political pressures, state governments retain the discretionary authority on whether to pursue criminal cases, and also determine the urgency and dilligence with which individual cases are pursued. In this regard, the electoral success of AA parties can influence the speed at which cases are brought to law courts, and the quality of evidence collected in support of the prosecution's case. We term changes in these various state responses that can affect crimes against SC/ST communities an "enforcement effect."[18]

State-level politicians can also have an "enforcement effect" that affects minority-targeted crime by changing police officers' attitudes towards low-caste subpopulations. Beaman, Chattopadhyay, Duflo, Pande, and Topalova (2009), for example, provides evidence that men's attitudes towards women improve in places where women served as government representatives. This intermediary effect of representation may induce police offers to be more responsive toward SC/ST victims, and more vigorous in their prosecution of caste-targeted crimes.

In the same vein, it is also possible that politicians may reduce identity-based violence by changing the composition of the bureaucracy in a manner that privileges the protection of party constituents (a "government composition" effect). A police officer may be more attuned to the incidence of violence suffered by his or her own socio-ethnic group. So, by hiring more SC/ST police officers, state governments may improve enforcement of laws as well as trust in the police.

Politicians' effects on law enforcement and judicial behavior at Stage 3 of a criminal incident can also affect victim decisions at Stage 2 of the incident. Increased representation in government may make minority victims more likely to report crimes. We expect this "reporting effect" if having representatives of SC/ST communities in positions of political power changes victims' expectations about how crimes will be prosecuted by the government. If SC/ST victims believe that the presence of politicians representing their interests will make police officers more likely to respond to their grievances and to pursue perpetrators more vigorously, this higher expected return to criminal offense reporting will lead to increased levels of actual reporting (even if the true number of crimes committed stays level).

An alternative explanation for increased minority victim reporting of crime may be that the presence of elected representatives for SC/ST communities inspires greater self-confidence

[17] The transfer of police officers and bureaucrats are often amongst the first administrative decisions made by AA parties when they come to power at the state level.

[18] Iyer, Mani, Mishra, and Topalova (2012) calls this is the "recording effect," as they focus primarily on how politicians change police recording behavior.

and reduced tolerance for being mistreated. The empowerment of communities through this "role model effect" may also lead to increased reporting of criminal victimizations.

At the criminal offender level (the first stage of a criminal episode), the aggregate effects of greater minority political representation (i.e., on law enforcement and victim behavior) may deter potential criminals from committing crimes altogether the commonly known "deterrence effect." As expected, strong deterrence would lead to a decline in the actual number of crimes against SC/ST individuals. However, greater political power for groups that have been historically discriminated against may also result in an increase in criminal offenses against them for other reasons. For instance, if potential high caste offenders experience resentment due to increased political power or assertiveness among low caste communities, forward castes may commit more caste-motivated crimes in retaliation to political mobilization. In addition to being an expression of resentment, such an act could be part of a calculated strategy to intimidate leaders and voters ("the retaliation effect").[19]

Finally, it is worth noting that the skill/human capital of politicians from low-caste parties has occasionally come into question (Acharya, Roemer, and Somanathan 2015). To the extent that these caste-based parties are less effective, one might think that they are *less* likely to be able to offer protection to underrepresented minorities.

Thus, the overall effect of minority political representation on the number of documented crimes against SC/ST populations is theoretically ambiguous. If the deterrence effect within disadvantaged caste and tribe communities dominates, one would expect to observe a decline in crimes against these subpopulations when they are politically represented in greater numbers within state legislatures. However, if the law enforcement and reporting effects outweigh the deterrence effect, we would actually see an increase in the total number of documented crimes against low-caste minorities. Similarly, if representation induces retaliation by upper caste communities, we would also observe a crime increase.

It is worth briefly noting that Indian crime data likely suffers from reporting problems that raise concerns about measurement error in our dependent variables. The most serious problem is underreporting of crimes, a prime reason for which is a lack of trust in the police (Ansari, Verma, and Dadkhah 2015). Survey evidence suggests that a substantial number of citizens are dissatisfied with how police respond to crime and treat victims (Singh 1996; Chockalingham 2003). Recording of crime by the police are also affected by limitations of criminal justice system resources, police hiring practices, as well as systematic biases in the police that may arise and persist due to multiple factors. We try in our analysis to address some concerns related to the criminal justice system. Moreover, we believe that spurious trends due to problems related to measurement of crime would bias estimates of the impact of a representation shock in a positive direction, since it is more likely that increased low-caste representation would increase reporting of low-caste crimes, as well as low-caste response of police to crimes. As such, to the extent that we actually observe a crime *decline* after an exogenous increase in low-caste political representation, we interpret our estimate as a likely lower bound on the magnitude that crime declines subsequently.

[19] This was often witnessed in the state of Bihar through the 1990s when private armies of upper castes clashed against an increasingly assertive Dalit populations who were supported by militant Leftist outfits.

2.4 Data and Empirical Strategy

The goal of this paper is to identify the effects of organized political representation for SC/STs, emanating from "Affirmative Action parties" and their explicitly pro-SC/ST policy agendas. We test whether these identity-based parties translate electoral success into targeted public good provision in the form of protection from acts of violence at the hands of forward castes. To summarize our approach, we construct a panel data set where the unit of analysis is $state * year$ for nineteen Indian states from 1994 to 2012. The outcome of interest is one of 5 categories of crime committed by a non-SC person against a SC/ST victim. The explanatory variable of interest is aggregate SC/ST political representation measured at the state level (which will be stable within state for a term). We use both a naive OLS approach that measures the raw correlation with the number of elected caste-party representatives, as well as quasi-exogenous variation coming from close electoral contests between candidates of AA parties and traditional mainstream political parties. The exogenous variation in representation in our primary specification comes from the "surprising" component of AA party wins in close elections (deviations around the expected 50-50 split).

2.4.1 Data

We construct a state-by-year time series-cross sectional dataset by merging data across three main sources: government crime reports, election returns, and the Indian Census.

Crime/Criminal Justice Data: Since we are interested in the incidence of race-based crimes, our primary dependent variable of interest is crimes committed by non-SC/STs against SC/ST victims. We measure this outcome at the state-level using data provided yearly by India's National Crime Records Bureau (NCRB) through its annual "Crime in India" publication. [20]. Since the 1950s, the NCRB has provide aggregated statistics for a wide array of violent and property crimes. Since the early 1990s, however, the NCRB has also released state-level statistics on the number of violent crimes perpetrated specifically against SC/ST individuals by non-SC/ST perpetrators.[21]

We use as outcomes a set of crimes for which research show SC/ST individuals to be vulnerable to biased and discriminatory behavior – in particular, murder, rapes, kidnappings, and assaults. These violent crime categories are typically discussed in the literature about caste-related abuse (Mangubhai and Lee 2012; Hanchinamani 2001). We digitize data from the "Crime in India" reports on the number of reported crimes with SC/ST victims/non-SC/ST offenders for each of these 4 violent crime categories. In addition, we also collect data and analyze as an outcome violations of India's oldest hate crime legislation, the 1960 Civil Rights Act (CRA). The CRA defined specific (typically violent) crimes against Scheduled Castes and Scheduled Tribes as "atrocities," and set forth punishments to deter these acts.

[20] The NCRB issues aggregate data at the state, district, and city levels. The data are based on First Information Reports (FIR) filed at various police stations within districts. FIRs form the basis of any investigation into a crime.

[21] Unfortunately, there are no records of crimes committed against groups formally classified as Other Backward Castes (OBCs), although admittedly these groups are sometimes considered constituents of low-caste political parties.

Examples of banned activities include forcing members of an SC/ST to eat an offensive substance (such as feces), parading SC/ST individuals naked in city streets, or compelling them to engage in bonded labor or scavenging jobs.

In order to draw general conclusions in a context of multiple crime outcomes, we also construct a z-score index of crimes that aggregates all five crime categories described above. The index is an equally-weighted average of the yearly z-scores of its disaggregated crime components (Dal Bo and Rossi 2011; Kling, Liebman, and Katz 2007). The z-scores are crime levels standardized using the mean and standard deviation for state-year observations. Standardization is done by normalizing each crime outcome using mean and variance for each crime category within a state, across the time period. Higher-crime measures have higher z-scores. Aggregating outcomes improves the statistical power to detect impacts and reduces the risk of "false positives" by reducing the number of statistical tests carried out (Kling, Liebman, and Katz 2007).

We also analyze measures of total crime (i.e., crime counts irrespective of victim/perpetrator caste). We collect this data because one natural concern with our dependent variable of interest–caste-targeted crime– is that it may capture trends in overall crimes that are unspecific to SC/ST communities. To rule out that changes in caste-related violence are part of broader trends in Indian crime, we collect NCRB crime data at the state level for the same time period. We again focus on the violent crime categories that allow us to compare our results on caste-targeted crimes: murder, rape, assault, and kidnapping. We also collect data on property crimes and economic crimes, which one might think are crime categories unlikely to be suffered disproportionately by SC/ST communities, given the literature on caste discrimination just referenced.

Finally, we collect additional state-level data on other criminal justice system variables from the NCRB. This includes police strength (by caste and overall), crime-specific arrest rates, and court activity data. These data allow us to test the policy/bureaucratic channels that may influence crime.

Political Representation Data: Our explanatory variable of interest is the fraction of representatives within Indian state legislatures (the Vidhan Sabha) that are members of low-caste political parties. As described above, these parties represent SC/ST policy interests. To compute our measure of low-caste political representation, we use data from the Election Commission of India which presents detailed election results for each electoral constituency in a state. An electoral constituency is the administrative area that a state legislator represents (analogous to, for example, a district that a state-level legislator represents within a US state general assembly). The raw electoral data allows us to determine the candidates contesting each constituency-level election within a given state and year, the parties of each candidate, and the votes received by each candidate. Using this data, we can calculate the winning candidate (and his/her party) and the winning vote margin. We can then aggregate to the state level the number of AA party state legislators and the overall share of seats won by each political party in an election cycle.

Low-caste political party representation is stable within an election cycle, and elections occur at five-year intervals. The electoral cycle varies across states, and in any given year, elections for state legislatures are held in 4-5 states nationwide. For all elections, we categorize each candidate as being either from an AA party or a non-AA party, based on our classification of AA parties described in Section 2. Thus, for each constituency race, we

categorize candidates from the BSP and the multiple political party subgroups of the JD as "AA party" candidates using a dichotomous indicator. Our implicit assumption is that AA party members are "low-caste representatives" in the sense that the parties of these representatives focus on policies to advance the interest of low-caste subpopulations.

Other Data: We also collect individual-level attitude data from the Indian Human Development Survey that allows us to consider channels of influence related to attitude changes. Moreover, to both demonstrate the exogeneity of our explanatory variation and to include as control variables in our regressions, we collect data for several demographic factors that may influence our outcome variables of interest. Data on six state-level measures–gender ratio, literacy rate, population density, rate of urbanization, percentage of the population that is in the workforce, and the percentage of the population who are SCs or STs–are collected from three rounds of the Decennial Census of India occurring in 1991, 2001 and 2011. Data for the intervening years are interpolated using a linear trend. [22]

Before proceeding to the analysis, we discuss some key features of the dataset used for the analysis. Our dataset contains 16,277 constituency level elections to multiple state legislative assemblies over twenty-one years. Each of these elections are first-past-the-post elections contested by multiple political parties, with the winner on average receiving 45 percent of the votes.[23] Each election in our sample had at least one AA party contesting and 2,019 elections - about 12 percent - can be deemed as 'close' according to the 2 percent threshold used in Iyer et al. (2013). Within this universe of close elections in our sample, 1,128 are close elections involving an AA party and 976 of them are close elections where an AA party is paired against a non-AA party. The remaining 152 close elections are those where both the winner and the closest runners up within a 2 percent margin are AA parties. We drop these elections from our sample as the focus of this paper is on the relative impact of AA parties vis-a-vis non-AA parties and close elections involving two AA parties guarantees the presence of an AA legislator. Similarly, we drop close elections not involving AA parties as regardless of which party wins, there is no AA legislator elected.[24] Table 1 presents state-level summary statistics, weighted by population. There is substantial variation across all variable categories.

We also include summary information on state-level electoral characteristics, since in this paper we exploit variation in the number of state-level close elections. Appendix Table 1 presents state-by-state data on elections. Geographically, we observe wide variation in party representation and political competition. On average, there are around 200 state-level elections in a given election year. Some legislatures are quite small. Jharkhand, for

[22] These six variables are identified by Dreze and Khera (2000) as being the principal determinants of murders in India, and so we assume them to be valid control variables for all crimes that we are interested in studying, which are mostly violent crimes.

[23] Typically, each election involves multiple independent candidates who are able to garner a very small number of votes. We control for the average number of political parties to account for political competition, in line with Laasko and Taagepera (1979).

[24] Conceptually, this can be thought of in the LATE framework of an instrumental variable. Close elections involving 2 AA parties are the compliers with the randomness in the outcome of close elections not altering the identity of the legislators. Likewise, close elections involving 2 non-AA parties are the non-compliers. The paper's focus therefore is on the elections where akin to a coin-flip, we have an AA (non-AA) legislator elected, as opposed to a non-AA (AA) legislator.

example, had two sets of elections for the state assembly over our data period for eighty-one constituencies, and Himachal Pradesh elected 68 legislators in each of five elections. Uttar Pradesh on the other hand, elected over 400 legislators in each of six elections.

There is also substantive regional variation in the number of close elections involving AA candidates. On average, states have thirty-eight close elections per cycle that feature an AA party. Looking at some of the states individually: of Uttar Pradesh's (average) 418 elections, 99 were close and involved an AA party candidate. In contrast, in Maharashtra, where the assembly size is 288, there were only 12 close elections with an AA party per cycle. We also see variation in the strength of the traditional political powers the right-wing BJP and left-of-centre Indian National Congress.

2.4.2 Empirical Strategy

Our goal is to test whether an increase in AA party success reduces crimes against SC/ST caste groups, India's most vulnerable ethnic minorities. As previously mentioned, data on crimes committed against SC/ST groups are aggregated to the state level. Our empirical strategy is in turn devised to test whether having a greater number of state legislators representing underrepresented minority subgroups significantly changes levels of crimes committed against these subpopulations.

Naive Model

To motivate our preferred specification based on the quasi-random variation in the number of AA party legislators elected to the state legislature, we begin with discussion of a "naive" statistical model. This baseline specification can be written as a standard OLS specification with state and year fixed effects:

$$\ln(CrimeRate_{st}) = \alpha_s + \delta_t + \beta TotalAALegislators_{sy} + \eta TotalSeats_{sy} + \gamma \mathbf{X}_{st} + \epsilon_{st} \quad (2.1)$$

The outcome variable is the natural logarithm of minority-victim crimes committed as a fraction of a state's SC/ST population. β is the coefficient of interest, and represents how the number of legislators belonging to an AA party is correlated with crime rate (we also control for the total size of the legislature, as the raw number of AA party wins may be correlated with number of winnable seats).[25] The coefficient on the explanatory variable of interest tells the percentage change in the crime rate with an additional low-caste representative.[26]

[25] We also estimate the same regression in which we use the proportion of AA party legislators as a fraction of total legislature size is the variable of interest:

$$PAA = \frac{\text{Legislators Belonging to AA Party in State}}{\text{Total Legislators in State}} \quad (2.2)$$

The results are qualitatively similar when using this form of the explanatory variable.

[26] For two reasons, we use regressions on crime rates expressed in natural logs. First, research on the determinants of crime suggest using the log transformation to account for the likely nonlinearities in the relationship between violent crime and its determinants (Fajnzylber, Lederman, and Loayza 2002). More

In this specification, s denotes the state, t denotes the year of observation, and y the year of election to the legislative assembly in state s. X is a vector of state-specific factors described in the data section that may also affect levels of crime, and is measured at 10-year intervals (alleviating traditional concerns about the endogeneity of time-varying controls).[27] α and δ indicate state and year fixed effects, respectively.

The explanatory variable of interest, $TotalAALegislators_{sy}$, remains unchanged for all years within an electoral cycle. Formally, if elections occur in state s in year y and the subsequent election occurs in year $y+5$, $TotalAALegislators$ is invariant for all $y \leq t \leq y+5$. The coefficient of interest is then β, and identifies if a higher state-level number of legislators belonging to AA parties significantly alters crimes perpetrated against HMCs relative to average crime outcomes. A natural concern is that as our explanatory variable varies only in five-year intervals, we are artificially improving our statistical power by using an outcome measure which varies annually. To alleviate this concern, we discuss below a robustness check in which we collapse our crime data to five-year averages that correspond with electoral cycles. Reassuringly, the results remain similar in both magnitude and precision.

Quasi-experimental Approach

It is possible that unobservable factors could be affecting the success of caste-based political movements within states and localities. These factors may be correlated with levels of crimes suffered by the constituents of these parties, biasing our estimate of β. For example, states that elect more legislators from AA parties may already have more favorable overall public opinions towards SC/ST communities. If this favorable opinion is also shared by law enforcement agencies, then a police force that is more protective of marginalized groups may already vigorously prosecute, and by extension deter, crimes directed towards these populations. In this scenario, our estimate of the effect of political representation on caste-related crimes would be biased away from zero. Another plausible problem for identification is that in areas where crimes against minorities are policed more carefully and prosecuted more vigorously, citizens are more supportive of low-caste politicians, leading to the success of caste-based AA parties. This would again induce a downward bias (away from zero) in our coefficient of interest.

Any unobservable factor correlated with both SC/ST political representation and the minority crime rate would bias the estimate of the effect of having AA party representation in a state assembly. To resolve this classic endogeneity concern, we exploit the quasi-random variation in AA party success within state legislatures using close contests between caste-based and traditional party candidates. The goal of this empirical exercise is to isolate the effect of an *unexpected* increase in representation for SC/ST communities.

As described in the introduction, we cannot map characteristics of the winning candidate

practically, natural logs are used so that coefficients can be read as percentage changes in the crime rate (Agan 2011). Importantly, though, the results are robust to other forms of the dependent variables. In an appendix, we show the impact of representation on raw crime rates, as well as the impact of representation on the logged crime levels.

[27] Another natural concern is that some of our controls may be endogenous with respect to political representation. To alleviate this concern, we conduct a robustness check in which we also measure these outcomes at the baseline year, and interact these variables with a linear trend.

(in this case, having SC/ST constituents) to the geographical area that a legislator represents: we have constituency-level electoral data and state-level outcome (crime) data. States are an aggregation of administrative constituencies. To resolve the issue of treatment-outcome level mismatch, we *aggregate* tight electoral contests to the state level to create a state-level measure of quasi-random exposure to low-caste political representation. We use the sum of all close AA party victories in state elections against non-AA parties, controlling for the total number of AA-non-AA party close contests. The intuition is that given N close elections involving one AA party candidate, we should expect the AA party candidate to win 50% of all these contests. Put another way, conditional on the total number of close contests, the number of observed AA party wins is the quasi-random variation in political representation for low-caste communities. Our strategy relies on the fact that this variation in state-level low-caste government representation is uncorrelated with state-level support for AA party candidates and also relies on the assumption that electoral contests decided by a very small margin are equivalent to being decided "randomly" (as long as there are no correlated shocks, which we argue later is not problematic in our setting).

To make our identification strategy concrete, we provide an example of how we derive our state-level measure of exogenous variation in AA party representation:

- Assume that out of 250 constituency-level elections in a state, 100 involve an AA party candidate (i.e., low-caste) against a non-AA party candidate, and 50 of these are decided by a margin of less than 5% (i.e., 50 are "close elections"). Assume that the AA party candidate wins 30 of these contests.

- If we assume that elections that are decided by a very small margin are equivalent to being decided "randomly," we expect the AA party candidate to win 50 % of these contests, or 25 contests.

- So, we can call 30-25 = 5 the unexpected/ *"exogenous"* increase in the number of representatives for low castes in this state legislature

Close elections are defined similarly to Asher and Novosad (2015) and Bhalotra, Clots-Figueras, Cassan, and Iyer (2014), who also study Indian legislative races. We limit to elections where the margin of victory between the winner and loser is less than or equal to two percent of the votes cast, and where one of the competing candidates is from an AA party.[28] The outcome of any close election is assumed to be "quasi-random," and so any party in expectation should have a 50 percent chance of winning a close election it contests. As such, the state-level proportion of close elections won by an AA party against non-AA parties should approach 0.5 as the number of contests grows (we show this empirically in the summary statistics). Close elections contested between two AA (non-AA) parties do not contribute to our final data sample. Controlling for the number of total close elections between AA and non-AA parties allows us to account for variation across states in the strength of caste-based political parties or other state-level factors that contributes to caste-based political competition.

[28] In the robustness checks of the main results, we show that the results are invariant to the choice of victory margin threshold.

Using this quasi-exogenous variation in low-caste political party representation, our main specification with quasi-random variation in low-caste representation becomes the following reduced form model:

$$\ln(CrimeRate_{st}) = \alpha_s + \delta_t + \beta CloseAAWins_{sy} + \eta TotalCloseElections_{sy} + \gamma \mathbf{X}_{st} + \epsilon_{st} \quad (2.3)$$

CloseAAWins is now the explanatory variable of interest as defined above, and the coefficient of interest is again β. This specification allows us to identify the impact of the marginal additional legislator from an AA party on crimes committed against SC/ST caste groups. In essence, the variable of interest constitutes the number of additional legislators that can be considered a "like-random" increase in SC/ST political representation. Assuming that the outcomes of close elections are equivalent to a coin-flip, this measure satisfies the necessary conditions to be used to measure the causal effect of the marginal low-caste representative on the outcome variable of interest - crimes committed against ethnic minorities. $CloseAAWins$ is also positively correlated with the *total* number of elections won by AA parties (we confirm this empirically; we also do so similarly for the proportion of AA legislators as a fraction of the total seats).[29]

2.5 Results: Effects of Organized Political Representation on SC/ST Criminal Victimization

We now turn to our analysis of how the political representation of Scheduled Castes and Tribes that comes from identity-based political parties affects targeted criminal victimization.

2.5.1 "Naive" Specification Results

We begin by discussing the results obtained by estimating β from Equation (2.1), our naive specification. This regression exploits all of the variation in low-caste representation. All the specifications estimated include year and state fixed effects and robust standard errors, clustered at the state level (shown within brackets in the tables). As discussed in Section 2.4.1, we examine 5 crime categories: murder, rape, assault, kidnapping, and civil rights violations, as well as a composite index of all crime categories to avoid concerns about multiple hypothesis testing.

Table 2.2 presents the results for this empirical exercise of exploring the correlation between low caste criminal victimization and political representation. Thinking first about the correlation between overall crime and low caste representation, Column (1) presents results for the naive specification where the outcome is the summary measure of crime. The results suggest that there is, overall, a negative correlation between low-caste representation and caste-based crime. Each of columns (2) through (6) presents point estimates in which the

[29] Note that this empirical strategy could also be implemented using an instrumental variables approach in which the first stage is the regression of total wins by AA parties on the number of close elections won by AA parties. In our case, however, the reduced form is as direct a study of what we hope to identify, which is the impact of a marginal AA seat. As such, we focus on this estimate.

outcome variable of interest is a disaggregated crime category for which SC/ST subpopulations have traditionally been targeted. The results collectively suggest a negative correlation between AA party success and crimes against SC/ST communities. Consistent with the summary crime results, the remaining estimates also suggest a negative correlation between low-caste political representation and reduced SC/ST criminal victimization, although the estimates are generally not very precise. The magnitudes of the coefficients are small (around 1% reduction in crime rates) and for rape, murder, and kidnapping, are either statistically significant or close to being statistically significant at conventional standards (which is not entirely unreasonable given our sample size in this empirical exercise).

The results here are suggestive of a relationship between targeted violent crime and minority political power. However, for the reasons of endogeneity discussed above, these results should be treated with caution. We now move to our primary empirical exercise of estimating the causal effects of AA party representation on caste-related violence.

2.5.2 Main Results

Validating the Exogeneity of Close Election Outcomes

The empirical strategy described in Section 2.4.2 relies on the assumption that the outcome of a close election is quasi-random. To confirm the internal validity of our results, we begin by exploring the identifying assumptions of our empirical model. We first show that the outcome of close elections between an AA party candidate and a non-AA party candidate can be considered "quasi-random" events. We subsequently demonstrate that the total number of close elections won by AA parties during a state electoral cycle is independent of observable political and demographic covariates of interest. Together, these results suggest that our measure of SC/ST representation–the total number of close, "like-random" AA party local wins–is a valid source of quasi-exogenous variation for making causal statements about the aggregate impacts of low-caste representation.

As the elections to the state legislative assembly are conducted at the constituency level, we first test the smoothness in the victory margin for AA parties - the running variable in the regression discontinuity design at the constituency level - across constituency-level observables, at the cutoff point of 0. At this threshold, the outcome of the election changes discontinuously from an AA party loss to an AA party win. We conduct a density test akin to McCrary (2007), where we test whether the density of constituency elections is continuous near this cutoff value (the idea behind this is a test for the absence of manipulation around the cutoff). If there is no evidence of a discontinuity in the density of units at the zero-cutoff, we can be more certain that there is likely not self-selection or non-random sorting of AA party wins and losers into control and treatment status.

The results of this test are presented in Figure 2.1. In our case, victory margin for AA parties is defined as difference in the vote share between an AA party and a non-AA party with a positive (negative) difference signifying a victory (defeat) for an AA party.[30]

[30] For elections involving multiple AA and non-AA parties, we use the following rule to determine the victory margin: if an AA party wins, the victory margin is the difference in vote share between the winning AA party and the non-AA party with the highest vote share amongst all non-AA parties. If a non-AA party

The running variable is partitioned into 500 bins in the interval between -0.25 and 0.25 and we plot the average value of the covariate corresponding to each bin. The graphical plots validating the exogeneity in the outcome of close elections for AA parties at the constituency level is shown in the first figure. The red line marks the cutoff point at which the victory margin is 0, signifying the discontinuous change in outcome from an AA party loss to an AA party victory and we use a second order local polynomial with a 95 percent confidence interval to identify the relation between the running variable and constituency level observables on either side of 0.

We observe from Figure 2.2 that the plots are extremely smooth for nearly all of the constituency-level characteristics. One slight exception is for the winning candidate's age (plot 9), where there is a modest drop at the point of discontinuity (i.e., winning AA candidates tend to be slightly younger). On the whole, however, we believe Figures 2.1 and 2.2 demonstrate that our key source of variation - the chance of an AA party winning a close election - is indeed "as good as random."

We also provide evidence demonstrating that our constructed measure of low-caste party representation, obtained by adding the total number of close wins by an AA party in any state electoral cycle, is also independent of any observable characteristics. First, we plot the actual number of total close wins against the expected number of total close wins by an AA party. If the outcome of a close election is drawn from a stochastic process, we should expect AA parties to win half the number of close elections contested by them. Figure 2.3 shows confirms our expectations in this regard as we see that the actual number of close wins is distributed around the 45-degree line and a linear fit explaining the relationship between the actual number of close number of wins and the expected number of wins does not deviate much from the 45 degree line.

To be more formal in this test of exogeneity, we also individually regress the total number of close wins during any state election on a number of demographic and political covariates. The results are presented in Tables 2.3a and 2.3b; each specification also controls for the total number of close elections occurring in that electoral period between an AA and a non-AA party, and the standard errors are clustered at the state level. Each column in this table is a regression of our quasi-random low-caste political representation on a single covariate, where we expect (if the variation is quasi-random) that the covariate should not be significantly predictive of the number of close AA wins.

In Table 2.3a, we first observe that the total number of close AA party wins is strongly predicted by the total number of close elections contested by AA parties against a non-AA party, with the coefficient being centered at 0.5. Importantly, no demographic covariate (population, SC/ST population share, gender ratio, literacy, unemployment, or % urban population) predicts the total number of close AA wins in a state. Similarly, political characersitics of each state do not affect our exogenous measure of low caste representation. The results presented in Columns (1) through (4) of Table 2.3b show that the share of constituencies reserved for low caste candidates, voter turnout, overall vote share for AA parties, and the number of parties competing (a proxy for political competition) have no

wins the election, the victory margin then is the difference in vote share between the victorious non-AA party and the vote share of the AA party which has the highest vote share amongst all AA parties

impact in the number of close elections won (lost) by AA parties. The magnitude of the coefficients are small and statistically insignificant, suggesting there are no political conditions that systematically lead to close AA party wins (or losses).

We pay particular attention to Columns (5) and (6) of these tables. Column (5) confirms that the overall level of AA party success in a state (as measured by total contests won) is not predictive of the quasi-exogenous variation we exploit, the aggregate outcome of close elections. Similarly, previous AA party success has no predictive power over current AA party success in close elections. In other words, these two regressions tell us that neither the current, nor the historical success of AA parties affect their likelihood of winning close elections in a given electoral period.

Collectively, this collection of estimates and statistics suggest that the state-level variation in low-caste representation coming from aggregated close AA party wins is indeed quasi-random. As such, we can interpret our estimate of β from Equation 2.3 as causal.

Quasi-Experimental Estimates of AA Party Impact

We now disuccus the main results of the paper, based on the exogenous variation in low-caste representation created by "unexpected" AA party success in close electoral contests. Table 2.4 present the core results for the paper, based on the reduced form model in Equation 2.3. The explanatory variable $TotalCloseAAWins$ in each table tells us the aggregated close election wins in each state, and provides us with a causal estimate of β. Again, each column indicates an estimation of the reduced form quasi-experimental regression with a different crime category as the dependent variable (for Column (1) the z-score crime index is the outcome). The coefficient estimate for each column has a straightforward interpretation: we estimate the percentage change in the rate of crime against SC/STs resulting from an unanticipated increase of one elected AA legislator, given that we control for the total number of close contests (which we view as necessary given the potential endogeneity concerns raised by comparing places with and without many closely AA party vs. non-AA party contests

The results in Table 2.4 suggest that an extra AA party representative reduces significantly violent crime suffered by Scheduled Castes and Tribes. Column (1) indicates that there is a clear reduction in overall crimes suffered by SC/ST, perpetrated by non-SC/ST individuals. Similarly, looking at the crime categories individually in columns(2)-(6) in Table 2.4, the regression results suggest that increasing minority political representation reduces the level of the major types of crimes committed against marginalized minorities (which is expected if political representation deters crimes against its intended constituents). Significant reductions are observed for murder, assault, and civil rights violations: a (conditionally) random increase of 1 AA party legislator leads to reductions varying between 8%, 14%, and 30% for these crimes, respectively. Kidnapping and rape rates also experience declines for a unit increase in quasi-random AA party representation, with significance levels just below standard conventions. These results collectively suggest increased political representation produces a modest reduction in certain types of violent crimes committed against ethnic minorities. The results are stable to the exclusion/inclusion of state-level controls.[31] In Panel

[31] In an Appendix, we repeat this estimation, but also include time dummies for years since the last election, to account for potential political cycle effects in government priorities regarding crime. The results

A of Table 2.4, we control for various state-level demographic factors that may be correlated with both political support to a low-caste political party and the state-level incidence of crime. In Panel B, we omit the controls, but the results look qualitatively similar, with violent crime reductions ranging between -4% and -14% (unsurprisingly, given our small sample size, the precision is slightly lower for each crime category).[32]

Our causal estimates in Table 2.4 are in general larger in magnitude from the OLS estimates in Table 2.2. As discussed previously, the OLS regressions that use the entire variation in AA party representation are likely to be biased due to omitted variables and potential reverse causality. The share of caste-party legislators is likely to be correlated with underlying voter preferences and other area characteristics. In particular, politicians representing low caste constituents are more likely to be elected in areas with larger low caste populations, which may also have a lower level of public goods provision (Das et al., 2011). There could also be reverse causality in the sense that if AA party politicians are more likely to provide law enforcement and other sources of crime/discrimination prevention, then they are more likely to win in places where public service provision is particularly poor and hence there is a greater underlying demand for such services. The quasi-experimental reduced form results take advantage of the fact that in some areas AA party candidates won by very few votes against traditional party candidates, and in other areas the reverse occurs. Such areas are likely to have similar characteristics, SC/ST population shares, similar candidate ability, and similar access to public goods that affect crime ex-ante. As such, there is significantly less concern about sources of bias, which often run in opposite directions.

The causal interpretability of the core results presented in this section is conditional on the exogeneity of the independent variable of interest - the number of low caste representatives elected through close elections (conditional on the total number of close contest). In addition to the exogeneity tests discussed previously (see Tables 2.3a and 2.3b), we demonstrate that no state-level political or demographic covariate can predict this variation. We now present an additional test in the spirit of Altonji, Elder, and Taber (2005) to demonstrate unobservable covariates are possibly not driving the core results of the paper. This test involves regressing our independent variable of interest - number of state-level low-caste representatives elected through close elections - our state-level electoral and demographic covariates, in addition to state and year fixed effects. The fitted values from this regression provide a measure of the number of state-level low-caste representatives elected through close elections, as predicted by observable covariates. Subsequently, we regress our measures of targeted crimes against SC/STs on these fitted values and compare the results with those obtained from regressing our crime measures on the actual number of state-level low-caste representatives elected through close elections. Intuitively, this tests informs us if the best

---

are unchanged.

[32] In the appendix, we also repeat the primary analysis using a two-stage least squares approach, in which we instrument for overall changes in AA party representation within a state legislature using the "quasi-random" variation based on deviations from the expected number of wins in close elections for AA parties. Table A2.1 presents results from the IV estimation, and again suggests a non-trivial decrease in overall crime as well as each individual crime category given an exogenous increase of one HMC legislator. The results presented in Table A2.1 are broadly consistent with the OLS results. The 2SLS results are in most cases slightly smaller in magnitude than the reduced form results. Notably, however, magnitudes are closer to the OLS results.

available prediction of our key independent variable using observable covariates can also explain the variation in our outcome measures of interest. The results are presented in Table 2.5 and show a positive correlation between targeted crimes against SC/STs and the predicted number of state-level low-caste representatives elected through close elections, contrary to the negative impact identified in Tables 7 and 8. Moreover, with the exception of murders, none of these results are statistically significant. As long as the remaining unobservable covariates are positively correlated with our present set of covariates, the results presented from the Altonji, Elder, and Taber (2005) test should assure us that the results presented in Table 2.4 are not driven by some unobserved correlation between some covariate and the number of state-level low-caste representatives elected through close elections.

2.5.3 Electoral Cycle-Level Results

Due to the periodic nature of elections to the state legislative assemblies occurring roughly once every 5 years, our independent variable of interest varies once every 5 years while our outcome measures of crime vary annually. To verify that our results are not an artifact of this differential variation in the dependent and independent variables, we conduct two tests to confirm that our core findings above are plausible. First, rather clustering our standard errors on state, we cluster on "$State * ElectionYear$", in order to account for potential correlations in crime outcomes over time, but also within a given state and election cycle (since our explanatory changes with each election cycle). The results of this regression are presented in Table 2.6, and reassuringly, the precision of our initial estimates is virtually unchanged.

Second, we reconstruct our dataset to be at state "$State * ElectionCycle$" level rather than the state-year level. Our explanatory variable is unchanged in this dataset. However, our outcome variables are slightly altered and now measure the average crime level between two election years, conditioning out the variation for individual years within an electoral cycle. One concern is that crime is affected by electoral cycles. To remove this concern, before computing the average crime across cycle, we first regress our outcome measures on a set of five dummies, where each dummy i - $0<i<6$ - is an indicator for the year after a state election year. The residuals from this regression represent the variation in targeted crimes against SC/STs which are unexplained by years lapsed after elections. After partialling out the time effects, we obtain the mean value of our measures of targeted crimes against SC/STs within each electoral period and regress them on the number of state-level low-caste representatives elected through close elections, in addition to state-fixed effects, an election year dummy and our set of state and time varying controls, which we also average across electoral periods. The results from this specification are presented in Panel B of Table 2.6. The sample sizes are understandably smaller, representing the number of state-level elections observed in the 20 year period studied in this paper. Reassuringly, we observe that the number of AA close victories still has a negative impact on targeted crimes against SC/STs and despite the significant reduction in sample size, the coefficient on murder and CRA are significant (murder at the 10 percent level) and the p-values for the remaining outcomes are all between 0.1 and 0.15 with the exception of rape, which too is under 0.2. The fact that we are still able to identify a fairly precise negative effect of the number of state-level low-caste representatives on targeted crimes against SC/STs validates the core

results presented in Table 2.4 and confirms that the results are not a function of the data structure.

2.5.4 Heterogeneous Effects of Low-caste Representation: The Role of Party Power

Thus far, our estimations assume the effects of representation are linear in nature. It is possible, however, that the effects of representation are nonlinear. Research on the effects of other historically underrepresented subpopulations have anticipated such a possibility; such research exists for the effect of female legislators (Kanter, 1977; Schwindt-Bayer and Mishler, 2005).

We explore this possibility in two ways. First, we test whether the marginal politician representing low caste interests has an even greater ability to improve the lives of constituents when his/her party is in power. We thus next test whether the power of an AA party representative increases with a controlling presence in the legislature. We do so exploiting that AA parties have commanded governing majorities in certain states over the past twenty years. In Table 2.7, we include an interaction term between our quasi-random variation (total close AA wins) and a dummy variable for whether an AA party is in power in a given state and year. We expect a negative sign on the interaction term if a marginal low caste politician has a greater effect (in terms of reducing crime) when his/her party is in power. Table 2.7 confirms that this is the case for 4 out of our 5 crime categories, and also for overall levels of violent crime based on our index.

Even if a regional party representing SC/ST communities explicitly is not in power, another possibility is that the marginal low caste representative in the state assembly has a greater ability to allocate resources towards minority crime deterrence when there is a critical mass of public support across the state. To test this, we also include an interaction between our measure of quasi-random caste-based representation and bins for high levels of vote share across the state. Results from this test are shown in the Appendix (see Table A2.2), and suggest that the marginal representative is even greater when there is a sizable base of public support for the AA political party.

2.6 Robustness Checks

We now perform three main robustness checks to assert our confidence in the validity of our main results. First, we test for political heterogeneity across India by including several political competition measures. Second, we assess the sensitivity of the main results to alterations of the two percent victory margin threshold which is used to classify close elections. Finally, we use a different measure of low-caste political party representation to assess the effect of an additional AA party representative on SC/ST criminal victimization.

2.6.1 Accounting for Cross-state Political Heterogeneity

One potential concern is that certain other political parties may be more or less likely to take actions that reduce crimes committed against HMCs. The presence of such parties

could be another possible confounding factor. According to a Human Rights Watch (2007) report, members of the Bharatiya Janata Party (BJP) called for the repeal of hate crimes legislation, the Prevention of Atrocities Act (POAA).[33] Given its antipathy towards the law, a strong BJP presence within a state may have a direct effect on either the documentation or prosecution of criminal offenses.

To test whether the effect of AA party representation on crimes against HMCs is affected by a strong presence of the BJP or the Indian National Congress (the other dominant political party in India), in Panel A of Table 2.8 re-estimates Equation 2.3, including a robust set of political variables for the strength of other political parties and measures of political competition. These variables include the percentage of legislators from the BJP and the INC - the two dominant national parties in India - the 'effective number of parties' participating in an election, and the number of candidates [34]. The results remain qualitatively and quantitatively similar to the main results. The effect magnitudes are somewhat smaller, and the level of precision declines for some crime categories.

2.6.2 Altering the Victory Margin Threshold to Determine Close Elections

To this point, all the results have been presented using a victory margin threshold of 2 percent to define close elections. We now test the sensitivity of the reduced form results under two alternative thresholds. The first restricts the definition of close elections to elections whose outcomes are decided by victory margins of 1 percent or less. The second expands the definition of close elections to elections whose outcomes are decided by victory margins of 3 percent or less.

As seen from Panels B and C of Table 2.8, re-estimating Equation 2.3 using these two alternate definitions of close elections does not cause any significant changes to the primary results. With a victory margin of 3%, the results are similar. The case for the crime-reducing effect of minority political representation also remains largely intact for even very small victory margins. Although generally less precisely estimated when limiting close elections to those decided by less than 1%, the relationship is statistically and economically significant for assault, and approaching conventional levels of significance for other violent crimes. These results give us confidence that the selection of a specific threshold for electoral contests which generate the quasi-random variation in exposure to low-caste representation does not drive the observed identity-based crime reductions that we identify.

[33] The BJP claimed that POAA was being used as a political tool to create support for officials and candidates amongst HMCs.

[34] Effective number of parties differ from the total number of candidates as the former is an index created by adding the inverse of the square of voteshare received by each contestant in an election. Henceforth, contestant winning a small share of votes do not contribute to the construction of this index

2.6.3 Varying the Structural Relationship between AA Party Representation and the Number of Close Elections

As an additional check of the stability of the primary findings, we use as an alternative explanatory variable of interest: the deviation of the actual number of AA party wins in close elections from the expected number (i.e., $\frac{1}{2}$ of the total number of close races involving 1 AA party candidate). We standardize the quasi-random variation (i.e., the increase/decrease over the expected number of AA wins) using the standard deviation of a binomial distribution where the expected value of the proportion of close AA wins is 0.5. In essence, this (standardized) difference between actual and expected close AA party wins constitutes the number of additional legislators that can be considered a "like-random" increase in HMC political representation. Assuming that the outcomes of close elections are equivalent to a coin-flip, this measure of exogenous variation also satisfies the necessary conditions to be used to measure the causal effect of the marginal AA party representative on the outcome variable of interest. [35] The difference between the main specification and this specification is that the latter imposes some structure between the random component of total number of close AA party wins and the expected number based on total close elections. The results of this test are presented in Panel D of Table 2.8, and are very similar to our main results. Again the signs are as expected, and indicate that the marginal increase in AA party representation leads to significant declines in both assaults and kidnappings. Overall, the point estimates are slightly smaller using this specification, but on the whole, they are consistent with our main results.

2.7 Mechanisms

We now explore potential policy mechanisms that may explain why legislative representation for India's historically marginalized communities may have plausibly contributed to the reduced incidence of violence suffered by these groups.

2.7.1 Changes in Crime Reporting and Attitudes towards Low-Caste Representatives/Communities

The data presented so far does not speak to whether political empowerment results in a greater willingness of SC/ST individuals to report crimes perpetrated against them. Indeed,

[35] An earlier version of this study used this variation as the primary explanatory variable. The variation can be formally defined as:

$$RAA_{sy} = \frac{\sum_{c=1}^{N} AACW_{scy} - 0.5 * \sum_{c=1}^{N} AACE_{scy}}{0.5 * \sum_{c=1}^{N} AACE_{scy}} \tag{2.4}$$

In Equation 2.4, N denotes the total number of electoral constituencies in state s. $AACW$ is a dummy variable equaling 1 if an AA party won a close election against a non-AA party in constituency c, while $AACE$ is a dummy variable equaling 1 if there was a close election involving an AA party and a non-AA party in constituency c and year y. RAA is also positively correlated with the state-level proportion of elections won by AA parties (we confirm this empirically).

we might reasonably expect that marginalized communities such as Scheduled Castes and Tribes are less willing to report victimization if either: (1) they feel that the state is not likely to respond, (2) they are unaware of what constitutes certain crimes (such as rape or assault), or (3) if they fear retaliation. Iyer, Mani, Mishra, and Topalova (2012), for example, argues that the *reporting* of gender-based crimes increases with increased female representation, while the actual rate of criminal activity stays constant.

In our case, however, the potential threat of underreporting likely *strengthens* our finding of reported crime declines. In particular, we expect that positive representation shocks should, if anything, *increase* the level of crime reporting by the now-represented groups. We believe this assumption to be plausible if AA party representation increases confidence in government institutions.

An alternative possibility that makes interpreting our core results more challenging could be that when politicians representing SC/ST groups come to power, they manipulate data in a manner that gives the appearance of increased SC/ST safety. While we view this scenario as unlikely (given that we can find no qualitative evidence for it), one might expect such government malfeasance to reduce trust in institutions.

To both rule out the possibility that AA party representation (i.e., low caste representation) reduces faith in government and provide evidence that the declines in reported crimes that we analyze above are a signal of the true decline in caste-targeted violence, which would be consistent with decreases in reported crime, and thus confound our estimates, we now present results suggesting that increases in representation from caste parties increases low-caste faith in government. To this end, we utilize data from two rounds (2004 and 2011) of the India Human Development Survey, which is a nationally representative sample of Indian households, and has questions on levels of trust in state and local elected representatives and local-level inter-caste conflicts.[36]

Table 2.9 presents our results using the same exogenous variation with this additional outcome data. We estimate linear probability regressions where each column presents a dependent variable SC/ST and upper-caste households. For each column, the outcome is a dichotomous variable for whether an individual experiences a conflict due to caste (Column 1) and whether an individual trusts a given institution (Columns 2-4) We first rule that caste/tribe-directed crime is differentially underreported after exogenous increases in low-caste/tribe representation. To do this, a dummy for whether a household has experienced a caste-related conflict is regressed on the exogenous caste party measure. This result is an implicit check on our main results, suggesting that hostile and potentially violent interpersonal interactions between SC/ST and non-SC/ST citizens does decline after the election of low-caste representations. In Column (1) of Table 2.9, we observe that reported inter-caste conflicts decline. Panel A shows the effect for SC/ST households, and encouragingly, we find evidence that affirms our prior conclusion that inter-caste conflicts decline in the presence of AA legislators. Column (1), Panel (B) presents the corresponding effect for upper-caste households. That both high and low caste respondents experienced reductions

[36] The survey is based on responses compiled from a panel of over 35,000 households across all Indian states. Households are sampled from towns with probability proportionate to their population. Town identifiers are embargoed but district identifiers are not.

in (self-reported) inter-caste conflicts in the aftermath of AA party representation (and by similar magnitudes) is strong suggestive evidence that the reported crime declines in Table 2.4 accurately represents the effect of minority representation.

In Columns (2)-(4) of Table 2.9, the outcome is a dummy variable indicating a high degree of trust in the listed institution. The results shown suggest that AA party representation increases trust among SCs and STs in politicians, police, courts, and the state government overall. Although we have a very limited time series for this data, we believe our results demonstrate increased trust in the state, and provide further reason to believe that there is no differential reductions in the reporting of crime to authorities when AA party legislators are in power. Panel B again shows the corresponding effects for upper-caste households, and show little effect of AA party success on upper caste perceptions of state, except with respect to courts.

2.7.2 Changes in Policing Behavior: Enforcement Effects and Police Composition

One of the main channels discussed in the conceptual framework that we can test is the "enforcement effect"- i.e., changes in law enforcement behavior. Importantly, as law and order is one of the main responsibilities of state governments, understanding this channel is crucial to understanding the political economy of crime control. As discussed, increased representation for low-caste subpopulations can reduce crime by changing how law enforcement officers respond to reported crimes. By increasing the likelihood that police will aggressively investigate reported crimes and pursue perpetrators, low caste representation may deter crime by demonstrating the state's commitment to preserving public order and protecting HMCs in particular.

We test whether political representation for low castes and tribes affect police responses by collecting Indian police data on arrest rates. After investigating a reported crime, a police officer in India must prepare a report for a magistrate judge that describes the results of his/her investigation. If the police report contains sufficient grounds for a criminal charge, an arrest can be made. Thus, arrest rates can be considered indicative of how vigorously police (and the state overall) protect minorities.

To explore if political representation changes how police respond to crimes against HMCs, we explore the effect of an exogenous increase in AA party legislators on arrest rates for crimes committed against HMCs. The results are in Table 2.10. For three of the four crimes for which arrest data is collected, we see that an increase in AA party representation significantly increases the proportion of arrests (as a function of crimes committed). This could be coming either through changed policy priorities of the politician, or changed attitudes towards HMCs by police officers.

It is also possible that politician identity may reduce identity-based violence by changing the identity composition of the law enforcement force. Beginning in 2001, India started to release aggregate data on the number of police officers within a state who belong to SC or ST communities. One might expect that police officers are particularly attuned to incidence of violence suffered by one's socio-ethnic group. To test the possibility that police force composition is another channel through which legislators affect crimes against HMCs, we

estimate our reduced form specification with the logged proportion of SC and ST police officers as our dependent variable. The results are presented in Table 2.11. Consistent with this possibility, AA party success is positively associated with increasing the proportion of police presence from SC and ST communities. That these estimates are not very precise is perhaps unsurprising given the reduced power owing to the curtailed time series of this data. These results however are not an artifact of overall increased police hiring, as shown in Column 3.

2.7.3 Changes in Court Processing

Another way that caste representation may reduce minority crime is by changing how the justice system adjudicates such matters. Potential criminals, for instance, may be affected not only by changes in the likelihood of arrest, but also in the likelihood of cases being prosecuted diligently.[37].

In India, district courts which forms the first step in the judicial hierarchy have jurisdiction over criminal matters. District courts are under the administrative control of the High Court of the state to which the district concerned belongs, providing state level politicians with some influence over court proceedings. While the Constitution guarantees judges independence from political authority, state governments retain the discretion to choose which criminal cases to pursue and determine the expediency with which to pursue the case. Also, as prosecution in criminal cases are initiated by the state government, the quality of lawyers, as well as the evidence presented remains under the state government's control. In this regard, state legislators can use their bargaining power to influence whether a case is presented in courts and also the nature of evidence presented to the courts. To this extent, we identify if a quasi-random shock in AA party representation affects either of these two outcomes.

Table 2.12 presents the impact of low caste representation on court activity. We see that exogenous increases in state-level representation from low-caste parties improves how the judiciary responds to criminal cases related to violence against Scheduled Castes and Tribes. The acquittal rate decreases, suggesting, for example, that cases are prosecuted more vigorously and the evidence provided is of a higher quality to withstand judicial scrutiny. The speed with which cases are prosecuted also seems to increase, as the fraction of criminal cases for violence against SC/STs that are pending at the end of the year also declines. Finally, the number of cases withdrawn also decreases, which again signals the state government's intent in pursuing such cases through the judicial process.

2.8 Ruling Out Alternative Hypotheses

We now conduct several falsification exercises to rule out potential alternatives to our argument that caste-based political parties are reducing targeted violence by improving bureaucratic processes related to criminal justice administration as well as improving attitudes (both of SC/STs towards the state, as well as between minority and non-minority citizens.)

[37] Donohue and Wolfers (2009) discuss the potential roles of prosecutors and courts in criminal deterrence.

2.8.1 Do Traditional Political Parties Also Reduce SC/ST Violence?

Our results thus far do not rule out the possibility that any political party that unexpectedly receives "extra power" in the state legislature from winning more close elections than expected might generate different political or policy outcomes. We now show one piece of evidence to suggest SC/ST crime declines we observe are truly capturing the effect of specifically low-caste party power on criminal victimization. To do this, we examine the effect of *non*-caste party-based representation. If SC/ST crime reductions are produced by political movements that privilege low-caste interests, one would expect no effect from quasi-random increases in traditional party power. As such, we examine the same type of exogenous variation arising from close elections between BJP and non-BJP candidates. The BJP is one of India's major parties that has not aggressively championed the interests of SC/ST communities. Historically, the party capitalized on Hindu nationalist tendencies in the Indian citizenry, and draws much of its leadership and support base from the elite upper castes. We also conduct the same test using the Indian National Congress (INC), which is India's other major political party. The results of these two placebo tests are presented in Table 2.13. The results suggest that exogenous changes in either BJP or INC representation at the state level have no effect (and perhaps even a modest minority crime-increasing effect).

2.8.2 Overall Crime Trends

We also have not ruled out the possibility that minority representatives simply fight crime more aggressive, and are not overly concerned with the protection of minority civil and human rights protection. Indeed, it is possible that the decline in documented crimes against low-caste communities could be part of an overall decline in crimes following increased representation. This could occur, for example, if increased political presence for AA parties in state legislatures led to a general increase in the focus on law and order (not just deterrence of crimes targeted towards lower castes) through policies such as changes to policing strategy, or a change in overall economic trends that impacts crime. To test this possibility, we examine the impact of a (quasi-random) increase in AA party representation on crimes in which the victims are members of the general population. We first examine the impact on overall levels of economic crimes (counterfeiting, cheating, and breach of trust), and crimes against property (e.g., robberies and burglaries). The literature on caste-based discrimination suggests SC/ST are not typically targeted for the latter two types of crimes. Unfortunately, data for these two crime categories were only available until 2007, so the panel is not as complete as for crimes committed against SC/ST individuals.

Table 2.14b provides the results for this exercise. The results suggest that an additional AA party representative has no effect on overall economic offenses or property crimes: there is no positive or negative statistically significant relationship between an increase in AA party legislators and any of the three crime categories. The point estimates are always smaller than 0.01. The results give us some confidence that our main results in Tables **??** and 2.4 capture the impact of enhanced political representation of marginalized groups on crimes against HMCs, rather than the effect of other concurrent reforms that may have been implemented, which could be affecting overall crime rates. We also have greater confidence

that the observed effect is not simply a secular declining trend in overall crime.

In Table 2.14a, we perform a similar exercise, using the same set of violent crimes in our main results, except looking at overall levels of these crimes rather than the subset where there is a HMC victim and a high caste perpetrator. The estimate in Column (1) presents the effect of an AA party legislator on overall crime as measured by a z-score index of murder, assault, rape, and kidnapping. This coefficient is not statistically different from zero. Moreover, we again see no economically or statistically significant effect on any individual crime categories in Columns (2)-(5) (murder, assault, rape, or kidnapping), with the sign coefficients fluctuating across crime categories.[38]

2.9 Conclusion

This paper provides evidence on the social effects of achieving a political voice for historically disadvantaged minorities. We estimate the impact of increased political representation through active party mobilization for India's Scheduled Castes and Tribes on violent crimes committed against members of these groups. To address concerns related to endogeneity, we exploit the quasi-random variation arising from close constituency elections, aggregating these elections to the state-level and controlling for the expected number of wins. Using this plausibly exogenous state-level variation, we can isolate the effect of a "random" increase in the number of elected AA officials. Using this variation, we find that state-level political representation has the effect of reducing the incidence of certain types of violent crimes in particular, kidnapping, assault, and rape against HMCs. These results are consistent with the possibility that political voice has a deterrent effect. We probe the robustness of these results in a few different ways.

Our results are broadly consistent with the deterrence of potential caste-crime perpetrators due to increased representation. Our results suggest one major reason that potential criminals may decline to commit crimes against low caste and indigenous groups is that the government is now more responsive to reports of these crimes: states with greater low caste representation arrest the perpetrators of such crimes at higher rates, and prosecute these cases with greater vigor. Given our data limitations, though, we are unable to test for all potential mechanisms that may explain exactly how political representation may deter criminals, or if competing explanations are also at play. For example, we are unable to isolate whether a crime-reducing effect of representation is dampened by a "retaliation effect" that would increase crimes against members of HMCs as their political voice increases. We might expect that in areas where elections are close but tilt toward greater AA party representation, members of higher castes who do not support AA parties are more likely to express resentment toward the constituents of AA party representatives. When states endogenously select more AA party representatives, though, we would expect the retaliation effect to be small. Thus, a retaliation effect is less likely to push the OLS results in a more positive direction relative to when more AA party representation is assigned in a "like random" manner. Given our results, we can conclude that on average, any retaliation is outweighed

[38] It is worth noting that the coefficient on these crimes is negative, although it is possible that this is a mechanical by-product of AA party representatives reducing the incidence of violence against HMCs.

by deterrence and other channels that lead to reductions in identity-based crime.

Finally, measurement error is a concern for all crime data, but in particular for data such as rape and assault, where there is abundant qualitative evidence that rape is underreported, particularly by HMCs (Desphande 2011). In the presence of classical measurement error, this would likely bias our estimate toward zero. The fact that we still observe an economically significant coefficient for several crime categories may suggest to us that AA legislators are indeed having some positive effect on the incidence of some violent crimes. Future work may be useful to test the validity of the results here using more reliable data, such as data from victimization surveys.

2.10 Figures

Figure 2.1: McCrary Test for Discontinuity of AA Victory Margin

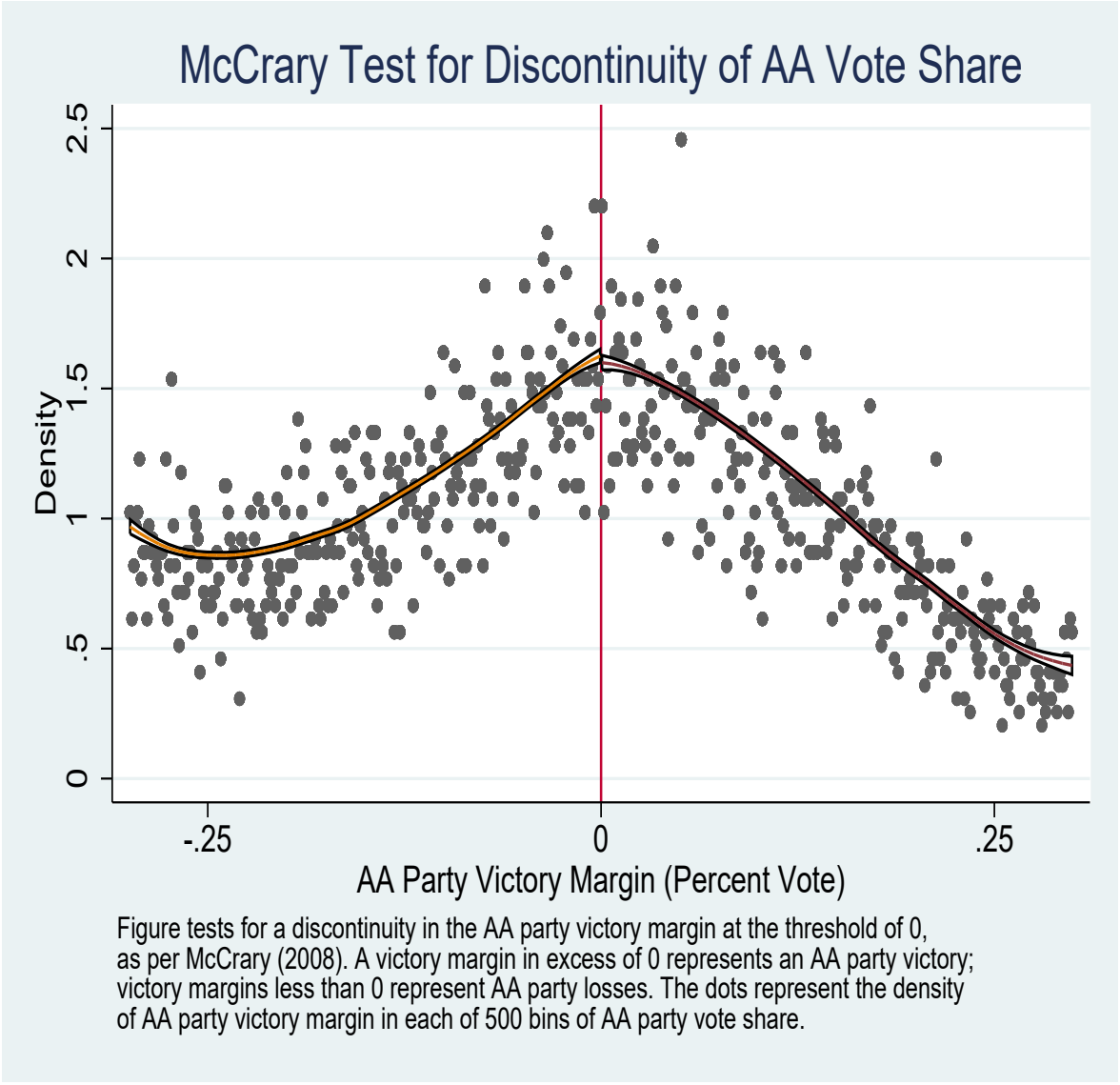
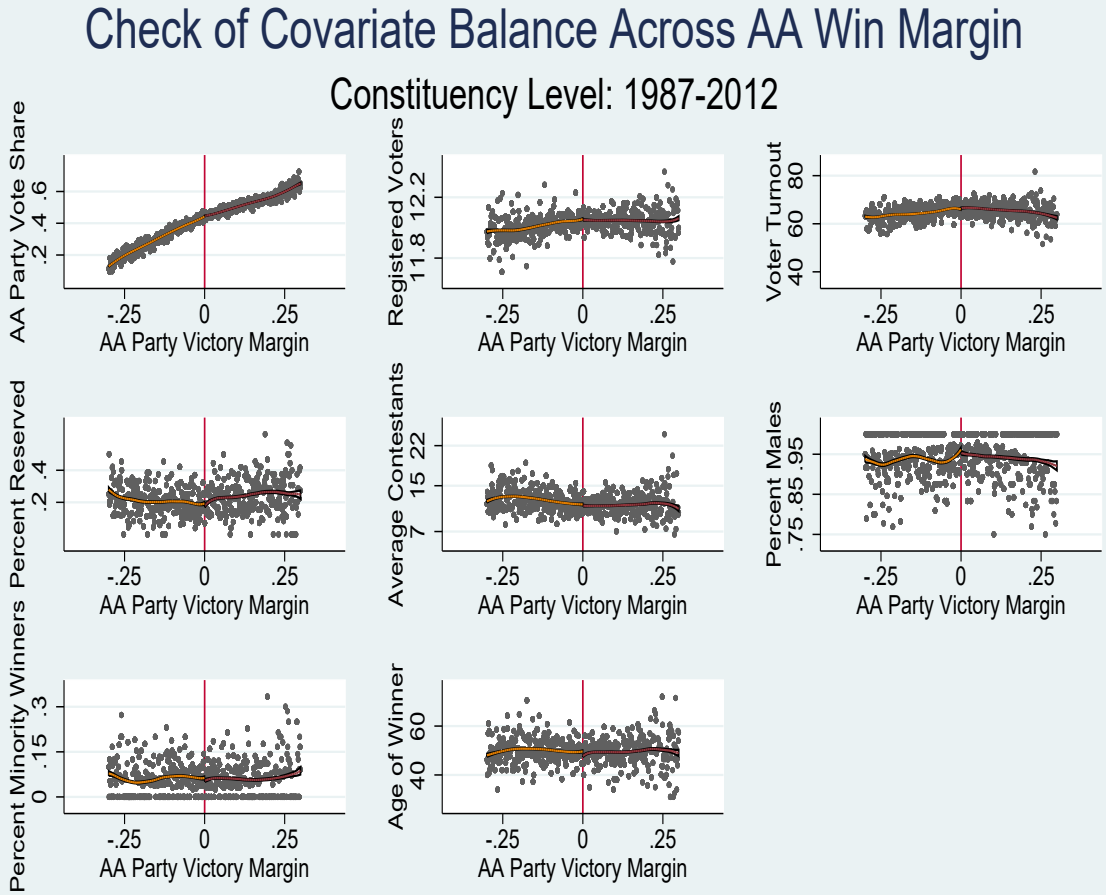
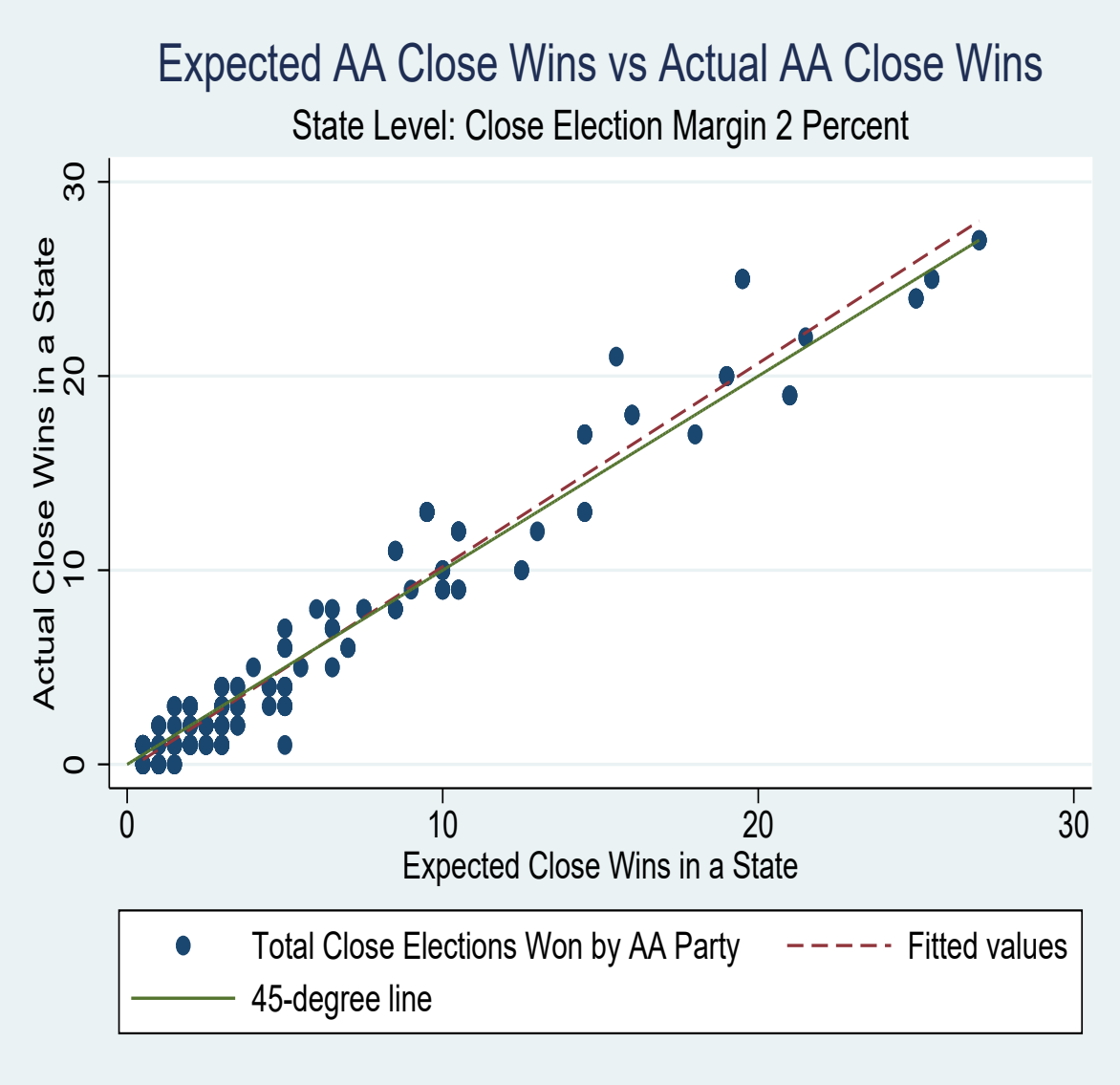


Figure 2.2: Check of Covariate Balance Across AA Win Margin



Covariate balance checks at the constituency level. The following covariates are tested: AA party vote share; number of registered voters; voter turnout; share of constituencies reserved for SC/ST candidates; number of candidates contesting; share of male winners; share of winners from SC/ST community; share of male winners. Each dot corresponds to unconditional means across 500 bins of AA party vote share. The solid lines represent 95 percent confidence intervals.

Figure 2.3: Expected Close AA Wins vs Actual Close AA Wins - 2 Percent Margin



2.11 Tables

Table 2.1: Summary Statistics of Key Variables

Variable	Mean	Std. Dev.	Min.	Max.	N
SC/ST Murder Rate per 100000	32.28	2.07	0.00	12.08	295
SC/ST Kidnap Rate per 100000	18.45	1.59	0.00	9.10	295
SC/ST Rape Rate per 100000	72.29	9.17	0.01	91.65	295
SC/ST Assault Rate per 100000	209.76	18.73	0.10	95.37	295
SC/ST Hate Crime Rate per 100000	30.72	5.87	0.02	49.58	295
Crime Index (Log)	0.67	2.86	-11.66	4.32	295
Total Elections in State	195.32	96.24	68	425	295
Total Close Elections Involving AA Party, 2 pc margin	8.85	11.94	0	54	295
Total Close Elections Won by AA Party, 2 pc margin	4.35	6.26	0	27	295
Proportion of Constituencies Reserved for SC/ST	0.26	0.08	0.1	0.48	295
State Voter Turnout	0.64	0.08	0.22	0.79	295
State Vote Share of AA Parties	0.22	0.17	0.01	0.67	295
Proportion of Elections Won by AA Party	0.17	0.2	0	0.87	295
Proportion of Elections Won by BJP Party	0.23	0.21	0	0.75	295
Proportion of Elections Won by INC Party	0.31	0.19	0	0.76	295
Average Candidates Contesting Election	10.17	4.38	4.82	28.07	295
Proportion of SC/ST Police	0.21	0.07	0	0.41	212

Table 2.2: OLS Results: Caste-party Representation and Crimes Against SC/STs, 1994-2012

	(1) Crime Index	(2) Murder	(3) Kidnapping	(4) Rape	(5) Assault	(6) CRA
Total AA Wins	-.0067***	-.0009	-.0058*	-.0001	-.0085**	-.0054
	(.0023)	(.0027)	(.0030)	(.0020)	(.0038)	(.0083)
Observations	295	295	295	295	295	295
R^2	.7198	.5984	.5546	.7437	.6904	.7222
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	15.65	32.28	18.45	72.29	209.76	30.72

Standard errors in parentheses
Errors are clustered at the state level. State and year fixed effects included.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.3a: Verifying State Level Demographic Characteristics Do Not Predict AA Close Wins at 2 Percent Margin

	(1) Close AA Wins	(2) Close AA Wins	(3) Close AA Wins	(4) Close AA Wins	(5) Close AA Wins	(6) Close AA Wins
Population (log)	.0592					
	(.1877)					
Share of SC/ST		-1.7345				
		(1.3294)				
Gender Ratio			2.0742			
			(2.9473)			
Literacy Rate				-.3743		
				(1.1781)		
Share of Workers					.0827	
					(.2586)	
Rate of Urbanization						-.6174
						(1.4306)
Observations	295	295	295	295	295	295
R^2	.9649	.9652	.9651	.9649	.9649	.9649
F-test	1307.0973	1319.9204	1547.9942	1166.3829	1303.7202	1187.6600

Standard errors in parentheses (clustered at the state level). Dependent variable for each column is percent of close elections won by AA parties at the 2 percent margin. All specifications control for the total number of close elections contested by AA parties at the 2 percent margin as well as state and year fixed effects.

Table 2.3b: Verifying State Level Political Characteristics Do Not Predict AA Close Wins at 2 Percent Margin

	(1) Close AA Wins	(2) Close AA Wins	(3) Close AA Wins	(4) Close AA Wins	(5) Close AA Wins	(6) Close AA Wins
State Vote Share of AA Parties	-.7846					
	(1.5674)					
State Voter Turnout		.2796				
		(2.1673)				
Effective Number of Parties			-.0299			
			(.2404)			
Reserved Elections for SC/ST				-1.4477		
				(1.1842)		
Total AA Wins - Current					.2225	
					(.6577)	
Total AA Wins - Lagged						-.5074
						(.6728)
Observations	83	83	83	83	83	65
R^2	.9626	.9624	.9624	.9627	.9624	.9692
F-test	2271.0197	962.7664	964.1766	1019.4798	1081.6858	1629.8780

Standard errors in parentheses (clustered at the state level). Dependent variable for each column is percent of close elections won by AA parties at the 2 percent margin. All specifications control for the total number of close elections contested by AA parties at the 2 percent margin. The regressions are conducted only for the years in which a state had an election

Table 2.4: Quasi-Experimental Results: Caste-party Representation and Crimes Against SC/ST

			Panel A: with Baseline Controls			
	(1)	(2)	(3)	(4)	(5)	(6)
	Crime Index	Murder	Kidnap	Rape	Assault	CRA
Total Close AA Wins	-.2333***	-.0799**	-.0788	-.0357	-.1365*	-.3173***
	(.0504)	(.0373)	(.0551)	(.0251)	(.0675)	(.1089)
R^2	.7251	.6709	.6209	.7743	.6751	.7069
Controls	Yes	Yes	Yes	Yes	Yes	Yes
			Panel B: No Baseline Controls			
	(1)	(2)	(3)	(4)	(5)	(6)
	Crime Index	Murder	Kidnap	Rape	Assault	CRA
Total Close AA Wins	-.2115***	-.0452	-.0646	-.0366	-.1394*	-.2911**
	(.0509)	(.0295)	(.0680)	(.0225)	(.0665)	(.1221)
R^2	.7168	.6562	.5966	.7639	.6718	.6780
Controls	No	No	No	No	No	No
Observations	295	295	295	295	295	295
Dependent Variable Mean	15.65	32.28	18.45	72.29	209.76	30.72

Standard errors in parentheses, clustered at the state level. State and year fixed effects included. The dependent variable in each specification is the log crime rate against SC/STs. Explanatory variable is the number of close (2%) AA party wins. In Panel A, estimation controls for the total number of close contests and state-level characteristics at baseline.

Table 2.5: Predicted Caste Party Representation and Crime Against SC/STs - Altonji-Elder-Taber (2005) Test

	(1) Crime Index	(2) Murder	(3) Kidnap	(4) Rape	(5) Hurt	(6) CRA
Predicted Close AA Wins	.0551	.0664**	.0492	.0134	.0124	.0413
	(.0389)	(.0229)	(.0349)	(.0261)	(.0215)	(.0415)
Observations	294	294	294	294	294	294
R^2	.0521	.1515	.0647	.0065	.0046	.0092

Standard errors in parentheses (clustered at the state level). The dependent variable in each specification is the logged rate of crimes (crime per million of population) targeted against SC/STs. Predicted AA close wins is the fitted values from a regression where predicted AA close wins is regressed on share of SCs and STs; total number of close contests; state-level political competition; rate of urbanization; gender ratio; total elections in the state; share of elections won by the national parties BJP and INC; rate of literacy and the share of workers in the population, in addition to state and year fixed effects. The results are weighted by state's population. The regression tests whether the variation in close AA wins predicted by the fixed effects and controls explains the variation in SC/ST crime rates.

Table 2.6: Caste-Party Representation and Crimes Against SC/STs - Accounting for Electoral Cycle Structure

			Panel A: Clustering on State $*$ Year			
	(1)	(2)	(3)	(4)	(5)	(6)
	Crime Index	Murder	Kidnap	Rape	Assault	CRA
Total Close AA Wins	-.2333***	-.0799**	-.0788	-.0357	-.1365**	-.3173***
	(.0634)	(.0354)	(.0542)	(.0297)	(.0559)	(.1089)
Observations	295	295	295	295	295	295
R^2	.7251	.6709	.6209	.7743	.6751	.7069
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	15.65	32.28	18.45	72.29	209.76	30.72
			Panel B: Collapsing Data Across Election Cycles			
	(1)	(2)	(3)	(4)	(5)	(6)
	Crime Index	Murder	Kidnap	Rape	Hurt	CRA
Total Close AA Wins	-.1142*	-.0935	-.1322	-.0402	-.1253	-.5074***
	(.0547)	(.0615)	(.0879)	(.0286)	(.0730)	(.1580)
Observations	83	83	83	83	83	83
R^2	.8940	.8595	.8920	.9225	.8447	.9405

Standard errors in parentheses, clustered at the state level. State and year fixed effects included. The dependent variable in each specification is the log crime rate against SC/STs. Explanatory variable is the number of close (2%) AA party wins. Estimation controls for the total number of close contests and state-level characteristics at baseline.

Table 2.7: Interaction between Marginal AA Party Representative & AA Party Control

	(1) Crime Index	(2) Murder	(3) Kidnap	(4) Rape	(5) Assault	(6) CRA
Total Close AA Wins	-.0959**	-.0611**	-.0439	.0008	-.0897**	-.3321**
	(.0343)	(.0288)	(.0408)	(.0256)	(.0418)	(.1326)
AA Party in Power	-.2915	-.4073*	-.0910	-.0171	-.3183	-1.1766**
	(.3699)	(.2337)	(.3177)	(.2579)	(.1967)	(.5243)
Marginal AA Rep X AA Party Control	-.0294*	-.0232*	-.0173	-.0085	-.0248**	.0924***
	(.0169)	(.0115)	(.0136)	(.0096)	(.0112)	(.0296)
Observations	295	295	295	295	295	295
R^2	.7346	.6726	.6225	.7745	.6754	.7171
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	1.57	3.23	1.85	7.23	20.98	3.07

Standard errors in parentheses, clustered at the state level. State and year fixed effects included. The dependent variable in each specification is the log crime rate against SC/STs. Explanatory variables are the number of close (2%) AA party wins, a dummy variable for whether an AA party controls the state legislatures, and the interaction between these two. Estimation controls for the total number of close contests and state-level characteristics at baseline.

Table 2.8: Robustness Checks: Political Controls, Alterations to Victory Margin Threshold and Alternative Specification of Exogenous Variation

			Panel A: Controlling for Political Competition			
	(1) Crime Index	(2) Murder	(3) Kidnap	(4) Rape	(5) Assault	(6) CRA
Total Close AA Wins	-.2218***	-.0568*	-.0623	-.0218	-.1401***	-.4753***
	(.0487)	(.0296)	(.0570)	(.0381)	(.0423)	(.1115)
Observations	295	295	295	295	295	295
R^2	.7256	.6760	.6271	.7753	.6769	.7163
			Panel C: Win Margin-Less than 1%			
	(1) Crime Index	(2) Murder	(3) Kidnap	(4) Rape	(5) Assault	(6) CRA
Total Close AA Wins	-.1930***	-.0654**	-.0530	-.0045	-.0976***	-.3024**
	(.0432)	(.0267)	(.0527)	(.0280)	(.0287)	(.1266)
Observations	295	295	295	295	295	295
R^2	.7246	.6697	.6202	.7736	.6722	.7092
			Panel B: Win Margin-Less than 3%			
	(1) Crime Index	(2) Murder	(3) Kidnap	(4) Rape	(5) Assault	(6) CRA
Total Close AA Wins	-.2737**	-.1832**	-.1213	-.0626	-.2820***	-.1068
	(.1192)	(.0665)	(.1131)	(.0546)	(.0885)	(.2237)
Observations	295	295	295	295	295	295
R^2	.7185	.6788	.6197	.7753	.6795	.6983
			Panel D: Dependent Variable - Wins Above Expectation			
	(1) Crime Index	(2) Murder	(3) Kidnap	(4) Rape	(5) Assault	(6) CRA
AA Wins Above Exp.	-.265**	-.151	-.102	-.0839	-.224*	-.309
	(.106)	(.0919)	(.105)	(.0707)	(.116)	(.215)
Observations	240	268	268	268	240	268
R^2	.6799	.6915	.5984	.7839	.6480	.7033

Standard errors in parentheses (clustered at the state level). State and year fixed effects included in each specification. Dependent variables for each column is natural log crime rate against SC/STs. The explanatory variable in Panel A is the number of close (2%) AA party wins; the explanatory variable in Panel B is the number of close (1%) AA party wins; the explanatory variable in Panel C is the number of close (3%) AA party wins; explanatory variable in Panel D is the total number of close (2%) AA party wins, less half the total number of close AA party contests.

Table 2.9: Caste Party Representation, Inter-personal Conflict, and Trust in Government

	Panel A: SC/ST Respondents			
	(1) Caste Conflicts	(2) Trust-Politicians	(3) Trust-Police	(4) Trust-Courts
Total Close AA Wins	-.0592***	.0201**	.0204**	.0476***
	(.0112)	(.0080)	(.0102)	(.0091)
Observations	21240	21275	21267	21298
R^2	.6382	.5977	.6234	.6615
	Panel B: Upper Caste Respondents			
	(1) Caste Conflicts	(2) Trust-Politicians	(3) Trust-Police	(4) Trust-Courts
Total Close AA Wins	-.0266*	.0042	.0109	.0287**
	(.0147)	(.0103)	(.0137)	(.0139)
Observations	16278	16275	16272	16019
R^2	.7026	.6805	.6840	.7340

Standard errors in parentheses (clustered at the household level). The dependent variable in each specification is a dummy representing whether either a SC/ST respondent (Panel A) or a non-SC/ST respondent (Panel B) reported caste conflicts in the village (Column 1), or whether they reported having confidence in public institutions (Columns 2-4). The explanatory variable of interest is the number of close elections won by AA parties at the 2 percent margin in the state. Estimation controls for household level of education, urban location, household income, total number of close contests, political competition, total elections in the state, and the fraction of constituencies reserved for SC/ST politicians, in addition to household and survey year fixed effects. The results are weighted by the household weights assigned by the IHDS.

Table 2.10: Caste-party Representation and Police Activity for Crimes Against SC/STs

	(1) Murder	(2) Kidnap	(3) Rape	(4) Assault
Total AA Close Wins	.0456***	.0557***	.0334***	.0016
	(.0060)	(.0085)	(.0048)	(.0064)
Observations	210	210	210	210
R^2	.6424	.5947	.7513	.7717

Standard errors in parentheses (clustered at the state level). Dependent variable for each column is natural log of the arrest rate for for a given crime against SC/STs. The explanatory variable is the number of close (2%) AA party wins. Estimation controls for the total number of close contests, state-level characteristics at baseline, as well as state and year fixed effects.

Table 2.11: Effect of AA Party Representation on SC/ST Police Presence

	(1) SC/ST Police Rate	(2) SC/ST Police Composition	(3) Overall Police Rate
Total AA Close Wins - 2	.0626	.0742	-.0149*
	(.0698)	(.0678)	(.0084)
Observations	225	225	225
R^2	.4824	.4450	.9224

Standard errors are clustered at the state level. Dependent variables for column (1) are the logged rate of Scheduled Caste or Tribe police officers per 100,000 SCST people and (2) the logged fraction of Scheduled Caste or Tribe police officers as a fraction of the total police. Explanatory variable is the number of close constituency elections won by AA parties, where close elections are contests decides by less than 2%. Estimation controls for the total number of close contests, state-level factors, as well as state and year fixed effects.

Table 2.12: Caste-party Representation and Court Activity for Crimes Against SC/STs

	(1) Convictions	(2) Acquittals	(3) Def.'s Awaiting Trial
Total AA Close Wins - 2	.0333	-.0731**	-.0252
	(.0317)	(.0338)	(.0271)
Observations	210	216	217
R-square	.8967	.6230	.9660

Standard errors in parentheses. Errors clustered at the state level. State fixed effects included.

Table 2.13: Major Party Representation and Crimes Against SC/STs

			Panel A: Impact of BJP			
	(1) Crime Index	(2) Murder	(3) Kidnap	(4) Rape	(5) Assault	(6) CRA
Total BJP Close Wins - 2	.0356	.0444	-.0131	-.0043	.0584	.0106
	(.0346)	(.0333)	(.0409)	(.0242)	(.0387)	(.1215)
R^2	.7142	.6682	.6214	.7737	.6741	.6976
			Panel B: Impact of INC			
	(1) Crime Index	(2) Murder	(3) Kidnap	(4) Rape	(5) Assault	(6) CRA
Total INC Close Wins	.0665	.0123	.0101	.0169	.0309	.0570
	(.0581)	(.0262)	(.0270)	(.0155)	(.0261)	(.0805)
R^2	.7187	.6665	.6198	.7759	.6792	.7007
Observations	295	295	295	295	295	295
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	1.57	3.23	1.85	7.23	20.98	3.07

Standard errors in parentheses, clustered at the state level. State and year fixed effects included. The dependent variable in each specification is the log crime rate against SC/STs. Explanatory variable is the number of close (2%) BJP/INC party wins. In Panel A, estimation controls for the total number of close contests and state-level characteristics at baseline.

Table 2.14a: Caste-party Representation and Overall Violent Crimes

	(1) Crime Index	(2) Murder	(3) Kidnap	(4) Rape	(5) Assault
Total AA Close Wins - 2	-.0473	-.0069	.0056	.0223	-.0248
	(.0846)	(.0097)	(.0220)	(.0205)	(.0285)
Observations	288	305	305	305	288
R^2	.8734	.9433	.9474	.9505	.9207

Standard errors in parentheses, clustered at the state level. State and year fixed effects included. The dependent variable in each specification is the log overall crime rate. Explanatory variable is the number of close (2%) AA party wins. Estimation controls for the total number of close contests and state-level characteristics at baseline.

Table 2.14b: Caste-party Representation and Overall Economic/Property Crimes

	(1) Overall Property Crimes	(2) Overall Economic Crimes
Total Close AA Wins	.0084	.0295
	(.0244)	(.0274)
Observations	208	208
R^2	.8758	.8949

Standard errors in parentheses, clustered at the state level. State and year fixed effects included. The dependent variable in each specification is the log overall crime rate. Explanatory variable is the number of close (2%) AA party wins. Estimation controls for the total number of close contests and state-level characteristics at baseline.

2.12 Appendix: Additional Results

Table A2.1: Caste Party Representation and Crimes Against SC/STs - IV Estimates

	(1) Crime Index	(2) Murder	(3) Kidnap	(4) Rape	(5) Assault	(6) CRA
Total AA Wins	-.0641**	-.0246	.0801**	-.0165	-.0144	-.2031**
	(.0395)	(.0233)	(.0327)	(.0212)	(.0308)	(.0936)
Observations	295	295	295	295	295	295
R^2	.7143	.6543	.5338	.7666	.6788	.6520

Standard errors in parentheses (clustered at the state level). The dependent variable in each specification is the logged rate of crimes (crime per million of population) targeted against SC/STs. Total elections won by AA parties in a state is instrumented by the total number of close elections at the 2 percent margin won by AA parties in the state. Estimation controls for the share of SCs and STs; total number of close contests; state-level political competition; rate of urbanization; gender ratio; total elections in the state; rate of literacy and the share of workers in the population, in addition to state and year fixed effects. The results are weighted by state's population.

Table A2.2: Interaction: Marginal Representative & AAParty Vote Share

	(1) Crime Index	(2) Crime Index	(3) Crime Index
Random AA Rep.	.0817	.0044	-.0321
	(.3207)	(.2736)	(.2619)
AA RepXHigh VS-10%	-.4380		
	(.3836)		
AA RepXHigh VS-15%		-.3691	
		(.3403)	
High VS-10%		-.1442	
		(1.5084)	
AA RepXHigh VS-20%			-.2609
			(.3359)
High Vote Share-20%			-1.6281
			(1.2644)
Constant	-7.6878	-23.7629	-25.7924
	(108.3994)	(99.0949)	(100.3436)
Observations	295	295	295

Standard errors in parentheses. Errors are clustered at the state level. Regressions include state and year fixed effects, controls for total close contests, and state-level characteristics.

Table A2.3: Testing for Non-Linearities in Effect of Caste Party Representation on Crime Against SC/STs

	(1) Crime Index	(2) Murder	(3) Kidnap	(4) Rape	(5) Hurt
Total Close Elections Won by AA Party, 2 pc margin	-.0272	-.0546*	.1086	-.0308	-.0904
	(.0591)	(.0272)	(.0856)	(.0278)	(.0525)
Close AA Wins*Close AA Wins Under 3	-.4270	-.5136	.0227	-.3440	-.1741
	(.7636)	(.6166)	(.5787)	(.4527)	(.5331)
Close AA Wins*Close AA Wins, 3-6	-.2089	-.0349	-.1206	-.0075	-.0733
	(.1931)	(.1016)	(.1240)	(.0904)	(.1318)
Observations	295	295	295	295	295
R^2	.7335	.6732	.5932	.7787	.6827

Standard errors in parentheses (clustered at the state level). The dependent variable in each specification is the crime index (logged) representing a measure of aggregated crimes targeted against SC/STs. Explanatory variable is the number of close (2%) AA party wins and its interactions with binary variables representing measures of the total number of close AA wins. Estimation controls for the total number of close contests, state-level characteristics at baseline, as well as state and year fixed effects

Chapter 3

Agricultural Productivity Growth and Rural Non-Farm Employment: Evidence from India

3.1 Introduction

The process of structural transformation has long been held by economists as a benchmark for economic development. It has been widely noted that as economies develop, there is an outflow of labour from the low productivity primary sector to the secondary and tertiary sectors which have higher levels of productivity. Increased employment of the labour force in higher productivity non-farm sectors leads to higher earnings, facilitating the process of economic development, growth and poverty alleviation. The contribution of off-farm employment to the process of poverty alleviation has been highlighted in multiple studies[1] which document how the rural poor enage in employment diversification to supplant farm incomes, which are not only lower than non-farm incomes but also subject to greater variance due to the dependence of agricultural production on weather in the majority of developing countries. The diversification of employment for rural workers from farm to non-farm therefore can serve both as a source of higher incomes and also a form of insurance against variable weather shocks to which the rural poor are disproportionately exposed (Dell et al. 2012).[2]

Recent research has also attempted to identify barriers to the process of structural transformation or strategies which can aid the rural poor to take advantage of non-farm opportunities offering a higher return on their labour. Banerjee and Duflo (2007) for instance cite the lack of capital endowments, combined with limited borrowing capabilities, as the reason for which the rural poor are either unable to shift to non-farm entrepreneurship or expand existing rural non-farm enterprises into profitable firms with have higher levels of productivity and firm size. Reardon et al. (2001) propose that the ability of the poor

[1] See for instance Banerjee and Duflo (2007); Christiansen et al. (2013); de Janvry and Sadoulet (2001); Himangshu et al. (2011); Kijima et al. (2006); Lanjouw and Lanjouw (2001); and Lanjouw and Murgai (2009).

[2] This has become of increasing importance in an era of rapid climate change.

to participate in rural non-farm opportunities can be enhanced through the imparting of education, skills and also the provision of credit and infrastructure for rural workers. Bryan et al. (2014) in their study in Bangladesh show that cash interventions incentivizing seasonal migration amongst rural agricultural workers can lead to the overcoming of barriers to migration, which are associated with high levels of risk aversion amongst the rural poor due to significant costs resulting from migration failure.

The present paper identifies whether growth in agricultural productivity can facilitate the creation of non-farm employment opportunities for rural workers using panel data from Indian districts. India provides a relevant context, offering a setting with low urbanization, combined with over two-thirds of the rural labour force being engaged in farm activities.[3] This becomes even more salient as the primary sector's share in national income over the past two decades has fallen below 20 percent, indicating low returns from farm activities and amplifying the role non-farm employment can play in raising the earnings of farm workers.

I collect data on agricultural production and employment outcomes from a panel of 273 Indian districts between 1987 and 2011 to identify the impact of variations in agricultural productivity on rural non-farm employment (RNFE). To obtain a district-level measure of agricultural productivity, I construct a composite productivity index aggregated across seven major crops, with district-level productivity growth in each crop being weighted by the long-run acreage allocated to the crop in the district. To generate exogenous variation in agricultural productivity, I instrument agricultural productivity with standardized deviations in the district's monsoon rainfall from the district's historical mean. Causal identification is based on changes to agricultural productivity caused due to variations in the district's monsoon rainfall realization, relative to the district's long-run average monsoon rainfall realization, akin to Jayachandran (2006).

Using the IV specification, I show that positive exogenous rainfall shocks to agricultural productivity has a significant impact on RNFE in the short-term. A rainfall-induced 1 percent increase in the district's agricultural productivity increases the share of rural workers employed in the manufacturing sector by 2 percent, across both male and female workers. A percent increase in agricultural productivity also causes a 2 percent increase in the share of rural female workers employed in the services sector. I also show that this increase in non-farm employment for rural workers is concentrated in unskilled activities, in both the manufacturing and the services sectors. Specifically, I show that for rural female workers, the unskilled manufacturing employment generated in response to higher agricultural productivity is located within the food and textiles sector. My results suggests that the increase in non-farm employment due to agricultural productivity for rural male workers is through salaried jobs while that for female workers is through self-employment.

Using a reduced form specification and testing for the impact of historical rainfall shocks on RNFE, my paper shows that the impact of rainfall-induced agricultural productivity on RNFE is limited to short-term effects. Consistent with the IV specification, the reduced form results show that rainfall shocks lagged by 1 time period have a positive effect on RNFE but none of the other 4 historical lagged rainfall shocks have any effect on RNFE. The coefficients

[3] Foster and Rosenzweig (2010) estimates that the surplus labour in the agricultural in India is 20 percent.

are also not jointly significant, ruling out any cumulative growth effects of rainfall shocks on RNFE as identified by Dell et al. (2012).

Through a disaggregation of my sample across household landholdings and workers' educational attainment, my paper also identifies the distributional implications of the RNFE generated in response to growth in agricultural productivity. I show that agricultural productivity growth impacts RNFE for rural workers across both large and small landowning households, but only for workers who have not completed secondary education. As only a fourth of rural households have landholdings in excess of 1 acre and under a third of rural workers have completed secondary education, these results signify that the RNFE opportunities generated by exogenous weather shocks to agriculture can be accessed by the majority of rural workers. By testing for the differential effects of agricultural productivity across districts' rates of urbanization and agricultural productivity prior to 1987, I show that the RNFE created through growth in agricultural productivity is concentrated in districts which had initially low levels of urbanization and agricultural productivity. This testifies that growth in agricultural productivity can lead to convergence across districts in terms of RNFE generation.

The paper also considers the potential channels explaining the IV results. I first show that growth in agricultural productivity has no impact on the share of rural workers working in the farm sector, ruling out the 'labour push' channel as a potential pathway explaining the results (Alvarez-Cuadrado and Poschke, 2011). I subsequently show that rainfall-induced shocks to agricultural productivity increases real rural wages in the manufacturing sector and is restricted to workers without secondary education. This indicates that the impact of agricultural productivity on RNFE identified in this paper is occuring through the aggregate demand channel whereby higher farm productivity increases farm incomes and leads to higher demand for locally produced goods and services, causing an increase in labour demand in these sectors.

The results obtained through the IV strategy are robust to alternate specifications of the rainfall instrument. I verify that the results are unchanged if agricultural productivity is instrumented using annual rainfall deviations and are also robust to a flexible functional form of the instrument. Finally, I re-estimate my results using district rice yields as the endogenous variable of interest and show that the results are stable to this alternate measure of agricultural productivity.

The paper adds to a growing literature studying the process of structural transformation and the impact of agricultural growth on non-farm employment.[4] The paper extends this literature by mapping the specific non-farm sectors which are affected by productivity shocks to agriculture and shows that the RNFE generated in response to such shocks is significantly greater for female than male workers. The paper also rules out any long-run impacts of rainfall shocks on rural non-farm employment. In this regard, the paper is closest to the recent works by Emerick (2016) and Colmer (2016). However, unlike Colmer (2016) who shows that negative temperature shocks to farm yields reduces labour demand in the agricultural

[4] For instance, Adelman and Vogel (1992); Alvarez-Cuadrado and Poschke (2011); Bustos et al. (2016); Cao and Birchenall (2013); de Janvry and Sadoulet (2010); Emerick (2016); Foster and Rosenzweig (2004); Gollin et al. (2002); Hazell and Roell (1983); Lanjouw and Lanjouw (2001); Mellor (1995); Datt and Ravallion (1998); Datt and Ravallion (2010).

sector and increases the employment of rural workers in the manufacturing sector, my paper estimates the RNFE generated through a positive rainfall shock to farm yields.

The remainder of the paper is structured as follows: Section 2 discusses the potential pathways through which agricultural productivity can impact RNFE; Section 3 presents the data used in the paper; Section 4 lays out the empirical strategy; Section 5 presents the core results of the paper; Section 6 provides a set of robustness checks; and finally, Section 7 offers some conclusions.

3.2 Conceptual Framework

The structural transformation literature documents three major channels explaining the outflow of farm labour into non-farm activities - these are namely the 'labour push', 'labour pull' and the aggregate demand channels. Alvarez-Cuadrado and Poschke (2011) define 'labour push' as rural non-farm employment generated when surplus labour leaves the primary sector as a threshold level of productivity is attained to satisfy the economy's food consumption. A similar model to the 'labour push' channel is explored by Bustos et al. (2016) who study the introduction of a new soybean variety in Brazil. The soybean variety as explained by Bustos et al. (2016) was comparable to a labour-saving technology, which resulted in the release of agricultural labour to the non-farm sector. In contrast, the 'labour pull' channel is the creation of non-farm employment opportunities due to a productivity shock in the non-farm sector - independent of any changes to farm yields - which raises non-farm wages and attracts farm labour into non-farm activities (Alvarez-Cuadrado and Poschke, 2011).

The third pathway of structural transformation is the aggregate demand channel discussed by Adelman and Vogel (1992), de Janvry and Sadoulet (2009), Gollin (2010), Himangshu et al. (2011), Irz et al. (2001), Mellor (1995), and others. This channel works through an increase in demand for consumption goods and services by farm households as a consequence of higher agricultural incomes emanating from higher farm yields. This has a direct effect on rural non-farm employment if the services demanded are non-tradables produced locally using labour-intensive technologies. (Hazell and Roell, 1983). An extension of this framework is considered by Gollin et al. (2002) who hypothesize that exogenous technological shocks in developed countries lead to higher agricultural production in developing countries. The increase in yields subsequently leads to greater domestic demand for industrial goods and causes an outflow of labor from the agricultural to the manufacturing sector.

In addition to rising consumer demand as an upshot of increasing agricultural incomes, there also exist forward and backward linkages between the farm sector and select manufacturing industries and services. For instance, in terms of forward linkages, both food and cash crops form the key input for multiple textile, agricultural and food processing industries. Subsequently, enhanced agricultural production can facilitate an expansion in these manufacturing industries, generating new employment opportunities. In the form of backward linkages, Mellor (1995) discusses the increasing role of mechanization in agriculture as a consequence of agricultural growth. This results in higher demand for machine tools and other capital goods, many of which are often locally produced in small and medium enterprises, promoting the growth of local rural industries. Similarly, the use of fertilizers

and multiple varieties of seeds as agricultural inputs contribute to the development of rural retail stores and other agricultural services. In Section 3.6, I discuss the role of each of these channels in explaining my core results.

A contrary model to the ones discussed above was developed by Foster and Rosenzweig (2004), who posit that positive shocks to farm yields increases the demand for farm labor, putting upward pressure on rural wages and reducing labor supply to the non-agricultural sector. The higher rural wages in areas of high farm productivity cause the owners of capital to invest in areas with lower rural wages - which are typically areas with low agricultural growth - to reduce labor costs, resulting in a negative relationship between agricultural growth and RNFE.[5] In light of the findings of Foster and Rosenzweig (2004), for agricultural productivity to have a net positive impact on RNFE, the positive effects of the labour-push or aggregate demand channels would have to be sufficiently large in order to surpass the negative impact of agricultural productivity on RNFE due to higher rural wages.[6]

3.3 Data

The data for this paper is compiled from two sources. The data on agricultural production is from the ICRISAT's district-level database for India which provides rainfall, output for major crops, agricultural land usage, fertilizer consumption and irrigation availability for a panel of 273 districts between 1966 and 2011. This is merged with data from the National Sample Surveys (NSS) which are nationally representative household surveys, undertaken by the Indian government. Schedule 10 of the NSS surveys records individual responses on employment and unemployment, in addition to educational attainment, sector of employment and type of occupation. This paper uses data from 6 quinquennial 'thick' rounds of the NSS surveys between 1987 and 2011,[7] each sampling approximately 100,000 households across India, to identify shifts in sectoral employment shares within districts over time. As the variation in agricultural output and rainfall incidence is measured at the level of the district, individual level data from the NSS surveys are aggregated to the level of district using the NSS-assigned household weights.[8]

To construct an annual district-level composite measure of agricultural productivity (subsequently referred to as 'yields'), the paper constructs a weighted sum of annual agricultural productivity of seven major crops - namely rice, wheat, maize, millets, sugarcane, groundnuts and pulses. Formally, the productivity measure, *Yields*, can be defined for any

[5] Conceptually though, a counter to this argument can be that low agricultural growth stymies rural income, which can subsequently thwart the development of both physical infrastructure and social capital, making such areas unattractive for investors and industries to locate in, due to transportation bottlenecks, and a shortage of skilled labor.

[6] The positive correlation between higher rainfall incidence and rural wages has also been documented by Kaur (2015).

[7] These are namely rounds 43 (1987-88), 50 (1993-94), 55 (1999-00), 61 (2004-05), 66 (2009-10) and 68 (2011-12).

[8] The NSS household weight is the inverse of the probability of a household getting sampled by the NSS.

year t and district d as:

$$Yields_{dt} = \sum_{c} \frac{Prod_{cdt}}{Prod_{c1981}} * \bar{A}_{cd} \tag{3.1}$$

In equation (3.1), c denotes the crop and *Prod* denotes productivity, defined as total output in tonnes, divided by total area in hectares. I account for 7 major crops spanning two-thirds of the total acreage in my data. The annual district-level productivity of each crop is scaled by the crop's average productivity across all districts in the year 1981, which serves as the base level of productivity. Thus, the first term in the summation exceeds 1 if the productivity of any crop in year t exceeds the national productivity for that crop in 1981. This forms a measure of growth in productivity for each crop over time, relative to the average productivity across districts in the base year. To obtain aggregate growth in district-level productivity relative to 1981, the paper sums the individual productivities of each crop, weighted by the average long-run acreage allocated to each crop in the district between 1981 and 2011 - $\bar{A}$.

Figure 3.1 plots individual yields for each crop and average annual district yields in the three decade period between 1982 and 2011. Over the past three decades, average district yields has grown at an annual rate of 1.9 percent, increasing from 0.48 in 1982 to 0.85 in 2011. The growth in yields dipped during the 1990s, falling under 1 percent and picked up subsequently in the 2000s, exceeding 3 percent. The key driver of yields are rice, wheat and maize.

To detect changes in employment trends, the paper relies on data from the NSS surveys detailing individuals' sector and type of occupation for the past year.[9] Based on individual responses, the paper constructs employment shares for broad employment categories - agriculture, manufacturing and services. In addition to the broad employment categories, employment trends in three other sectors are also observed - construction, food and the textiles sector.[10] The latter is an amalgam of both manufacturing and service activities, covering the manufacture and retail sales of food, beverages and textiles. Collectively, the construction, food and textiles sectors form the the three largest components of rural non-farm employment and the paper seperately estimates the impact of agricultural productivity on employment in these sectors.

Along with the broad employment categories, the paper also identifies whether the off-farm employment is in skilled or unskilled activities. To distinguish between skilled and unskilled work, the paper relies on the National Classification of Occupations (NCO) reported in the NSS survey for individual activities and combines it with the individual's level of education. Individuals who have not completed secondary education and individuals who work in 'elementary occupations' - services such as vending, garbage collection, mining, construction labour, domestic help, building caretaker and others - as defined by the NCO are not deemed as skilled workers. Henceforth, for the purposes of the paper, skilled workers are

[9] The NSS asks respondent their primary activity in the past year in which they have been employed for over 6 months (primary status) and any secondary activity which they might have been involved in for a period exceeding under 3 months (subsidary status). These two measured are combined, as per the NSS reports, to construct a single occupational category for the past year (usual status).

[10] The food and textile sectors are grouped together as a single sector.

those who work in non-elementary occupations and have also completed secondary education.

Figure 3.2 plots the labour force participation rates (LFP) for rural workers between 1987 and 2011 based on the six NSS survey rounds. The LFP for working-aged[11] rural males is significantly higher than rural females through the entire two decade period. While 85 percent of rural males reported participating in the labour force, the corresponding figure for rural females is 48 percent. Additionally, while the LFP for rural males has been invariant over time, the LFP for rural females has been steadily declining, falling from 50 percent in 1999-00 (round 55) to 40 percent in 2011-12 (round 68).

Conditional on participation in the labour force, Figure 3.3 plots aggregate sectoral employment for rural workers in this period. The first plot presents the trends for male workers and the second plot is for female workers. For both male and female workers, the vast majority of the workforce remains employed in the agricultural sector. While there has been a steady decline in the share of male workers employed in farm activities since 1999, the trend is much more sluggish for female workers, consistent with the observations of Gollin (2010). The decline in the share of workers in farm activities has been accompanied by an increase in the share of rural workers employed in the service sector. Employment in the manufacturing sector has remained stagnant in this period, for both males and females, with under 10 percent of the rural workforce employed in this category during the period of study. Figure 3.3 also informs that unskilled non-farm activities strictly dominates skilled non-farm activities during this period for both males and females. While the employment gap in skilled and unskilled non-farm activities has been steadily decreasing for rural male workers, this has not been the case for rural female workers.

To have a more precise understanding of the type of rural non-farm opportunities being created, Figure 3.4 plots the employment shares in the construction, and the food and textiles sectors, along with skilled and unskilled manufacturing and services. Collectively, the construction and the food and textile sectors provide employment to 55 percent of the rural non-farm workforce. While the construction sector is contained within the services sector, the food and textiles sector includes both manufacturing and service activities. From Figure 3.4, it is observed that the food and textiles sector forms the single largest source of non-farm employment for rural women while for rural men, unskilled services is the leading source of non-farm employment. While the share of men employed in the food and textiles sector has been stable in the past two decades, it has been a steadily rising source of employment for rural women. Two other leading sources of non-farm employment for rural women are the unskilled manufacturing and services sectors while rural males have seen a sharp rise in employment in the construction and skilled services sectors. Additional summary statistics for all the outcome variables are presented in Section 10.1 (Appendix).

3.4 Empirical Strategy

To empirically identify the short-term impact of growth in agricultural productivity on the structural transformation process, I regress sectoral employment shares on district yields. The primary equation of interest identifying the impact of short-term yields on non-farm

[11] This includes all individuals aged between 18 and 65.

employment can be expressed as:

$$ln(Y^{j}_{dst}) = \alpha_d + \theta_s t + \delta_t + \beta ln(Yields_{dst-1}) + \gamma X_{dst} + \epsilon_{dst} \tag{3.2}$$

In equation (3.2), Y is the employment share in sector j, measured in the NSS survey year t, and *Yields* is agricultral yields measured in the year prior to the NSS survey year, defined as in equation (3.1). Identification is conditional on district and survey round fixed effects, denoted by α and δ, in addition to the state-specific trend, $\theta_s t$. The log-log specification implies that β estimates the elasticity of district employment shares with respect to yields. X includes district-level time-varying agricultural and socio-economic covariates.[12] The identification in equation (3.2) comes from variations in yields over time, unique to the district as the state time trend captures state specific changes in employment outcomes over time. To determine the type of employment generated through increases in agricultural productivity and identify which workers are most affected, additional specifications are undertaken by disaggregating the sample of workers by educational attainment and household landholdings.

However, despite the set of fixed effects and district-level time-varying covariates, the presence of time-varying omitted variables correlated with both district non-farm employment and district yields can result in biased estimation of the coefficient β in equation (3.2).[13] To counter the threat of endogeneity between rural non-farm employment and gains in agricultural productivity, the paper instruments yields with deviations in the district's monsoon rainfall from the district's long-run mean.[14] The first stage of the instrumental variables (IV) specification can be expressed as:

$$ln(Yields_{dst-1}) = \alpha_d + \theta_s t + \delta_t + \phi StandRainDev_{dst-1} + \psi X_{dst} + \epsilon_{dst} \tag{3.3}$$

In equation (3.3), *StandRainDev* is standardized deviations in monsoon rainfall, defined as:

$$StandRainDev_{dst} = \frac{Rainfall_{dst} - \bar{Rainfall}_{ds}}{SD(Rainfall_{ds})} \tag{3.4}$$

In (3.4), *Rainfall* is the incidence of monsoon rainfall in district d, located in state s and year t. Monsoon rainfall incidence is subtracted from the district's long-run mean level of monsoon rainfall, $\bar{Rainfall}_{ds}$, measured between 1966 and 2011, and subsequently normalized by the standard deviation in the district's long-run monsoon rainfall.

For *StandRainDev* to be a valid instrument, moonsoon rainfall deviations should be a strong predictor of agricultural yields. This has been established in a number of prior

[12] The full set of district-level time-varying covariates included are: Simpson diversity index for crops; share of non-agricultural land in the district; share of agricultural land which is irrigated; rate of urbanization; average land possession by rural households; share of Muslims, low caste and indigenous populations; share of adults with secondary education; average age of labour force.

[13] The direction of the bias is hard to predict a priori. For instance, if agricultural productivity leads to higher levels of road connectivity which in turn increases industrial activity, this could increase rural employment in the manufacturing sector, biasing the estimated β coefficient upwards. Similarly, if the region declines in terms of political relevance over time, leading to under-investment in public infrastructure, this can diminish both returns from farm and off-farm activities, biasing the estimated β downwards.

[14] Monsoon rainfall is the sum of rainfall incidence in the months of June, July and August.

studies,[15] both in India and across other developing countries. The criticality of natural rainfall is accentuated in the Indian context as irrigation coverage in India remain limited. For instance, in the three decades between 1981 and 2011, only 40 percent of the agricultural land on average is covered by irrigation, which makes the majority of agricultural production rainfed. Monsoon rainfall is the most relevant to agricultural productivity as the bulk of agricultural production - particularly the rice crop - is dependent on monsoon rainfall. This is validated empirically in Table (3.1) where the I test the first stage by regressing logged yields on monsoon rainfall deviations as per equation (3.3). To justify my choice of instrumenting yields with contemporaneous rainfall deviations only, I also include in columns (2) and (4) three lags in monsoon rainfall and one lead. Finally, in columns (5) and (6), I test a flexible functional form where I include 4 categorical variables, corresponding to the top four quintiles of monsoon rainfall. All specifications include district, survey round and state-specific time trends while columns (3), (4) and (6) also include other district-level covariates.

Columns (1)-(4) inform that contemporaneous deviations in monsoon rainfall is a strong predictor of yields and the first stage relationship is stable to the inclusion of covariates. While historical lags one and two years prior have no impact on current yields, there is a negative impact of monsoon rainfall three years prior on current yields. The coefficient however is significantly smaller than the one for contemporaneous monsoon rainfall and the sum of the three historical lags are also not statistically significant. To this effect, the core results of the paper are estimated using a single instrument for yields - contemporaneous monsoon rainfall deviations. As part of my robustness checks, I verify that the exclusion of the historical lags do not affect my results.

In columns (5) and (6) where I estimate the impact of monsoon rainfall on yields with a flexible functional form, the coefficients are benchmarked to the lowest quintile and are monotonically increasing in magnitude, signifying the positive impact of higher incidences of monsoon rainfall on yields. While I am unable to reject the equality of the coefficients for the fourth and the fifth quintiles, the coefficient for the third and fifth quintiles are significantly different from one another. I however do not use monsoon rainfall quintiles in my core specifications to avoid econometric complexities which may arise from the use of multiple instruments but do perform a robustness check to verify that the results are stable to this flexible specification of the instrument.

The second condition for a valid IV strategy is the satisfaction of the exclusion restriction which essentially implies that after conditioning on the observables, rainfall realizations do not affect sectoral employment outcomes through any other channel with the exception of yields. This is an assumption which we impose on the econometric model of the paper. This assumption can be violated if for instance, infrastructural growth due to government policies is correlated over time with positive rainfall deviations, and subsequently facilitate the growth of rural non-farm employment, independent of yields. This would be a particular concern with regard to road connectivity, which has greatly expanded in the past decade across India. To account for this and for other location specific factors, I include district fixed effects and state specific time trends. This implies that any violation of the exclusion restriction would have to occur through district-level time-varying factors which are correlated with both

[15] For instance Jayachandran (2006) and Emerick (2016).

rainfall incidence and sectoral employment outcomes. The paper counters this through the inclusion of the district's rate of urbanization as a covariate. Growth in urbanization can impact RNFE independent of yields through population growth, agglomeration spillovers and upstream and downstream linkages (Christiansen et al., 2013; Michaels et al. 2012). As Henderson et al. (2017) shows that changes in moisture also impacts urbanization, it implies that the omission of urbanization as a covariate can lead to a violation of the exclusion restriction. Moreover, it is plausible that increases in physical infrastructure such as transport or financial connectivity would be correlated with growth in urbanization so observing the rate of urbanization would be a relevant covariate. Nonetheless, it is impossible to rule out that there are other unobservables affecting the estimation and all the results should be treated as an upper bound of the impact of yields on rural non-farm employment.

Finally, as the independent variable of interest in equation 3.2 is yields in the year prior to the NSS survey year, the β coefficient obtained from the IV specification captures the short-term impact of agricultural productivity on RNFE. To identify whether there is a medium or long-term impact of agricultural productivity on RNFE, the paper uses the following lagged reduced form specification motivated by Dell et al. (2012):[16]

$$ln(Y^j_{dst}) = \alpha_d + \theta_{st} + \delta_t + \sum_{k=0}^{k=4} \phi_k StandRainDev_{dst-k} + \gamma X_{dst} + \epsilon_{dst} \tag{3.5}$$

In equation (3.5), standardized monsoon rainfall deviations in the year of the NSS survey year is included, in addition to the standardized monsoon rainfall deviations in the previous 4 years. As monsoon rainfall incidence is an unambiguously exogenous event, the ϕ coefficients from equation (3.5) can be interpreted as the causal impact of a one standard deviation increase in monsoon rainfall on the share of rural workers in sector j.

3.5 Results

The key results of the paper are documented here. I will first present the short-term impact of yields on RNFE, and subsequently discuss the impact of historical rainfall shocks on RNFE. Next, I will present the differential impacts of yields on RNFE by historical levels of urbanization and agricultural advancement in the district, testing whether growth in yields leads to convergence or divergence across districts in terms of RNFE generation. Finally, I will discuss the roles of the potential channels outlined in Section 3.2 in explaining my results by identifying the impact of agricultural productivity on rural non-farm wages.

3.5.1 Short-Term Impact of Yields on RNFE

The causal impact of yields on RNFE in the short-run is presented in Tables 3.2, 3.3 and 3.4. Table 3.2 presents the estimates for all rural workers while Tables 3.3 and 3.4 present

[16] An alternative method would have been to include multiple lags of agricultural productivity, each instrumented by standardized monsoon rainfall deviations corresponding to the endogenous variable of interest. This is not the preferred specification of the paper due to econometric concerns arising from the instrumenting multiple endogenous variables.

estimates separately for male and female workers. In each table, panel A reports estimates in the absence of covariates and panel B contains estimates with covariates. The dependent variable in column (1) of each panel is the logged share of labour force participants aged between 18 and 65 in the district while the dependent variable in the remaining columns is the logged share of workers in each employment category. All the coefficients are estimated using two-stage least squares with the second stage corresponding to equation (3.2) and the first stage corresponding to equation (3.3). Standard errors, clustered at the level of the district, are presented in the parentheses.

In all three tables, the addition of covariates has little impact on the size of the coefficient on logged yields with the exception of skilled non-farm employment.[17] To this effect, all subsequent results include covariates. For both males and females, as well as overall rural workers, yields has a positive and significant impact on employment in the manufacturing sector (column 5). Interpreting the coefficients, a percent increase in district yields in the year prior to the NSS survey year increases the share of rural male workers in the manufacturing sector by 2 percent and female workers by 6 percent. While the elasticities are large, they should be interpreted in the context of the relatively small share of rural workers employed in the manufacturing sector. For the sake of consistency, results from the estimation of equation (3.2) using OLS is presented in the Appendix (Table A3.5). For all the dependent variables, the OLS estimates are biased downwards.

To get a sense of the relative magnitudes, there was a 33 percent increase in the share of female workers employed in the manufacturing sector between 1987 and 2011. In this respect, a one percent increase in annual yields explains a fifth of the total increase in manufacturing employment of rural female workers during this period. In terms of jobs created, at the mean level of district manufacturing employment, a percent increase in yields increases the number of rural male and female workers employed in the manufacturing sector by 1,700 and 2,600 respectively.

Column (4) of Tables (3.2), (3.3) and (3.4) offers evidence that the increase in manufacturing employment is concentrated in unskilled activities. In terms of magnitude, a 1 percent increase in yields increases the share of rural male and female workers employed in unskilled non-farm activities by 1 and 3 percent respectively. In terms of gross employment, this is equivalent to an additional 2,000 jobs for male workers and 2,300 jobs for female workers. For rural female workers, there is also a positive and statistically significant effect of yields on employment in the services sector, but the estimated elasticity is significantly less than that for employment in the manufacturing sector.

As the NSS surveys are not a household panel but a repeated cross-section, it is not possible to precisely establish whether individuals are shifting out of farm work and into non-farm work in response to an increase in yields. However, the regression coefficients do not suggest that rainfall-induced agricultural productivity increases the share of rural workers working in the agriculture sector (column 2) for either males or females even though the point estimate is negative. However, there is suggestive evidence that an increase in

[17] The coefficient is halved upon the inclusion of covariates, particularly for rural males, indicating that the exclusion restriction is possibly being violated in this case. Subsequently, the IV coefficients of yields on rural non-farm employment is treated with caution.

yields results in an increase in labour force participation for rural male workers (Table 3.3, column 1). While the coefficient is not statistically significant (p-value of 0.2), it suggests that a 1 percent increase in yields raises labour for participation for male workers by 0.05 percent.

To ascertain the validity of the exclusion restriction of the IV specification, I test the reduced form impact of monsoon rainfall deviations on rural non-farm employment for male and female workers with and without covariates. The results are presented in Tables (3.5) and (3.6). In each table, panel A excludes all covariates with the exception of the fixed effects and the state-specific time trends while the results in panel B includes district-level time-varying covariates. The results are consistent with those obtained through the IV specification in Tables (3.3) and (3.4). Importantly, with the exception of skilled non-farm employment, the addition of covariates has no impact on either the precision or the magnitude of the reduced form coefficients, implying that the impact of monsoon rainfall shocks on rural non-farm employment is not affected by key observables such as the district's rate of urbanization or share of population with secondary education.

To have a precise understanding of the type of employment generated by positive shocks to agricultural productivity, I identify the effect of yields on disaggregated rural non-farm employment categories in Table 3.7. Panel A contains results for rural male workers while Panel B contains the results for rural female workers. Consistent with the results in Tables (3.3) and (3.4), columns (1)-(4) of Table 3.7 documents that an increase in yields has a positive causal impact on the employment of both rural male and female workers in only unskilled non-farm activities. In addition to the positive effect of yields on employment in the unskilled manufacturing sector, there is also a positive impact on employment in unskilled services for both male and female workers, albeit significant at the 10 percent level. The unskilled manufacturing employment for rural female workers is generated within the food and textile sectors (column 6) and there is no corresponding impact of yields on employment in other manufacturing activities (results not shown). For rural males however, the employment in unskilled manufacturing activities is not through the food and textile sectors. Surprisingly, even though yields have a causal positive effect on employment in unskilled services, this does not occur in the construction sector (column 5) which accounts for over 20 percent of the rural non-farm workers in this period.

Finally, I also explore the impact of short-term yields on the broad employment types of rural workers. This is based on the NSS' classification of workers across 3 broad activities - self-employed, salaried worker and casual workers working for a wage. While all rural salaried workers are employed in the non-farm sector, both farm and non-farm activities are reported by those who are self-employed or employed as casual workers. In this regard, I disaggregate each of these two categories into farm and non-farm work and identify the impact of yields on the share of rural workers employed in each category. The results in Table 3.8 shows that yields have a positive and significant impact on the share of self-employed rural female workers (Panel B, column 2) engaged in non-farm activities. For rural male workers, the results suggest that yields has a positive impact on employment in salaried non-farm activities although the coefficient is significant only at the 15 percent level.

3.5.2 Short-Term Impact of Yields on RNFE by Worker Characteristics

In the results presented till now, a rainfall-induced positive shock to yields has a positive causal impact on RNFE in unskilled activities for both male and female workers, mainly in the manufacturing sector. For rural female workers, this is concentrated within the food and textile sectors. Moreover, there is also evidence that an exogenous positive shock to yields increases non-farm employment for rural male and female workers in the unskilled services sector. These non-farm opportunities generated by a positive rainfall shock to yields take the form of self-employment for rural female workers and salaried employment for rural male workers. Having identified the sectors in which rural workers gain employment in response to a positive shock to yields, I now determine the distributional implications of such a shock by identifying the type of workers who are most impacted by it. To this effect, I disaggregate the sample of rural workers by household landholdings and workers' educational attainment.

For the purposes of this paper, rural landholdings serves as a proxy for household assets. In the lack of other tangible assets, land is a key asset for rural households, particularly as a collateral while obtaining credit from formal sources (Banerjee and Duflo, 2007). The distribution of household land possession is heavily skewed for rural households with the median land holding being 0.5 acres and the mean land holding being 2 acres. As 75 percent of households own under 1 hectare of land, the paper considers households holding under 1 hectare of land as small landowning households, and the remainder as large landholding households. I estimate the IV specification separately for workers from large and small landowning households to identify which workers are most affected by a positive weather shock to agricultural productivity.

The results are presented in Tables 3.9 and 3.10. In both tables, Panel A contains the results for rural male workers while Panel B contains the results for rural female workers. The results show that workers from both small and large landowning households are affected by a positive exogenous weather shock to yields. An increase in yields in the year prior to the survey year raises the share of rural male and female workers in the manufacturing sector (column 5) for workers from both large and small landowning households. Yields also have a positive impact on employment in unskilled non-farm activities (column 4) for rural workers from small landowning households, with the coefficient almost significant at the 10 perent level for male workers and at the 5 percent level for female workers. A corresponding effect is detected for rural female workers from large landowning households but not male workers. Finally, yields have a positive and significant impact on the employment of rural female workers from small landowning families in the services sector (column 6) but no corresponding impact on those from large landowning families. Based on these results, it can be inferred that the impact of a positive shock to yields generates rural non-farm employment opportunities across both large and small landowning households.

With respect to educational qualifications, the existing literature has repeatedly documented the vitality of educational attainment to the process of structural transformation (Erosa et al., 2010; Lee and Malin, 2013; Matsuyama, 1992; Restuccia et al. 2008). However, the results in Section 3.5.1, detect a positive impact of yields only on rural non-farm employment in unskilled activities. In light of this finding, one would expect that workers with relatively low levels of education would be most impacted by a positive shock to agricultural

yields. To test if this is indeed the case, I split the sample of workers by their completion of secondary education.

The results for workers without secondary education are presented in Table 3.12 and those with secondary education are in Table 3.11. Panel A in each table presents the results for male workers and Panel B presents the results for female workers. The results are consistent with the initial hypothesis: yields have a positive impact on rural non-farm employment only for workers without secondary education. Consistent with the overall results, the yields impact both the manufacturing[18] and the service sectors (columns (3) and (4)) and for rural female workers without secondary education, there is a positive and significant impact on employment in the food and textile sectors.

The results in Tables 3.11 and 3.11 are quite significant given the educational profile of the rural workforce. Across the six NSS survey rounds, only 30 percent of rural working-age individuals have completed secondary education with the corresponding statistic being even lower for females. This implies that exogenous shocks to yields generates employment opportunities which can be accessed by the majority of the rural workforce. This implies that growth in agricultural productivity alone can play a significant role in facilitating employment diversification across the rural workforce.

3.5.3 Impact of Historical Rainfall Shocks on Rural Non-Farm Employment

All the results presented in Sections 3.5.1 and 3.5.2 estimate the short-term impact of yields on RNFE. In this section, the paper explores whether there is a longer term impact of yields on RNFE. To this effect, I estimate equation (3.5) to identify the impact of historical lagged rainfall shocks on current employment in non-farm activities.

To facilitate interpretation, the results are presented both as coefficient plots in Figures 3.5 and 3.6 and also in Table 3.13. The results do not provide any evidence of historical rainfall shocks impacting current non-farm employment outcomes. While deviations in monsoon rainfall in the year prior to the NSS survey year (lag 1) has a statistically significant and positive impact on the share of rural males and females employed in the manufacturing sector and unskilled non-farm activities, there is no significant impact of earlier lags. These results underline that the impact of rainfall induced shocks to agricultural productivity affect rural non-farm employment in the short-run but not the medium or long run. This would be consistent with employment generation through the aggregate demand channel with positive rainfall shocks boosting agricultural production and farm incomes, leading to an immediate increase in the demand for consumption goods, causing higher labour demand in the non-farm sector.

Dell et al. (2012) use a lagged temperature shock model to identify the impact of long-run temperature shocks on GDP growth. In their estimation, the individual coefficients on the lagged variables present level effects, while the sum of the lags serves as a test for growth effects. Adopting this framework, I test for the joint significance of the sum of the five lags in monsoon rainfall shocks and am unable to reject the null hypothesis that the sum of the

[18] The coefficient for male workers is almost significant at the 10 percent level.

coefficients equal 0 for either male or female workers, for any of the non-farm employment categories. Consequently, the paper is unable to detect any evidence of a cumulative growth effect of rainfall shocks on RNFE.

3.5.4 Differential Effects of Agricultural Productivity on Rural Non-Farm Employment Across Districts

The empirical results in the previous two sections established that rainfall induced shocks to agricultural productivity has a positive short-term impact on the employment of rural workers in unskilled non-farm activities, mostly in the manufacturing sector. In this section, I test for differential effects of shocks to agricultural productivity across districts based on district characteristics observed in the period prior to my study period. This is to identify whether exogenous shocks to agricultural productivity can assist in the process of convergence across districts in terms of RNFE, or whether the structural transformation process emanating from higher yields is restricted to districts which were at a a relatively advanced stage of development. I test for differential effects across two characteristics - urbanization in 1987 and average agricultural productivity between 1970 and 1985. The differential effects are tested using interaction terms. The formal specification is the following:

$$ln(Y^j_{dst}) = \alpha_d + \theta_s t + \delta_t + \beta_1 ln(Yields_{dst-1}) + \beta_2 DistChar^k_{ds1987} * ln(Yields_{dst-1}) + \gamma X_{dst} + \epsilon_{dst} \tag{3.6}$$

In equation (3.6), *DistChar* is a binary variable categorizing districts based on the characteristic of interest. *DistChar* equals 1 if district d's score exceeds the median district's score in 1987 (1970-1985 for agricultural productivity). To illustrate, the median rate of urbanization in 1987 across the sample was 22 percent. Thus, districts whose rates of urbanization exceeded 22 percent in 1987 receive a value of 1 and the rest receive the value 0. Equation (3.6) is estimated using the IV specification outlined in Section 3.4 with *Yields* being instrumented by standardized deviations in contemporaneous monsoon rainfall. As *Yields* appear twice in equation (3.6), this equation has two endogenous variables and two instruments.

Differential Effects by Rates of Urbanization

Urbanization is a key outcome of the structural transformation process and plays a vital role in the creation of non-farm employment through spillovers from urban demand, migration opportunities and agglomeration economies (Christiansen et al. 2013). To this extent, the paper identifies if RNFE's response to a positive shock to yields is higher in districts with an initially high rate of urbanization. The results are presented in Table 3.14. Panel A presents the results for male workers and Panel B contains the results for female workers. The results assert that the positive effect of short-term yields on RNFE is concentrated in districts which had lower rates of urbanization in 1987. The coefficient on β_1 is positive and statistically significant for employment in the manufacturing sector and unskilled non-farm activities for both male and female workers. While the coefficient on the interaction term is not signficant, the sum of the coefficients is jointly significant for rural female workers in the

manufacturing signifying that a positive shock to yields impacts the employment of rural female workers in the manufacturing sector across both sets of districts.

Differential Effects by Historical Levels of Agricultural Productivity

The second test of differential effects is by districts' historical levels of agricultural productivity. Historical agricultural productivity of the district is the district's average level of agricultural productivity in the fifteen year interval between 1970 and 1985. Historical agricultural productivity measures the level of agricultural advancement in the district. As per Gollin et al (2002), a threshold level of agricultural productivity needs to be attained after which the remaining farm labour exits to non-farm occupations. This specification tests whether the present data supports this theoretical prediction. Additionally, as high agricultural productivity boosts rural incomes, this specification also tests whether the positive impact of agricultural productivity on the structural transformation process is restricted to relatively richer districts.[19]

The results are presented in Table 3.15 with Panel A containing the results for rural male workers and Panel B containing the results for rural female workers. The results suggest that the positive impact of yields on RNFE is restricted to districts which had lower levels of agricultural productivity in the 1970-85 period. The interaction term is significant only for the employment of rural females in the manufacturing sector (column 5) and the coefficients are jointly significant at the 10 percent level. For both rural male and female workers, the results show that a positive shock to farm yields raises unskilled non-farm and manufacturing employment only in historically low agricultural productivity areas.

Collectively, the results presented in this section offer evidence that the non-farm employment generated through rainfall induced positive shocks to farm yields are concentrated in districts which were less urbanized and had lower levels of agricultural productivity. While the non-farm employment generated is limited to unskilled activities which are likely to have lower wages relative to skilled jobs, it also highlights the poverty alleviating aspect of growth in agricultural productivity through the creation of RNFE in areas which have lagged behind in terms of farm yields and urbanization.

3.6 Potential Channels

In Section 3.2, the paper had considered three primary channels outlined in the literature through which agricultural productivity affects non-farm employment. The results in Section 3.5 detected no impact of yields on the share of rural workers employed in the agriculture sector. As the 'labour push' channel relies on surplus workers leaving the farm sector, the lack of a significant effect of yields on farm employment suggests that the 'labour push' channel is not the primary pathway driving the paper's results.

The aggregate demand channel predicts that higher farm yields translates into higher rural incomes, resulting in higher consumer demand for goods and services, which in turn

[19] On average, households in districts classified as high agricultural productivity districts have 7 percent higher monthly per capita expenitures.

increases the demand for labour in the non-farm sector as long as the additional goods and services are locally produced using a labour-intensive production function. If wages are not rigid, higher demand for non-farm labour would result in higher wages for non-farm activities. Thus, if agricultural productivity increases rural non-farm employment through the aggregate demand channel, I would expect to see a corresponding increase in wages for rural non-farm workers.

To test this hypothesis, I run Mincerian wage regressions by regressing individual daily wages of workers measured in 2012 rupees on agricutural productivity, while controlling for workers' age and years of schooling[20] using a quadratic specification. I rely on data at the level of the individual due to the paucity of observations at the level of district, particularly for female workers.[21] The formal specification is the following:

$$ln(wage^{j}_{isdt}) = \mu_s + \alpha_d + \theta_d t + \beta ln(Yields_{dt-1}) + \gamma X_{isdt} + \epsilon_{isdt} \tag{3.7}$$

In equation (3.7), i is the individual and s the subround,[22], equivalent to quarters. The unit of observation is the individual-subround-district-round, with μ denoting subround fixed effects. As I am using individual data, I include district-specific time trends to capture time-varying district-specific shocks to real wages,[23] in addition to district fixed effects α which control for time-invariant district-specific shocks to wages. Finally, individual worker characteristics such as age, years of schooling[24] and household size are included in the vector X. District-specific yields in the year prior to the survey year is considered to be a common shock to all individuals in the district and instrumented using contemporaneous standardized deviations in the district's monsoon rainfall, as described in Section 3.4.

I estimate equation (3.7) seperately for rural male and female workers in Table 3.16, as well as for workers with and without secondary education (Tables 3.18 and 3.17). The results are consistent with the aggregate demand channel being the pathway through which yields affect rural non-farm employment. For both rural male and female workers (Table 3.16, column 2), a percent increase in yields increase real wages in the manufacturing sector by 1 percent. The coefficient is statistically significant at the 15 percent level for rural female workers but highly significant for rural male workers. Consistent with the results obtained in Section 3.5.2, the positive impact of yields on wages applies only to rural workers without

[20] The NSS data only provides levels of schooling completed by responedents so I convert educational level into years of schooling. I assume 0 years of schooling for illiterate workers; 2 years of schooling for literate workers without primary education; 4 years of schooling for workers having primary education; 8 years for schooling for workers with secondary education and 12 years of schooling for those who have completed higher degrees.

[21] For instance, 15 percent of the districts in my sample have no observations for weekly non-farm wages for rural female workers in the manufacturing sector. The lack of observations is due to the non-recording of wages for self-employed work which comprises 40 percent of rural male non-farm workers and 50 percent of rural female non-farm workers.

[22] NSS surveys are undertaken in four waves through the year with each wave randomly sampling a fourth of the households from the full sample. The subrounds correspond to quarters in duration.

[23] The NSS surveys enquires respondents for daily wages earned in the past 7 days. I aggregate total wages and divide by 7 to estimate daily wages for each individual. These daily wages are normalized using a state-level price index to obtain real wages.

[24] Both age and years of schooling are included as quadratic specifications.

secondary education (Table 3.17) and not those with secondary education (Table 3.18).[25]

The third channel discussed in Section 3.2 is the 'labour pull' channel, ocurring due a productivity shock in the non-farm sector with the increased wage differential between non-farm and farm activity attracting farm labour to non-farm jobs (Alvarez-Cuadrado and Poschke, 2011). As rainfall affects the entire spectrum of economic activity, it is plausible that the detected positive impact of yields on RNFE is explained by this channel through a violation of the exclusion restriction.[26] However, the IV results estimating the impact of yields on workers' wages makes this channel unlikeley. First, the productivity shock in the non-farm sector induced through monsoon rainfall has to be restricted only to a subset of workers as there is no impact of a rainfall-induced increase in yields on the real wages of rural workers with secondary education. Second, I find no impact of an increase in yields on real wages for urban workers (results not shown), implying that for the 'labour pull' channel to hold, the rainfall-induced shock to yields would have to affect the productivity of rural non-farm workers but not urban workers. While these above explanations do not fully rule out the 'labour pull' channel, it does make it an unlikely mechanism to explain my results.

3.7 Robustness Checks

In this section, I present a set of robustness checks to determine the empirical validity of my results in Section 3.5. I test for robustness across alternate rainfall instruments and a separate measure of agricultural yields.

3.7.1 Robustness Checks: Alternate Rainfall Instruments

All the IV results presented in this paper uses standardized deviations in monsoon rainfall from the district's historical mean as an instrument for the district's agricultural productivity. To verify that the results are not sensitive to this choice of instrumentation, I undertake robustness checks using three separate choices of instruments. The first is using standardized deviations in annual rainfall instead of monsoon rainfall. The instrument is defined akin to equation (3.4) except that I now consider annual rainfall instead of monsoon rainfall. The results in Table (3.19) confirm that the results are stable to this alternate instrument.

I undertake a second test of the IV strategy by relaxing assumptions regarding the functional form of the instrument and using 4 binary variables as instruments instead, each corresponding to the top four quintiles of the district's monsoon rainfall in the year prior to the NSS survey year. As shown in columns (5) and (6) of Table (3.1), the four categorical variables are valid instruments, all of which have a positive and statistically significant impact on farm yields, alleviating in the process any concerns regarding multiple weak instruments.

[25] There is no impact of yields on wages for rural female workers in the food and textile sectors (column 5) despite the strong positive impact of yields on the share of women employed in these two sectors, documented in Section 3.5.1. This is possibly because 80 percent of female workers employed in these sectors are self-employed for who wages are not recorded.

[26] An immediate pathway is that a higher incidence of rainfall reduces mean temperatures making workers more productive in non-farm activities and raising the marginal product of labour, resulting in higher real wages.

Table (3.20) confirms that the results are stable to this flexible specification of the rainfall instrument. Farm yields in the year prior to the NSS survey year has a positive and significant impact on the employment of both rural male and female workers in unskilled non-farm activities and the manufacturing sector. There is also a positive and statistically significant impact on the share of rural female workers employed in the services sector.

The final test of the IV strategy undertaken is to use historical rainfall lags as instruments, in addition to the contemporaneous standardized deviations in monsoon rainfall. I use three lags and 1 lead of standardized deviations in monsoon rainfall as instruments, in addition to contemporaneous standardized deviations in monsoon rainfall. The first stage corresponding to this choice of instrumentation is presented in columns (2) and (4) of Table 3.1. The lead, as well as the first and second lags are not statistically significant, which can lead to concerns about multiple weak instruments. However, the coefficients are jointly significant and the F-statistic for the first stage also exceeds 30. The results from this choice of instrumentation are presented in Table 3.21 and confirm that the core results obtained in Section 3.5 are stable to the use of lagged rainfall shocks as instruments.

3.7.2 Robustness Check: Alternate Definition of Agricultural Productivity

This paper uses a composite yields index across 7 crops to measure district-level agricultural productivity. As there is a lack of reliable data on historical district-level crop prices, I construct a volume-weighted index of productivity, relying on the share of land allocated to each crop as weights. To verify that the results are not an artefact of this specification of agricultural productivity, I re-estimate my IV results using rice yields in the year prior to the NSS survey year as a proxy for district-level agricultural productivity. I choose rice as it is the single largest crop in terms of acreage with one-sixth of the average district's cropped area allocated to rice production. Additionally, rice is also a water-intensive crop with the major cropping season corresponding exactly with the onset of the monsoon.[27] In this regard, the impact of monsoon rainfall would be most relevant for the rice crop.

Rice yields are defined as the total output of rice in the district (in tonnes), divided by the total area in the district (in hectares) allocated to rice production. The results are presented in Table 3.22 and are very similar to those obtained with the composite measure of yields. The coefficient for female workers employed in unskilled non-farm activities is significant at the 20 percent level but the coefficient remains positive. The specification is also able to detect a positive and significant impact of rice yields on employment in the manufacturing sector for rural workers as well as for rural male workers in unskilled non-farm activities.

3.8 Conclusion

This paper shows the positive causal impact of rainfall-induced agricultural productivity on rural non-farm employment using panel data from 273 districts in India over a twenty-five

[27] The Kharif rice crop is planted between June and July and harvested between November and December.

year period. The increase in rural non-farm employment is located primarily in the manufacturing sector, in unskilled activities, across both male and female workers. The paper also identifies the nature of non-farm employment generated by weather shocks to agricultural productivity - while male workers gain employment in salaried non-farm occupations, female workers are primarily self-employed. Using historical lagged rainfall shocks, the paper documents that the impact of agricultural productivity on rural non-farm employment is restricted to short-term effects. Finally, by identifying the causal impact of agricultural growth on rural wages, the paper offers suggestive evidence that the positive impact of agricultural growth on rural non-farm employment is driven by the aggregate demand channel.

With regard to policy implications, the paper's results highlight the poverty alleviating aspect of agricultural productivity. The impact on non-farm employment is restricted to rural workers with low levels of education, highlighting its accessibility to a broad range of rural workers, the majority of whom have not completed secondary education. Moreover, my results show that the positive effect of agricultural productivity is salient in regions which were at lower levels of development, measured through historical rates of urbanization and agricultural productivity, signifying that sustained growth in agricultural productivity can offer an effective pathway for lagging regions to catch up with those which have proceeded to a higher threshold of development.

3.9 Figures

Figure 3.1: Growth in Agricultural Productivity Over Time

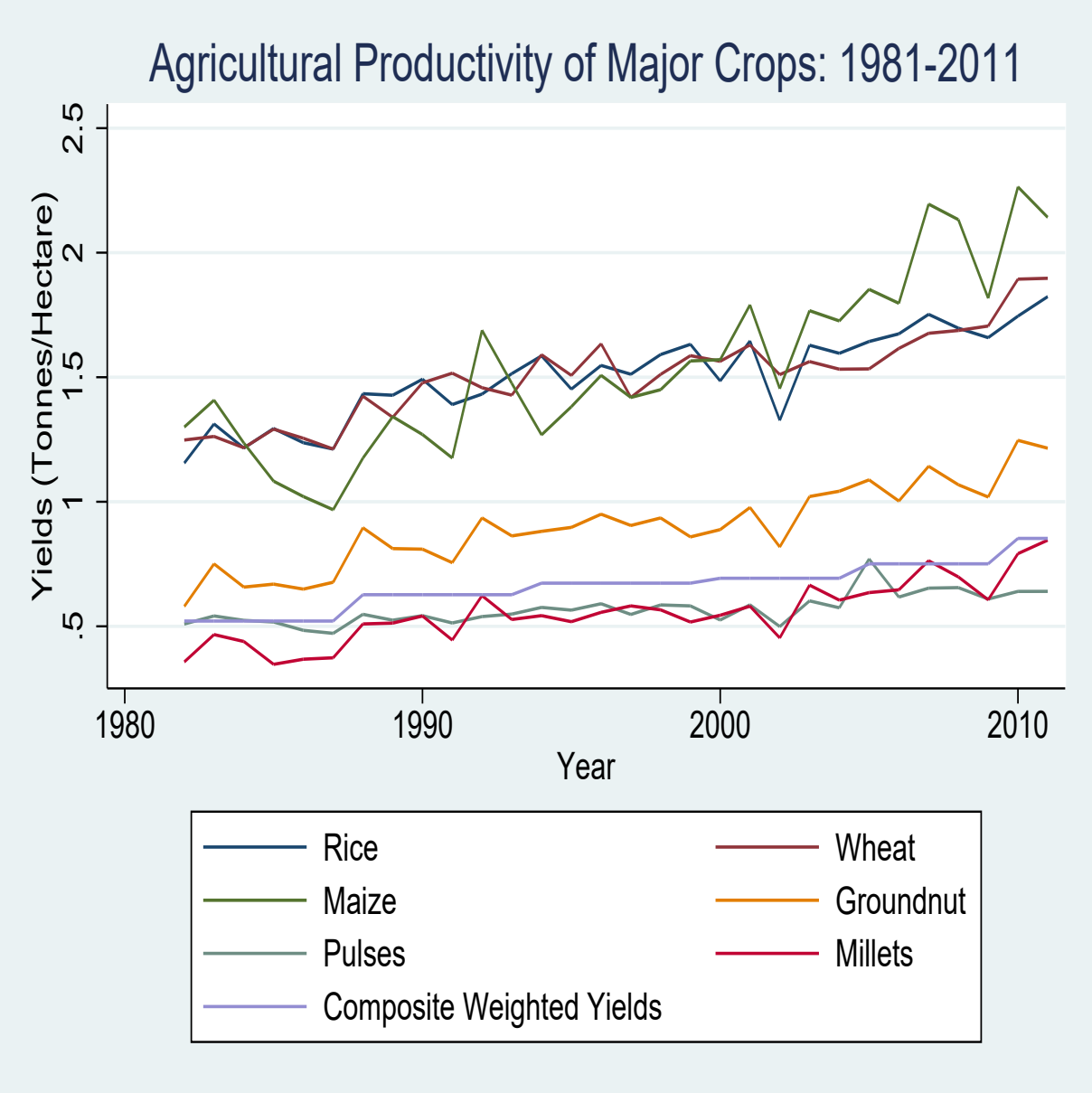


Figure 3.2: District Labour Force Participation Over Time

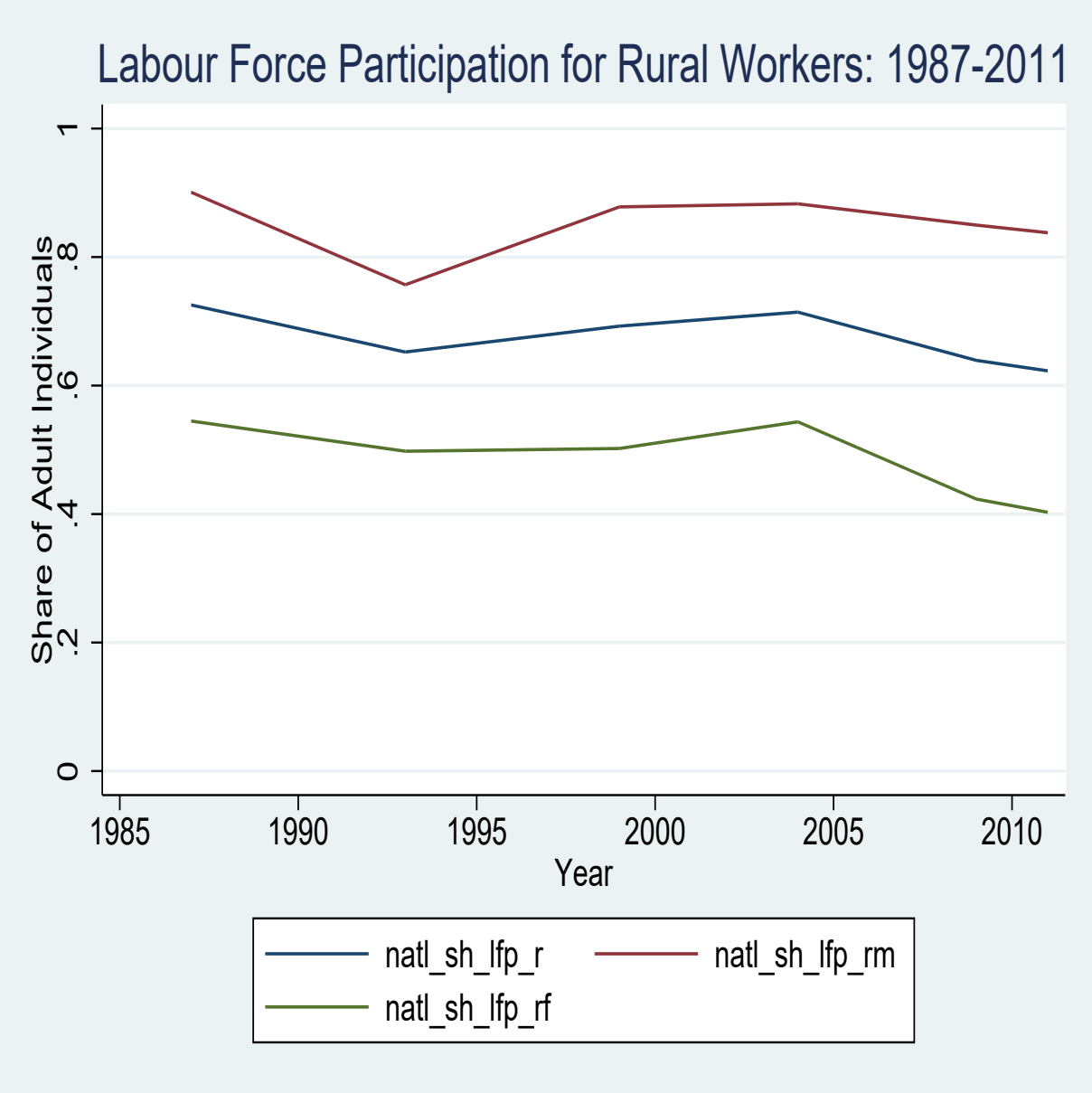


Figure 3.3: Rural Non-Farm Employment Over Time, Aggregate Categories

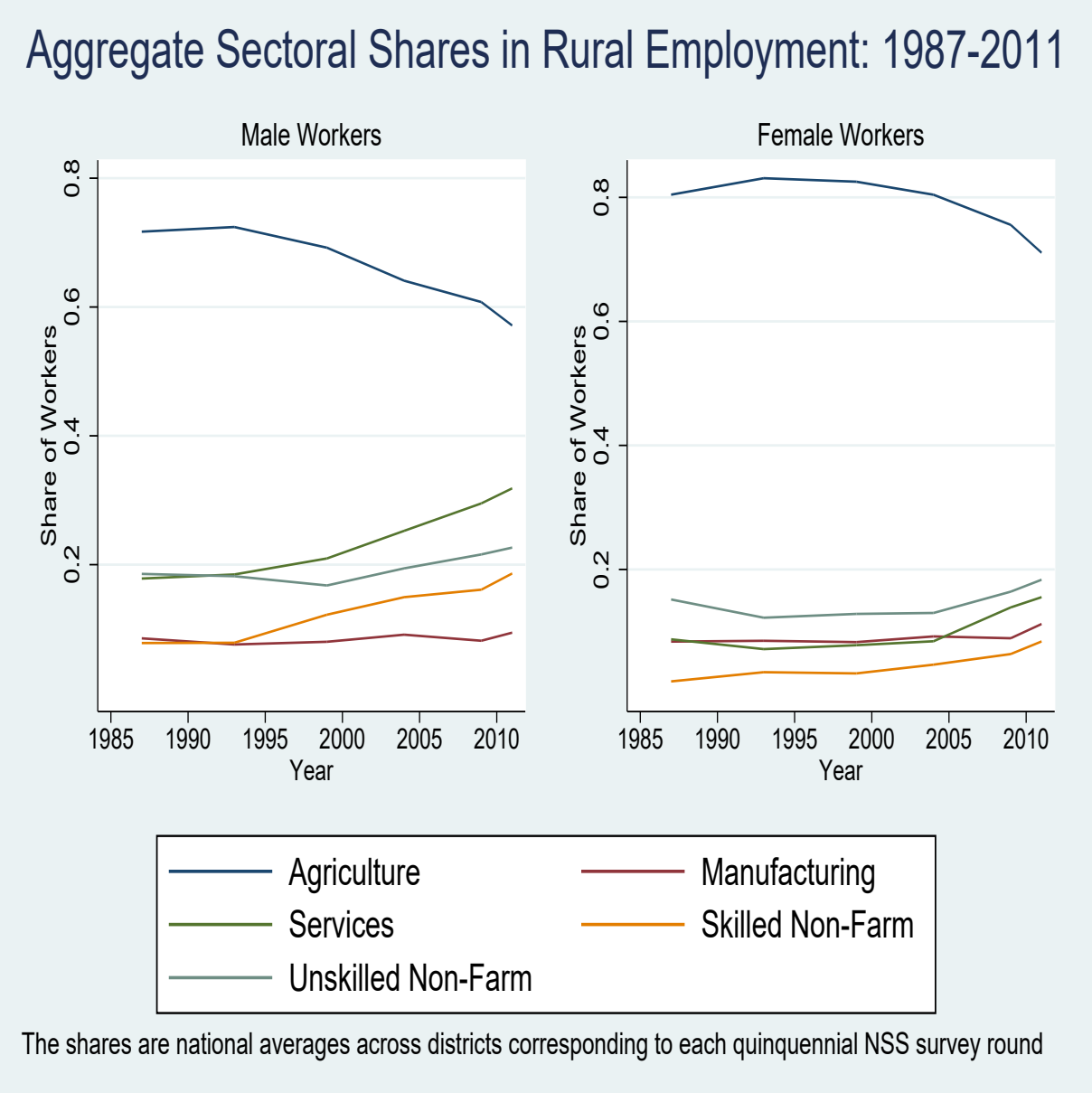


Figure 3.4: Rural Non-Farm Employment Over Time, Detailed Categories

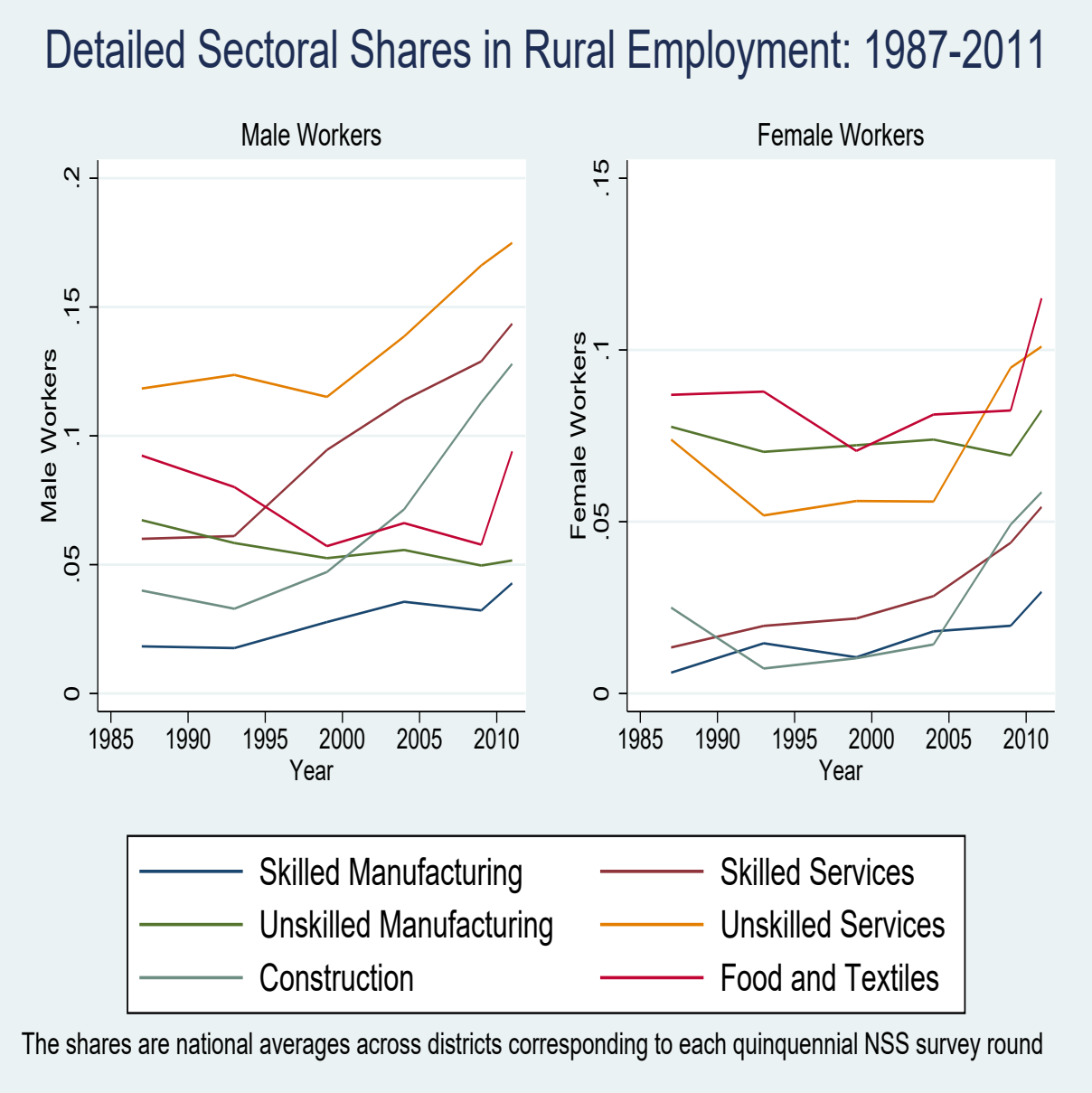
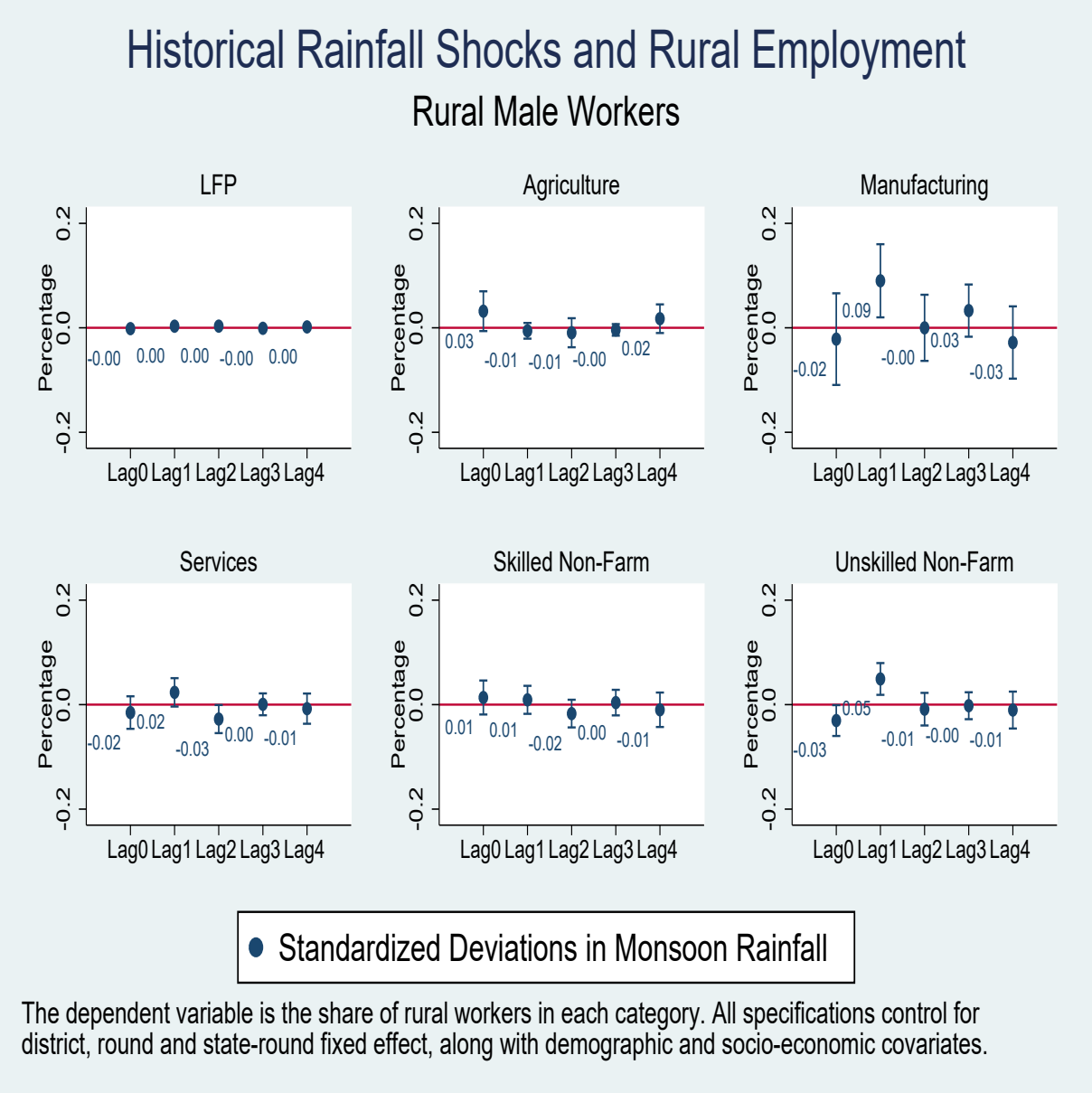
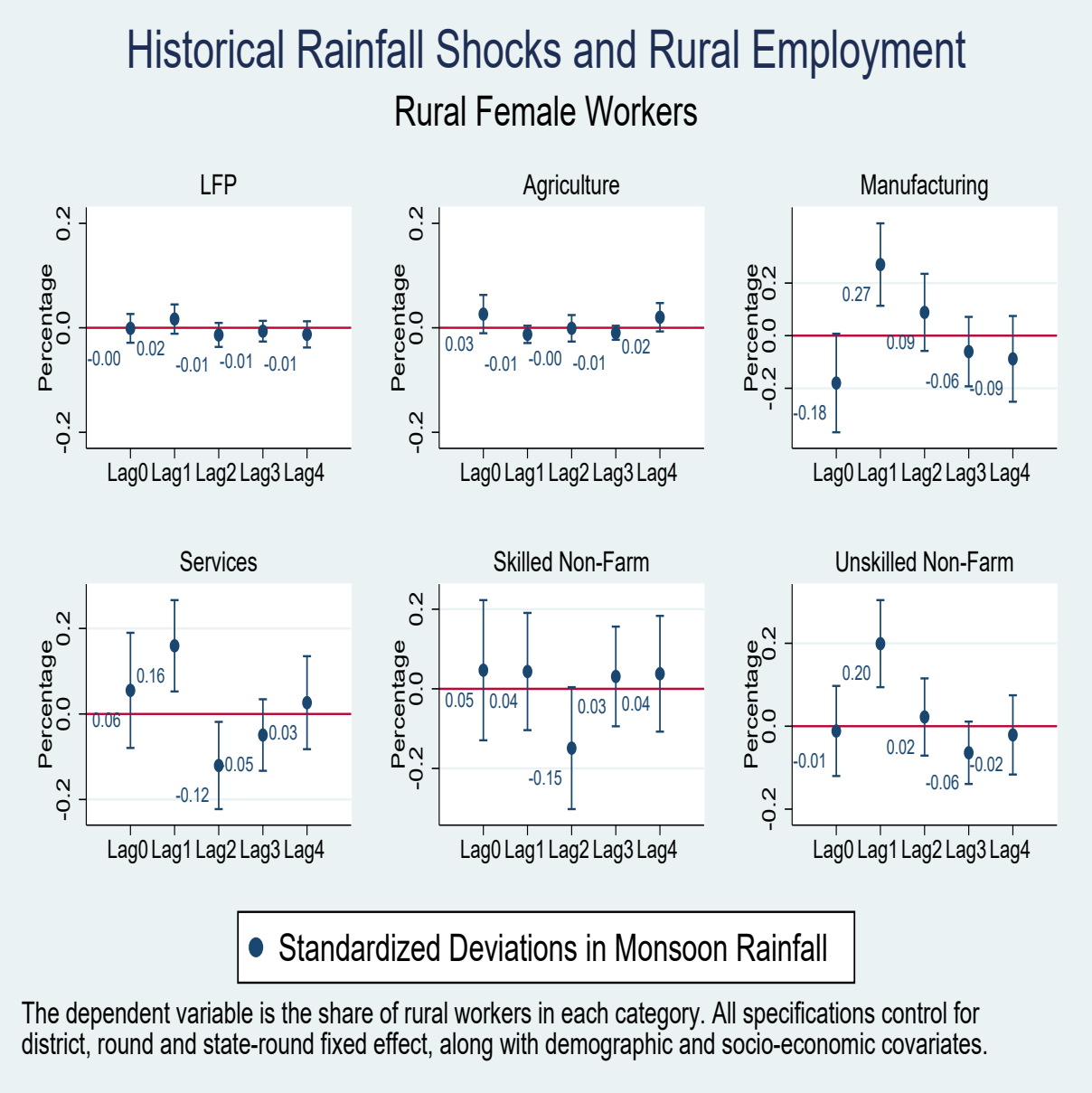


Figure 3.5: Historical Shocks and Rural Non-Farm Employment: Male Workers



The dependent variable is the share of rural workers in each category. All specifications control for district, round and state-round fixed effect, along with demographic and socio-economic covariates.

Figure 3.6: Historical Shocks and Rural Non-Farm Employment: Female Workers



The dependent variable is the share of rural workers in each category. All specifications control for district, round and state-round fixed effect, along with demographic and socio-economic covariates.

3.10 Tables

3.10.1 Main Results

Table 3.1: Impact of Monsoon Rainfall on Agricultural Productivity - First Stage Estimates

	(1) Yields	(2) Yields	(3) Yields	(4) Yields	(5) Yields	(6) Yields
Deviations in Rainfall - Lead 1		.012		.014		
		(.011)		(.010)		
Deviations in Rainfall - Contemporaneous	.042***	.037***	.044***	.039***		
	(.008)	(.009)	(.008)	(.009)		
Deviations in Rainfall - Lag 1		-.004		-.004		
		(.010)		(.010)		
Deviations in Rainfall - Lag 2		.007		.008		
		(.007)		(.007)		
Deviations in Rainfall - Lag 3		-.020**		-.017*		
		(.010)		(.010)		
Rainfall Quintile 2, Contemporaneous					.047**	.049***
					(.019)	(.018)
Rainfall Quintile 3, Contemporaneous					.073***	.081***
					(.023)	(.023)
Rainfall Quintile 4, Contemporaneous					.099***	.102***
					(.024)	(.023)
Rainfall Quintile 5, Contemporaneous					.129***	.137***
					(.026)	(.026)
Observations	1614	1344	1613	1343	1614	1613
R^2	.85	.85	.86	.86	.85	.86
Covariates	No	No	Yes	Yes	No	Yes
Dependent Variable Mean (Tonnes/Hectare)	.721	.721	.721	.721	.721	.721

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is logged composite district-level agricultural yields measured in the year prior to the NSS survey year. The independent variables of interest are contemporaneous and lagged standardized deviations in district's monsoon rainfall from the district's historical mean. Specifications in columns (3) and (4) control for district-level socio-economic and demographic covariates. District, state time trends and survey round fixed effects are also included.

Table 3.2: Short-Term Agricultural Productivity and Non-Farm Employment: Rural Workers, Broad Categories

	Panel A: Without Covariates					
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.052	-.113	.601*	.729**	2.336***	.412
	(.066)	(.185)	(.324)	(.362)	(.780)	(.306)
Observations	1614	1614	1614	1614	1614	1614
R^2	.71	.43	.67	.55	.34	.63
Covariates	No	No	No	No	No	No
Dependent Variable Mean (Share)	.672	.698	.106	.179	.085	.200
	Panel B: With Covariates					
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.079	-.089	.283	.751**	2.011***	.323
	(.058)	(.197)	(.257)	(.358)	(.743)	(.291)
Observations	1613	1613	1613	1613	1613	1613
R^2	.75	.51	.75	.58	.39	.67
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.672	.698	.106	.179	.085	.200

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects. Specifications in Panel B also control for district-level socio-economic and demographic covariates.

Table 3.3: Short-Term Agricultural Productivity and Non-Farm Employment - Rural Male Workers

			Panel A: Without Covariates			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.049	-.225	.592*	.908**	2.292***	.555*
	(.038)	(.263)	(.327)	(.386)	(.884)	(.329)
Observations	1614	1614	1614	1614	1614	1614
R^2	.68	.49	.66	.52	.33	.62
Covariates	No	No	No	No	No	No
Dependent Variable Mean (Share)	.853	.648	.137	.199	.086	.250
			Panel B: With Covariates			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.046	-.105	.301	.844**	2.202***	.412
	(.035)	(.202)	(.259)	(.348)	(.823)	(.296)
Observations	1613	1613	1613	1613	1613	1613
R^2	.70	.57	.74	.56	.35	.66
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.853	.648	.137	.199	.086	.250

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age males in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects. Specifications in Panel B also control for district-level socio-economic and demographic covariates.

Table 3.4: Short-Term Agricultural Productivity and Non-Farm Employment - Rural Female Workers

			Panel A: Without Covariates			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.195	-.171	1.007	3.183**	6.214***	2.395*
	(.295)	(.271)	(1.427)	(1.383)	(1.945)	(1.240)
Observations	1613	1613	1613	1613	1613	1613
R^2	.69	.43	.52	.22	.27	.31
Covariates	No	No	No	No	No	No
Dependent Variable Mean (Share)	.480	.782	.050	.149	.092	.107
			Panel B: With Covariates			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.189	-.132	1.342	2.964**	5.955***	2.359**
	(.267)	(.215)	(1.371)	(1.256)	(1.787)	(1.177)
Observations	1612	1612	1612	1612	1612	1612
R^2	.70	.49	.55	.26	.31	.32
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.480	.782	.050	.149	.092	.107

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age females in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity growth in year prior to survey year. Agricultural productivity growth in year prior to survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects. Specifications in Panel B also control for district-level socio-economic and demographic covariates.

Table 3.5: Monsoon Rainfall Shocks and Non-Farm Employment - Reduced Form Estimates, Rural Male Workers

			Panel A: Without Covariates			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Deviations in Rainfall - Lag 1	.002	-.006	.026*	.035**	.093***	.021
	(.002)	(.009)	(.014)	(.014)	(.032)	(.014)
Observations	1614	1614	1614	1614	1614	1614
R^2	.69	.53	.68	.60	.50	.65
Covariates	No	No	No	No	No	No
Dependent Variable Mean (Share)	.853	.648	.137	.199	.086	.250
			Panel B: With Covariates			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Deviations in Rainfall - Lag 1	.002	-.005	.013	.037**	.098***	.018
	(.002)	(.010)	(.012)	(.015)	(.035)	(.014)
Observations	1613	1613	1613	1613	1613	1613
R^2	.71	.59	.74	.63	.51	.68
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.853	.648	.137	.199	.086	.250

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age males in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is lagged standardized deviations of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects. Specifications in Panel B also control for district-level socio-economic and demographic covariates.

Table 3.6: Monsoon Rainfall Shocks and Non-Farm Employment - Reduced Form Estimates, Rural Female Workers

			Panel A: Without Covariates			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Deviations in Rainfall - Lag 1	.010	-.003	.037	.120**	.222***	.103**
	(.013)	(.010)	(.067)	(.049)	(.075)	(.048)
Observations	1613	1613	1613	1613	1613	1613
R^2	.69	.45	.52	.41	.48	.40
Covariates	No	No	No	No	No	No
Dependent Variable Mean (Share)	.480	.782	.050	.149	.092	.107
			Panel B: With Covariates			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Deviations in Rainfall - Lag 1	.008	-.006	.059	.131**	.264***	.104**
	(.013)	(.010)	(.067)	(.051)	(.076)	(.051)
Observations	1612	1612	1612	1612	1612	1612
R^2	.70	.51	.56	.43	.49	.42
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.480	.782	.050	.149	.092	.107

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age females in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is lagged standardized deviations of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects. Specifications in Panel B also control for district-level socio-economic and demographic covariates.

Table 3.7: Short-Term Agricultural Productivity and Non-Farm Employment: Rural Workers, Detailed Employment Categories

			Panel A: Male Workers			
	(1) Manuf. Skilled	(2) Services Skilled	(3) Manuf. Unskilled	(4) Services Unskilled	(5) Constr.	(6) Food and Textiles
Ag Productivity, 1 Lag	1.604	.110	2.429**	.831*	-.078	.650
	(1.550)	(.310)	(1.066)	(.428)	(1.304)	(.726)
Observations	1608	1608	1608	1608	1608	1608
R^2	.51	.69	.35	.50	.50	.44
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.031	.106	.055	.144	.079	.074
			Panel B: Female Workers			
	(1) Manuf. Skilled	(2) Services Skilled	(3) Manuf. Unskilled	(4) Services Unskilled	(5) Constr.	(6) Food and Textiles
Ag Productivity, 1 Lag	3.298*	1.225	5.313***	3.077*	1.498	4.298**
	(1.808)	(1.614)	(1.830)	(1.581)	(2.364)	(1.698)
Observations	1607	1607	1607	1607	1607	1607
R^2	.51	.53	.35	.26	.50	.41
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.018	.033	.075	.075	.030	.088

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects, alongwith controls for district-level socio-economic and demographic covariates. Panel A present estimates for rural male workers; panel B presents estimates for rural female workers.

Table 3.8: Short-Term Agricultural Productivity and Non-Farm Employment: Rural Workers, Employment Types

	Panel A: Male Workers				
	(1) Self-Emp. Ag	(2) Self-Emp. Non-Ag	(3) Casual Work Ag	(4) Casual Work Non-Ag	(5) Salary Work Non-Ag
Yields, Lag 1	.065	.385	-.374	-.238	.619
	(.249)	(.409)	(.661)	(.962)	(.422)
Observations	1536	1536	1536	1536	1536
R^2	.65	.50	.64	.49	.55
Covariates	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.391	.139	.257	.109	.087
	Panel B: Female Workers				
	(1) Self-Emp. Ag	(2) Self-Emp. Non-Ag	(3) Casual Work Ag	(4) Casual Work Non-Ag	(5) Salary Work Non-Ag
Yields, Lag 1	.245	4.085***	.305	2.016	1.127
	(.476)	(1.547)	(1.092)	(2.145)	(1.669)
Observations	1535	1535	1535	1535	1535
R^2	.50	.32	.64	.43	.47
Covariates	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.469	.103	.313	.057	.040

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force in each employment type. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects in addition to district-level socio-economic and demographic covariates.

Table 3.9: Short-Term Agricultural Productivity and Non-Farm Employment: Rural Workers, Small Landowning Households

			Panel A: Male Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	-.000	-.117	.064	.641	2.189**	.310
	(.042)	(.237)	(.678)	(.413)	(.946)	(.509)
Observations	1611	1611	1611	1611	1611	1611
R^2	.65	.66	.56	.46	.31	.47
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.855	.568	.156	.258	.108	.306
			Panel B: Female Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.031	-.364	2.094	3.071**	5.381***	3.478**
	(.281)	(.341)	(1.566)	(1.377)	(1.991)	(1.443)
Observations	1608	1608	1608	1608	1608	1608
R^2	.65	.36	.55	.22	.34	.25
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.480	.745	.054	.182	.108	.128

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects, alongwith controls for district-level socio-economic and demographic covariates. Small landowning households are those with less than 1 hectare of land.

Table 3.10: Short-Term Agricultural Productivity and Non-Farm Employment: Rural Workers, Large Landowning Households

			Panel A: Male Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.046	-.114	1.655*	1.497	5.505***	.622
	(.062)	(.215)	(.889)	(1.637)	(2.130)	(.791)
Observations	1611	1611	1611	1611	1611	1611
R^2	.51	.37	.39	.41	.27	.34
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.853	.821	.102	.064	.038	.128
			Panel B: Female Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	-.091	-.128	-.126	3.537*	4.768**	.528
	(.453)	(.469)	(1.733)	(2.130)	(2.367)	(1.967)
Observations	1589	1589	1589	1589	1589	1589
R^2	.67	.39	.47	.34	.33	.39
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.468	.856	.059	.063	.041	.081

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects, alongwith controls for district-level socio-economic and demographic covariates. Large landowning households are those with an excess of 1 hectare of land.

Table 3.11: Short-Term Agricultural Productivity and Non-Farm Employment: Rural Workers Without Secondary Education

	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Manuf.	Services	Constr.	Food and Textiles
Ag Productivity, 1 Lag	.009	-.234	1.628	1.056**	-.915	.827
	(.033)	(.218)	(1.007)	(.453)	(1.440)	(1.018)
Observations	1613	1613	1613	1613	1613	1613
R^2	.80	.48	.39	.53	.51	.42
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.912	.701	.081	.211	.092	.070
			Panel B: Female Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Manuf.	Services	Constr.	Food and Textiles
Ag Productivity, 1 Lag	.279	.056	5.487***	3.134**	1.053	4.625**
	(.269)	(.299)	(1.879)	(1.593)	(2.329)	(1.864)
Observations	1611	1611	1611	1611	1611	1611
R^2	.68	.38	.35	.28	.51	.41
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.524	.812	.092	.090	.034	.096

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. The sample is restricted to workers who have not completed secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects, alongwith controls for district-level socio-economic and demographic covariates. Panel A present estimates for rural male workers; panel B presents estimates for rural female workers.

Table 3.12: Short-Term Agricultural Productivity and Non-Farm Employment: Rural Workers with Secondary Education

	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Manuf.	Services	Constr.	Food and Textiles
Ag Productivity, 1 Lag	.143*	.010	1.904	-.027	-.160	-.196
	(.075)	(.239)	(1.483)	(.289)	(1.617)	(.926)
Observations	1613	1613	1613	1613	1613	1613
R^2	.30	.59	.42	.46	.50	.33
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.774	.548	.092	.328	.060	.122
			Panel B: Female Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Manuf.	Services	Constr.	Food and Textiles
Ag Productivity, 1 Lag	-.100	-1.657	2.904	1.654	2.635*	1.400
	(.449)	(1.287)	(1.852)	(1.513)	(1.594)	(1.640)
Observations	1561	1561	1561	1561	1561	1561
R^2	.56	.34	.40	.39	.32	.39
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.313	.607	.101	.237	.014	.126

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. The sample is restricted to workers who have completed secondary education or above. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects, alongwith controls for district-level socio-economic and demographic covariates. Panel A present estimates for rural male workers; panel B presents estimates for rural female workers.

Table 3.13: Historical Rainfall Shocks and Non-Farm Employment - Rural Workers

			Panel A: Male Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Deviations in Rainfall - Lead 1	-.002	.032	.014	-.030*	-.022	-.015
	(.002)	(.023)	(.020)	(.018)	(.053)	(.019)
Deviations in Rainfall - Contemporaneous	.003	-.006	.009	.049***	.090**	.023
	(.002)	(.009)	(.016)	(.018)	(.042)	(.016)
Deviations in Rainfall - Lag 1	.003	-.009	-.017	-.009	-.000	-.028*
	(.002)	(.017)	(.016)	(.019)	(.038)	(.016)
Deviations in Rainfall - Lag 2	-.001	-.004	.004	-.002	.033	.001
	(.002)	(.007)	(.015)	(.016)	(.030)	(.013)
Deviations in Rainfall - Lag 3	.002	.017	-.010	-.010	-.028	-.008
	(.002)	(.017)	(.020)	(.021)	(.042)	(.018)
Observations	1343	1343	1343	1343	1343	1343
R^2	.75	.55	.73	.65	.55	.68
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.853	.648	.137	.199	.086	.250
			Panel B: Female Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Deviations in Rainfall - Lead 1	-.001	.026	.047	-.012	-.180	.055
	(.017)	(.022)	(.107)	(.066)	(.113)	(.082)
Deviations in Rainfall - Contemporaneous	.017	-.013	.043	.199***	.270***	.159**
	(.017)	(.010)	(.089)	(.064)	(.095)	(.065)
Deviations in Rainfall - Lag 1	-.014	-.001	-.149	.022	.088	-.121*
	(.014)	(.015)	(.093)	(.056)	(.089)	(.062)
Deviations in Rainfall - Lag 2	-.006	-.009	.031	-.064	-.061	-.049
	(.012)	(.008)	(.076)	(.046)	(.080)	(.051)
Deviations in Rainfall - Lag 3	-.013	.020	.038	-.021	-.088	.026
	(.015)	(.016)	(.088)	(.058)	(.099)	(.066)
Observations	1342	1342	1342	1342	1342	1342
R^2	.74	.58	.57	.46	.51	.43
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.480	.782	.050	.149	.092	.107

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. The independent variables of interest are lagged standardized deviations in the district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects, in addition to district-level socio-economic and demographic covariates.

Table 3.14: Differential Effects of Short-Term Agricultural Productivity on Non-Farm Employment by Historical Levels of Urbanization: Rural Workers

			Panel A: Male Workers			
	(1)	(2)	(3) Non-Farm	(4) Non-Farm	(5)	(6)
	LFP	Ag.	Skilled	Unskilled	Manuf.	Services
Yields, Lag 1	.078**	-.148	.193	.986**	1.984**	.487
	(.038)	(.233)	(.263)	(.394)	(.797)	(.312)
Yields, Lag 1*High Urban, 1987	-.092*	.126	.317	-.412	.634	-.217
	(.056)	(.191)	(.418)	(.499)	(1.285)	(.439)
Observations	1613	1613	1613	1613	1613	1613
R^2	.68	.57	.73	.55	.34	.66
Joint Significance - 10pc	No	No	No	No	No	No
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.853	.648	.137	.199	.086	.250
			Panel B: Female Workers			
	(1)	(2)	(3) Non-Farm	(4) Non-Farm	(5)	(6)
	LFP	Ag.	Skilled	Unskilled	Manuf.	Services
Yields, Lag 1	.283	-.165	1.219	3.052**	4.369**	2.854**
	(.254)	(.240)	(1.285)	(1.401)	(1.870)	(1.337)
Yields, Lag 1*High Urban, 1987	-.272	.094	.353	-.253	4.551	-1.422
	(.416)	(.240)	(2.357)	(1.769)	(2.819)	(1.702)
Observations	1612	1612	1612	1612	1612	1612
R^2	.70	.49	.55	.26	.27	.31
Joint Significance - 10pc	Yes	No	No	No	Yes	No
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.480	.782	.050	.149	.092	.107

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity growth in year prior to survey year. Agricultural productivity growth in year prior to survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state-survey round and survey round fixed effects, alongwith controls for district-level socio-economic and demographic covariates. Panel A present estimates for rural male workers; panel B presents estimates for rural female workers.

Table 3.15: Differential Effects of Short-Term Agricultural Productivity on Non-Farm Employment by Historical Levels of Agricultural Productivity: Rural Workers

	Panel A: Male Workers					
	(1)	(2)	(3) Non-Farm	(4) Non-Farm	(5)	(6)
	LFP	Ag.	Skilled	Unskilled	Manuf.	Services
Yields, Lag 1	.047	-.214	.319	.927***	1.914***	.568*
	(.031)	(.234)	(.238)	(.338)	(.733)	(.293)
Yields, Lag 1*High Yields, 1970-85	-.004	.619	-.097	-.470	1.635	-.883
	(.080)	(.416)	(.663)	(.676)	(2.047)	(.659)
Observations	1613	1613	1613	1613	1613	1613
R^2	.70	.53	.74	.55	.33	.63
Joint Significance - 10pc	No	No	No	No	No	No
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.850	.676	.120	.189	.075	.234
	Panel B: Female Workers					
	(1)	(2)	(3) Non-Farm	(4) Non-Farm	(5)	(6)
	LFP	Ag.	Skilled	Unskilled	Manuf.	Services
Yields, Lag 1	.158	-.213	1.975	2.091**	4.005***	2.030*
	(.169)	(.228)	(1.263)	(1.017)	(1.466)	(1.114)
Yields, Lag 1*High Yields, 1970-85	.156	.415	-3.260	4.494	10.036**	1.694
	(.813)	(.459)	(3.406)	(3.110)	(5.009)	(2.399)
Observations	1612	1612	1612	1612	1612	1612
R^2	.70	.48	.53	.21	.23	.32
Joint Significance - 10pc	No	No	No	No	Yes	No
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.487	.805	.041	.139	.079	.101

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects, alongwith controls for district-level socio-economic and demographic covariates. Panel A present estimates for rural male workers; panel B presents estimates for rural female workers.

Table 3.16: Short-Term Agricultural Productivity and Rural Wages: Rural Workers

			Panel A: Male Workers		
	(1)	(2)	(3)	(4)	(5)
	Ag.	Manuf.	Services	Constr.	Food and Textiles
Yields, Lag 1	.2269	.7849**	-.0368	.1474	.8697
	(.1583)	(.3400)	(.1653)	(.1509)	(.5637)
Observations	70131	16524	67112	24903	8633
R^2	.38	.44	.43	.51	.54
Covariates	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	102	174	252	146	158
			Panel B: Female Workers		
	(1)	(2)	(3)	(4)	(5)
	Ag.	Manuf.	Services	Constr.	Food and Textiles
Yields, Lag 1	.0426	1.5984	-.3333	-.2745	-1.0833
	(.1614)	(1.1029)	(.2835)	(.2269)	(1.3990)
Observations	42730	4792	15250	4678	3101
R^2	.39	.55	.48	.60	.63
Covariates	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	71	85	246	96	79

Standard errors in parentheses (clustered at the district level). The unit of observation is individual-district-subround. The dependent variable in each specification is the logged daily wage of individuals, measured in 2012 rupees. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, sub-round, survey round fixed effects in addition to district-specific time trends. Individual level characteristics such as age, years of schooling and household size are also included as covariates.

Table 3.17: Short-Term Agricultural Productivity and Rural Wages: Rural Workers without Secondary Education

			Panel A: Male Workers		
	(1) Ag.	(2) Manuf.	(3) Services	(4) Constr.	(5) Food and Textiles
Yields, Lag 1	.1721	1.0773**	.0528	.0905	.7193
	(.1887)	(.4613)	(.2043)	(.1629)	(1.1339)
Observations	57308	9028	28790	16853	4683
R^2	.40	.50	.33	.50	.61
Covariates	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	97	150	159	137	141
			Panel B: Female Workers		
	(1) Ag.	(2) Manuf.	(3) Services	(4) Constr.	(5) Food and Textiles
Yields, Lag 1	.0380	2.0725	.0690	-.1868	-2.5402
	(.1633)	(1.5179)	(.2929)	(.2275)	(2.6303)
Observations	39046	3632	7535	4104	2252
R^2	.40	.65	.41	.61	.71
Covariates	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	70	83	220	95	77

Standard errors in parentheses (clustered at the district level). The unit of observation is individual-district-subround. The dependent variable in each specification is the logged daily wage of individuals, measured in 2012 rupees. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, sub-round, survey round fixed effects in addition to district-specific time trends. Individual level characteristics such as age, years of schooling and household size are also included as covariates.

Table 3.18: Short-Term Agricultural Productivity and Rural Wages: Rural Workers with Secondary Education

			Panel A: Male Workers		
	(1)	(2)	(3)	(4)	(5)
	Ag.	Manuf.	Services	Constr.	Food and Textiles
Yields, Lag 1	.2209	.2284	.0215	.2989	.5354
	(.1619)	(.4193)	(.1851)	(.2692)	(.4099)
Observations	12823	7496	38322	8050	3950
R^2	.44	.54	.45	.64	.62
Covariates	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	125	205	327	170	178
			Panel B: Female Workers		
	(1)	(2)	(3)	(4)	(5)
	Ag.	Manuf.	Services	Constr.	Food and Textiles
Yields, Lag 1	-.2873	.7131	.6445	1.4427	.5369
	(.3665)	(.7536)	(.7438)	(4.8549)	(.3957)
Observations	3684	1160	7715	574	849
R^2	.69	.48	.62	.95	.58
Covariates	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean	88	90	279	106	87

Standard errors in parentheses (clustered at the district level). The unit of observation is individual-district-subround. The dependent variable in each specification is the logged daily wage of individuals, measured in 2012 rupees. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, sub-round, survey round fixed effects in addition to district-specific time trends. Individual level characteristics such as age, years of schooling and household size are also included as covariates.

3.10.2 Robustness Checks

Table 3.19: Short-Term Agricultural Productivity and Non-Farm Employment for Rural Workers: Robustness Check with Annual Rainfall Instrument

			Panel A: Male Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.026	-.117	.534**	.499**	1.519***	.406*
	(.028)	(.153)	(.219)	(.253)	(.558)	(.234)
Observations	1613	1613	1613	1613	1613	1613
R^2	.71	.57	.72	.60	.43	.66
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.853	.648	.137	.199	.086	.250
			Panel B: Female Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.067	.128	1.843	.766	3.374**	.986
	(.229)	(.203)	(1.217)	(.766)	(1.316)	(.841)
Observations	1612	1612	1612	1612	1612	1612
R^2	.70	.51	.54	.41	.43	.39
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.480	.782	.050	.149	.092	.107

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's annual rainfall from the district's historical mean. All specifications include district, state time trends and survey round fixed effects, alongwith controls for district-level socio-economic and demographic covariates.

Table 3.20: Short-Term Agricultural Productivity and Non-Farm Employment for Rural Workers: Robustness Check with Monsoon Rainfall Quintiles Instrument

			Panel A: Male Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.052*	-.015	.349	.668**	1.635**	.437
	(.031)	(.128)	(.233)	(.305)	(.658)	(.273)
Observations	1613	1613	1613	1613	1613	1613
R^2	.70	.58	.73	.58	.41	.66
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.853	.648	.137	.199	.086	.250
			Panel B: Female Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.208	-.150	2.152*	2.327**	5.784***	2.013**
	(.220)	(.141)	(1.197)	(.980)	(1.531)	(.950)
Observations	1612	1612	1612	1612	1612	1612
R^2	.70	.49	.54	.32	.32	.35
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.480	.782	.050	.149	.092	.107

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by four binary variables, corresponding to the top four quintiles of monsoon rainfall in the district in the year prior to the survey year. All specifications include district, state time trends and survey round fixed effects, alongwith controls for district-level socio-economic and demographic covariates.

Table 3.21: Short-Term Agricultural Productivity and Non-Farm Employment for Rural Workers: Robustness Check with Historical Rainfall Lags as Instrument

			Panel A: Male Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.052*	-.015	.349	.668**	1.635**	.437
	(.031)	(.128)	(.233)	(.305)	(.658)	(.273)
Observations	1613	1613	1613	1613	1613	1613
R^2	.70	.58	.73	.58	.41	.66
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.853	.648	.137	.199	.086	.250
			Panel B: Female Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.319	-.211	1.292	3.291***	3.695*	2.963**
	(.285)	(.153)	(1.633)	(1.263)	(1.976)	(1.238)
Observations	1342	1342	1342	1342	1342	1342
R^2	.73	.54	.56	.26	.42	.30
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.480	.782	.050	.149	.092	.107

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by contemporary and 3 historical lags in standardized monsoon rainfall in the district. All specifications include district, state time trends and survey round fixed effects, alongwith controls for district-level socio-economic and demographic covariates.

Table 3.22: Short-Term Agricultural Productivity and Non-Farm Employment for Rural Workers: Robustness Check with Rice Yields as Measure of Agricultural Productivity

			Panel A: Male Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Rice Yields, Lag 1	.045	.087	.239	.803**	2.109**	.324
	(.039)	(.114)	(.286)	(.362)	(.890)	(.321)
Observations	1508	1508	1508	1508	1508	1508
R^2	.70	.89	.74	.58	.41	.69
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.853	.648	.137	.199	.086	.250
			Panel B: Female Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Rice Yields, Lag 1	.187	.077	1.124	1.291	4.228***	1.017
	(.355)	(.200)	(1.084)	(.905)	(1.571)	(.861)
Observations	1507	1507	1507	1507	1507	1507
R^2	.76	.70	.59	.44	.39	.43
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.480	.782	.050	.149	.092	.107

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged rice yields in the year prior to the survey year. Rice yields in the year prior to the survey year is instrumented by standardized deviations in monsoon rainfall in the district. All specifications include district, state time trends and survey round fixed effects, alongwith controls for district-level socio-economic and demographic covariates.

3.11 Appendix

3.11.1 Summary Statistics

Table A3.1: Summary Statistics of Key Independent Variables

Variable	Mean	Std. Dev.	Min.	Max.	N
Yields, Lag 1	0.792	0.382	0	5.064	1546
Monsoon Rainfall Deviations, Lag 0	-0.31	0.894	-3.196	4.742	1619
Monsoon Rainfall Deviations, Lag 1	0.027	0.999	-3.329	5.356	1619
Monsoon Rainfall Deviations, Lag 2	-0.046	1.059	-3.097	4.113	1348
Monsoon Rainfall Deviations, Lag 3	0.361	1.178	-3.329	5.356	1348
Monsoon Rainfall Deviations, Lag 4	0.262	1.005	-2.299	3.842	1353
Simpson Crop Diversity Index	0.809	0.32	-5.847	0.997	1621
Share of Non-Agricultural Land	0.369	0.172	0.018	1.177	1624
Percent Land Irrigated	0.479	0.321	0.002	1	1624
Rate of Urbanization	0.744	0.176	0	1	1623
Share, Rural Workers, Secondary Educated	0.301	0.143	0.017	0.814	1619
Share, Rural Male Workers, Secondary Educated	0.369	0.154	0.031	0.877	1619
Share, Rural Female Workers, Secondary Educated	0.166	0.145	0	1	1618

Table A3.2: Summary Statistics of Key Outcome Variables - Rural Workers

Variable	Mean	Std. Dev.	Min.	Max.	N
LFP	0.672	0.111	0.353	0.953	1619
Share in Agriculture	0.698	0.145	0	0.993	1619
Share in Skilled Non-Farm	0.106	0.076	0	0.504	1619
Share in Unskilled Non-Farm	0.179	0.092	0	0.721	1619
Share in Manufacturing	0.085	0.062	0	0.488	1619
Share in Services	0.2	0.103	0.002	0.666	1619
Share in Skilled Manufacturing	0.026	0.027	0	0.226	1619
Share in Skilled Services	0.08	0.058	0	0.438	1619
Share in Unskilled Manufacturing	0.059	0.047	0	0.442	1619
Share in Unskilled Services	0.12	0.07	0	0.6	1619
Share in Construction	0.063	0.066	0	0.574	1619
Share in Food and Textiles	0.076	0.058	0	0.501	1619

Table A3.3: Summary Statistics of Key Outcome Variables - Rural Male Workers

Variable	Mean	Std. Dev.	Min.	Max.	N
LFP	0.853	0.063	0.563	0.982	1619
Share in Agriculture	0.648	0.152	0	0.987	1619
Share in Skilled Non-Farm	0.137	0.091	0	0.631	1619
Share in Unskilled Non-Farm	0.199	0.097	0	0.731	1619
Share in Manufacturing	0.086	0.056	0	0.379	1619
Share in Services	0.25	0.12	0.001	0.751	1619
Share in Skilled Manufacturing	0.031	0.031	0	0.333	1619
Share in Skilled Services	0.106	0.073	0	0.541	1619
Share in Unskilled Manufacturing	0.055	0.04	0	0.346	1619
Share in Unskilled Services	0.144	0.08	0	0.694	1619
Share in Construction	0.079	0.077	0	0.617	1619
Share in Food and Textiles	0.074	0.047	0	0.395	1619

Table A3.4: Summary Statistics of Key Outcome Variables - Rural Female Workers

Variable	Mean	Std. Dev.	Min.	Max.	N
LFP	0.48	0.211	0	0.942	1619
Share in Agriculture	0.782	0.189	0	1.001	1618
Share in Skilled Non-Farm	0.05	0.064	0	0.519	1618
Share in Unskilled Non-Farm	0.149	0.138	0	1	1618
Share in Manufacturing	0.092	0.123	0	0.927	1618
Share in Services	0.107	0.099	0	1	1618
Share in Skilled Manufacturing	0.018	0.032	0	0.395	1618
Share in Skilled Services	0.033	0.044	0	0.465	1618
Share in Unskilled Manufacturing	0.075	0.106	0	0.877	1618
Share in Unskilled Services	0.075	0.079	0	1	1618
Share in Construction	0.03	0.069	0	0.787	1618
Share in Food and Textiles	0.088	0.119	0	0.927	1618

3.11.2 Additional Results

Table A3.5: Short-Term Agricultural Productivity and Non-Farm Employment, OLS Estimates: Rural Workers

			Panel A: Male Workers			
	(1)	(2)	(3) Non-Farm	(4) Non-Farm	(5)	(6)
	LFP	Ag.	Skilled	Unskilled	Manuf.	Services
Yields, Lag 1	.008	.122	.023	.020	-.113	.011
	(.007)	(.093)	(.053)	(.057)	(.190)	(.052)
Observations	1613	1613	1613	1613	1613	1613
R^2	.71	.59	.74	.62	.50	.68
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.853	.648	.137	.199	.086	.250
			Panel B: Female Workers			
	(1)	(2)	(3) Non-Farm	(4) Non-Farm	(5)	(6)
	LFP	Ag.	Skilled	Unskilled	Manuf.	Services
Yields, Lag 1	.010	.181*	.274	-.333*	-.005	-.341**
	(.041)	(.102)	(.282)	(.177)	(.264)	(.162)
Observations	1612	1612	1612	1612	1612	1612
R^2	.70	.51	.56	.43	.49	.42
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.480	.782	.050	.149	.092	.107

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the rural labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. All specifications include district, state time trends, survey round fixed effects and district-level socio-economic and demographic covariates.

Table A3.6: Short-Term Agricultural Productivity and Non-Farm Employment: Urban Workers, Broad Categories

			Panel A: Male Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	.102*	-.721	.064	-.136	1.110	.040
	(.059)	(1.353)	(.096)	(.161)	(.816)	(.095)
Observations	1369	1369	1369	1369	1369	1369
R^2	.39	.41	.86	.67	.40	.55
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.799	.079	.511	.372	.198	.685
			Panel B: Female Workers			
	(1)	(2)	(3)	(4)	(5)	(6)
	LFP	Ag.	Non-Farm Skilled	Non-Farm Unskilled	Manuf.	Services
Yields, Lag 1	-.183	1.039	-.773	-.031	-1.306	-1.112
	(.339)	(1.838)	(1.046)	(.827)	(1.684)	(.803)
Observations	1361	1361	1361	1361	1361	1361
R^2	.54	.47	.57	.42	.38	.29
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Dependent Variable Mean (Share)	.228	.192	.315	.435	.233	.517

Standard errors in parentheses (clustered at the district level). The unit of observation is district-survey round. The dependent variable in each specification is the logged share of working age individuals in the urban labour force employed in each employment category. Skilled workers are those in non-elementary occupations and having completed a secondary education. The independent variable of interest is logged agricultural productivity in the year prior to the survey year. Agricultural productivity in the year prior to the survey year is instrumented by standardized deviation of district's monsoon rainfall from the district's historical mean. All specifications include district, state time trends, survey round fixed effects and district-level socio-economic and demographic covariates.

Chapter 4

References

[1] Acemoglu, D. and J. Robinson (2008). Persistence of Power, Elites and Institutions. *American Economic Review 98(1)*: 267-293.
[2] Acemoglu, D. and J. Robinson (2008). The Role of Institutions in Growth and Development. *Working Paper No. 10, Commission on Growth and Development, The World Bank.*
[3] Acemoglu, D., S. Naidu, P. Restrepo and J. Robinson (2013). Democracy, Redistribution and Inequality. Unpublished Working Paper.
[4] Acharya, A., J. E. Roemer and R. Somanathan (2012). Caste, Corruption and Political Competition in India. *Research in Economics 69(3)*: 336-352.
[5] Adelman, I. and S. Vogel (1992). "The Relevance of ADLI for sub-Saharan Africa" in *Industrialization Based on Agricultural Development: African Development Perspectives Yearbook*, H. Bass et al. ed. Bremen: Research Group on African Development Practices.
[6] Adena, M., R. Enikopolov, M. Petrova, V. Santarosa and E. Zhuravskaya (2016). Radio and the Rise of the Nazis in Pre-War Germany. *The Quarterly Journal of Economics*: 1885-1939.
[7] Agan, A. Y. (2011). Sex Offender Registries: Fear Without Function? *Journal of Law and Economics 54(1)*: 207-239.
[8] Altonji, J. G., T. E. Elder and C. R. Taber (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy 113(1)*: 1-54.
[9] Ansari, S., A. Verma and K. M. Dadkhah (2015). Crime Rates in India: A Trend Analysis. *International Criminal Justice Review 25(4)*: 318-336.
[10] Asher, S. and P. Novosad (2017). Politics and Local Economic Growth: Evidence from India. *American Economic Journal: Applied Economics 9(1)*: 229-273.
[11] Alvarez-Cuadrado, F. and M. Poschke (2011). Structural Change out of Agriculture: Labor Push Versus Labor Pull. *American Economic Journal: Macroeconomics 3(3)*: 127-158.
[12] Aterido, R., M. Hallward-Diemeier and C. Pages (2007). Investment Climate and Employment Growth: The Impact of Access to Finance, Corruption and Regulation Across Firms. *IZA DP No. 3138: Institute for the Study of Labour.*
[13] Aschauer (1989). Is Public Expenditure Productive? *Journal of Monetary Economics 23(2)*: 177-200.
[14] Autor, D., D. Dorn and G. Hanson (2013). The China Syndrome: Local Labor Market

Effects of Import Competition in the United States. *American Economic Review 103(6)*: 2121-2168.
[15] Banerjee, A. and E. Duflo (2007). The Economic Lives of the Poor. *Journal of Economic Perspectives 21(1)*: 141-168.
[16] Bannister, G. and K. Thuggee (2001). International Trade and Poverty Alleviation. *International Monetary Fund Working Paper WP/01/54*.
[17] Bardhan, P. (2016). State and Development: The Need for a Reappraisal of the Current Literature. *Journal of Economic Literature, 54(3)*: 862-892.
[18] Bardhan, P., D. Mookherjee and P. T. Monica (2010). Impact of Political Reservations in West Bengal Local Governments on Anti-Poverty Targeting. *Journal of Globalization and Development, 1(1)*: 1-38.
[19] Bastos, F. and J. Nasir (2004). Productivity and Investment Climate: What Matters Most? *World Bank Policy Research Paper 3335*: The World Bank.
[20] Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande and P. Topalova (2009). Powerful Women: Does Exposure Reduce Bias? *Quarterly Journal of Economics 124(4)*: 1497-1540.
[21] Beland, L. (2015). Political Parties and Labour Market Outcomes: Evidence from U.S. States. *American Economic Journal: Applied Economics 7(4)*: 198-220.
[22] Beland, L. (2016) and S. Oolomi. Party Affiliation and Public Spending: Evidence from U.S. Governors. *Economic Inquiry 55(2)*: 982-995.
[23] Besley, T. and R. Burgess (2000). Land Reform, Poverty Reduction and Growth: Evidence from India. *The Quarterly Journal of Economics 115(2)*: 389-430.
[24] Besley, T., R. Pande, L. Rahman and V. Rao (2004). The Politics of Public Goods Provision: Evidence from Indian Local Governments. *Journal of the European Economic Association 2(2-3)*: 416-426.
[25] Bhagwati, J. and T. Srinivasan (2002). Trade and Poverty in Poor Countries. *American Economic Review: Papers and Proceedings*, May 2002.
[26] Bhalhotra, S., G. Cassan, I. Clots-Figueras, and L. Iyer (2013). Religion, Politician Identity and Development Outcomes: Evidence from India. *Working Paper 13-102, Harvard Business School*.
[27] Bhalla, G.S. (1995): "Agricultural Growth and Industrial Development in Punjab." in *Agriculture on the road to Industrialization*, Mellor, J.W., ed. (p. 68-117). London: International Food Policy Research Institute.
[28] Bosworth, B. and S. Collins (2008). Accounting for Growth: Comparing India and China. *Journal of Economic Perspectives 22(1)*: 45-66.
[29] Bosworth, B., S. Collins and A. Virmani (2007). Sources of Growth in the Indian Economy. *NBER Working Paper No. 12901*. National Bureau of Economic Research (NBER), Cambridge, MA. http://www.nber.org/papers/w12901
[30] Bros, C. and M. Coutteneir (2015). Untouchability, Homicides and Water Access. *Journal of Comparative Economics 43(3)*: 549-558.
[31] Bryan, G., S. Chowdhury, and A. M. Mobarak (2014). Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh. *Econometrica 82(5)*: 1671-1748.
[32] Bursztyn, L. (2016). Poverty and Political Economy of Public Education Spending: Evidence from Brazil. *Journal of the European Economic Association 14(5)*: 1101-1128.
[33] Bustos, P., B. Caprettini and J. Ponticelli (2016). Agricultural Productivity and Struc-

tural Transformation: Evidence from Brazil. *American Economic Review 106(6)*: 1320-65.
[34] Calderon, C., E. Moral-Benito and L. Servin (2011). Is Infrastructure Capital Productive? A Dynamic Heterogeneous Approach. *Policy Research Working Paper 5862*. Development Research Group, Macroeconomics Growth Team, The World Bank.
[35] Cao, K. H. and J. A. Birchenall (2013). Agricultural Productivity, Structural Change and Economic Growth in Post-Reform China (2013). *Journal of Development Economics 104(2013)*: 165-180.
[36] Chand, R. and S. Parappurathu (2012). Historical and Spatial Trends in Indian Agriculture: Growth Analysis at National and State Level. *Economic and Political Weekly XLVIII (26 and 27)*: 55-63.
[37] Chandra, K. *Why Ethnic Parties Succeed: Patronage and Ethnic Head Counts in India*. New York: Cambridge University Press.
[38] Chattopadhyay, R. and E. Duflo (2004). Evidence from a Nationwide Randomized Experiment. *Economic and Political Weekly, 39(9)*: 979-986.
[39] Chattopadhyay, R. and E. Duflo (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India.*Econometrica, 72(5)*: 1409-1443.
[40] Chin, A. and N. Prakash (2011). The Redistributive Effects of Political Reservations for Minorities: Evidence from India. *Journal of Development Economics 96(1)*: 265-277.
[41] Chockalingham, K. (2003). Criminal Victimization in Four Major Cities in Southern India. *Forum on Crime and Society 3(1-2)*: 117-126.
[42] Christiansen, L., J. De Weerdt and Y. Todo (2013). Urbanization and Poverty Reduction: The Role of Rural Diversification and Secondary Towns. *Agricultural Economics 44(2013)*: 435-447.
[43] Clots-Figueras, I. (2011). Women in Politics: Evidence from Indian States. *Journal of Public Economics 95*: 664-690.
[44] Cole, S. (2009). Fixing Market Failures or Fixing Elections? Agricultural Credit in India. *American Economic Journal: Applied Economics 1(1)*: 219-250.
[45] Colmer, J. (2016). Weather, Labor Reallocation and Industrial Production: Evidence from India. *Unpublished Working Paper*.
[46] Datt, G. and M. Ravallion (1998). Farm Productivity and Rural Poverty in India. *Journal of Development Studies 34(4)*: 62-85.
[47] Datt, G. and M. Ravallion (2010). Has India's Growth Become More Pro-Poor in the Wake of Economic Reforms? *World Bank Economic Review 25(2)*: 157-189.
[48] Dal Bo, E. and M. Rossi (2011). Term Length and the Effort of Politicians. *Review Economic Studies 78(4)*: 1237-1263.
[49] de Janvry, A. and E. Sadoulet (2001). Income Strategies among Rural Households in Mexico: The Role of Non-Farm Activities. *World Development 29(3)*: 467-480.
[50] de Janvry, A. and E. Sadoulet (2010). Agricultural Growth and Poverty Reduction: Additional Evidence. *The World Bank Research 25(1)*: 1-20.
[51] Deither, J., M. Hirn and S. Straub (2010). Explaining Enterprise Performance in Developing Countries with Business Climate Survey Data. *The World Bank Research Observer*: 1-52.
[52] Dell, M., B. F. Jones and B. Olken (2012). Temperature Shocks and Economic Growth: Evidence from the Last Half Century. *American Economic Journal: Macroeconomics 4(3)*: 66-95.

[53] Dell, M., B. F. Jones and B. Olken (2014). What do we Learn from the Weather? The New Climate-Economy Literature. *Journal of Economic Literature 52(3)*: 740-798.
[54] Deshpande, A. (2011). *The Grammar of Caste: Economic Discrimination in Contemporary India*. New Delhi: Oxford University Press.
[55] Dollar, D., M. Hallward-Driemeier and T. Mengistae (2005). Business Climate and Firm Performance in Developing Economies. *Economic Development and Cultural Change 54(1)*: 1-31.
[56] Donohue, J. J. and S. D. Levitt (2001). The Impact of Race on Policing and Arrests. *Journal of Law and Economics 44(2)*: 233-261.
[57] Dreze, J. and A. Sen (2013). *An Uncertain Glory: India and its Contradictions*. Princeton: Princeton University Press.
[58] Dreze, J. and R. Khera (2000). Crime, Gender and Society in India. *Population and Development Review 26(2)*: 335-352.
[59] Duflo, E. (2012). Women Empowerment and Economic Development. *Journal of Economic Literature, 50(4)*: 1051-1079.
[60] Dunning, T. and J. Nilekani (2013). Ethnic Quotas and Political Mobilization: Caste, Parties and Distribution in Indian Village Councils. *American Political Science Review 107(1)*: 35-56.
[61] Emerick, K. (2016). Agricultural Productivity and Sectoral Reallocation of Labour in Rural India. *Unpublished Working Paper*.
[62] Enikolopov, R. (2014). Politicians, Bureaucrats and Targeted Redistribution. *Journal of Public Economics 120*: 74-83.
[63] Erosa, A., T. Koreshkova, and D. Restuccia (2010). How Important is Human Capital? A Quantitative Assessment of World Income Inequality. *Review of Economic Studies 77(4)*: 1421-1449.
[64] Escribano, A. and Guasch, J. (2005). Assessing the Impact of Investment Climate on Productivity Using Firm Level Data Methodology and the Cases of Guatemala, Honduras, and Nicaragua. *Discussion Paper WPS 3621*:The World Bank.
[65] Escribano, A., J.L. Guasch, and J. Pena (2010). Assessing the Impact of Infrastructure Quality on Firm Productivity in Africa. *Policy Research Working Paper 519, The World Bank*.
[66] Fajnzylber, P., D. Lederman, and N. Loayza (2002). Inequality and Violent Crime. *Journal of Law and Economics 45(1)*: 1-39.
[67] Foster, A. D. and M. R. Rosenzweig (2010). Agricultural Productivity Growth, Rural Diversity and Economic Reforms: India, 1970-2000. *Economic Development and Cultural Change 52(3)*: 509-542.
[68] Foster, A. D. and M. R. Rosenzweig (2010). Is there Surplus Labour in Rural India? *Center Discussion Paper No. 901*: Economic Growth Center, Yale University.
[69] Ferreira, F. and J. Gyourko (2009). Do Political Parties Matter? Evidence from U.S. Cities. *Quarterly Journal of Economics, February 2009*: 399-422.
[70] Frias, J., D. Kaplan, and E. Verhoogen (2012). Exports and Within-Wage Plant Distributions: Evidence from Mexico. *American Economic Review: Papers and Proceedings 102(3)*: 435-440.
[71] Hanchinamani, B. (2001). Human Rights Abuses of Dalits in India. *Human Rights Brief 8(2)*: 15-19.

[72] Gollin, D., S. Parente and R. Rogerson (2002). The Role of Agriculture in Development. *American Economic Review 92(2)*: 160-164.
[73] Gollin, D. (2010). "Agricultural Productivity and Economic Growth." in *Handbook of Agricultural Economics, Vol. 4*, R. E. Evanson and P. Pingali ed. Amsterdam: North Holland.
[74] Hazell, P. B. S. and A. Roell (1983). Rural Growth Linkages: Household Expenditure Patterns in Malayasia and Nigeria. *Research Report No. 41*: International Food Policy Research Institute.
[75] Henderson, J. V., A. Storeygard and U. Deichmann (2017). Has Climate Change Driven Urbanization in Africa? *Journal of Development Economics 124(2017)*: 60-82.
[76] Himangshu, P. Lanjouw, A. Mukhopadhyay and R. Murgai (2011). Non-Farm Diversification and Rural Poverty Decline: A Perspective from Indian Sample Survey and Village Study Data. *Asia Research Centre Working Paper 44*: Asia Research Centre, London School of Economics.
[77] Hodler, R. and P. Raschky (2014). Regional Favouritism. *Quarterly Journal of Economics*: 995-1033.
[78] Irudayam, A., J. P. Mangubhai, and J. G. Lee (2012). *Dalit Women Speak Out: Caste, Class and Gender Violence in India*. Zubaan: Zubaan.
[79] Iyer, L., A. Mani (2012). Traveling Agents: Political Change and Bureaucratic Turnover in India. *The Review of Economics and Statistics 94(3)*: 723-739.
[80] Iyer, L., A. Mani, P. Mishra, and P. Topalova (2011). The Power of Political Voice: Womens Political Representation and Crime in India. *Working Paper 11-092, Harvard Business School*.
[81] Iyer, L., T. Khanna, and A. Varshney (2011). Caste and Entrepreneurship in India. *Working Paper 12-028, Harvard Business School*.
[82] Iyer, L. and P. Topalova (2014). Poverty and Crime: Evidence from Rainfall and Trade Shocks in India. *Harvard Business School Working Paper 14-067*.
[83] Irz, X., L. Lin, C. Thirtle and S. Wiggins (2001). Agricultural Productivity Growth and Poverty Alleviation. *Development Policy Review 19(4)*: 449-466.
[84] Jayachandran, S. (2006). Selling Labour Low: Wage Responses to Productivity Shocks in Developing Countries. *Journal of Political Economy 114(3)*: 538-575.
[85] Jacoby, H. (2013). Food Prices, Wages and Welfare in Rural India. *Policy Research Working Paper 6412*: Development Research Group, The World Bank, Washington D.C.
[86] Jaffrelot, C. (2003). *Indias Silent Revolution: The Rise of the Lower Castes in North India*. New York: Columbia University Press.
[87] Jaffrelot, C. (2006). The Impact of Affirmative Action in India: More Political than Socio-economic. *India Review 5(2)*: 173-189.
[88] Jensenius, F. R. (2015). Development from Representation? A Study of Quotas for the Scheduled Castes in India. *American Economic Journal: Applied Economics 7(3)*: 196-220.
[89] Kaur, S. (2015). Nominal Wage Rigidity in Village Labor Markets. *Unpublished Working Paper*.
[90] Kijima, Y. (2006). Caste and Tribe Inequality: Evidence from India, 1983-1999. *Economic Development and Cultural Change 54(2)*: 369-404.
[91] Kijima, Y., T. Matsumoto and T. Yamano (2006). Non-Farm Employment, Agricultural Shocks and Poverty Dynamics: Evidence from Rural Uganda. *Agricultural Economics*

35(2006): 459-467.
[92] Kling, J., J. Liebman and L. Katz (2007). Experimental Analysis of Neighborhood Effects. *Econometrica 75(1)*: 83-119.
[93] Kotwal, A., B. Ramaswami and W. Wadhwa (2011). Economic Liberalization and Indian Economic Growth: Whats the Evidence? *Journal of Economic Literature 49(4)*: 1152-1199.
[94] Krishnan, M. (2014). Rape and India's Untouchable Caste.
[95] Kumar, N. (2014). Politics and Real Firm Activity: Evidence from Distortions in Bank Lending in India. *Job Market Paper, Booth School of Business, University of Chicago*
[96] Krishnan, N. (2007). Political Reservations and Rural Public Good Provision in India. Working Paper.
[97] Lanjouw, J. O. and P. Lanjouw (2001). The Rural Non-Farm Sector: Issues and Evidence from Developing Countries. *Agricultural Economics 26(1)*: 1-23.
[98] Lanjouw, P. and R. Murgai (2009). Poverty Decline, Agricultural Wages and Non-Farm Employment in Rural India: 1983-2004. *Agricultural Economics: 40(2)*: 243-263.
[99] Laasko, M. and R. Taagepera (1979). Effective Number of Parties. *Comparative Political Studies 12(1)*: 3-27.
[100] Lee, D. (2001). The Electoral Advantage to Incumbency and Voters Valuation of Politicians Experience: A Regression Discontinuity Analysis of Elections to the U.S. House. *Working Paper 8441, National Bureau of Economic Research.*
[101] Lee, D., E. Moretti, and M.J. Butler (2004). Do Voters Affect or Elect Policies? Evidence from the U.S. House. *The Quarterly Journal of Economics, 119(3)*: 807-859.
[102] Lee, S and B. A. Malin (2013). Education's Role in China's Structural Transformation. *Journal of Development Economics 101(2013)*: 148-166.
[103] Lizzeri, A. and N. Persico (2001). The Provision of Public Goods Under Alternative Electoral Incentives. *American Economic Review 91(1)*: 225-239.
[104] Mas, A. and E. Moretti (2009). Racial Bias in the 2008 Presidential Election. *American Economic Review 99(2)*: 323-329.
[105] Matsuyama, K. (1992). Agricultural Productivity, Comparative Advantage and Economic Growth. *Journal of Economic Theory 58 (2)*: 317-334.
[106] McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics 142(2)*: 698-714.
[107] Mellor, J. W. (1995). "Introduction" in Agriculture on the Road to Industrialization, Mellor, J. W. ed (p. 2-22) London: International Food Policy Research Institute.
[108] Meltzer, A.H. and S.F. Richard (1981). A Rational Theory of the Size of Government. *The Journal of Political Economy, 89(5)*: 914-927.
[109] Meyersson, E. (2014). Islamic Rule and the Empowerment of the Poor and Pious. *Econometrica 82(1)*: 229-269.
[110] Michaels, G., F. Rauch and S. J. Redding (2012). Urbanization and Structural Transformation. *The Quarterly Journal of Economics 127(2)*: 535-586.
[111] Nagavarapu, S. and S. Sekhri (2014). Informal Monitoring Mechanisms in Public Service Delivery: Evidence from the Public Distribution System in India. Working Paper
[112] Novosad, P. and S. Asher (2013). Politics and Local Economic Growth: Evidence from India. Working Paper.
[113] Nyiri, Z. and C. English (2007). Is Europe a Good Place for Racial and Ethnic Minorities?

[114] Pande, R. (2003), Can Mandated Political Representation Provide Disadvantaged Minorities Policy Influence? Theory and Evidence from India. *American Economic Review, 93 (4)*: pp. 1132-1151.
[115] Persson, T. and G. Tabellini (2000). "Political Economics: Explaining Economic Policy." Cambridge: The MIT Press.
[116] Rehavi, M. (2007). Sex and Politics: Do Female Legislators Affect State Spending? *Unpublished Working Paper.*
[117] Reardon, T., J. Berdegue and G. Escobar (2001). Rural Non-Farm Employment and Incomes in Latin America: Overview and Policy Implications. *World Development 29(3)*: 395-409.
[118] Restuccia, D., D. T. Yang, and X. Zhu (2008). Agriculture and Aggregate Productivity: A Quantitative Cross Country Analysis. *Journal of Monetary Economics 55(2)*: 234-250.
[119] Sharma, S. (2015). Caste-based Crimes and Economic Status: Evidence from India. *Journal of Comparative Economics 43(1)*: 204-226.
[120] Singh, D. (1996). The International Crime (Victim) Survey in Bombay. in U. Zvekic and A. A. D. Frate (ed.) *Criminal Victimization in the Developing World* p. 89-134. Darby, PA: Diane Publishing
[121] Stephens-Davidowitz, S. (2014). The Cost of Racial Animus on a Black Candidate: Evidence using Google Search Data. *Journal of Public Economics 118*: 26-40
[122] Thorat, S. and K. Newman (2010). *Blocked by Caste: Economic Discrimination in Modern India.* New Delhi: Oxford University Press.
[123] Topalova, P. (2010). Factor Immobility and Regional Effects of Trade Liberalization: Evidence on Poverty and Inequality from Indian Districts. *American Economic Journal: Applied Economics 2(4)*: 1-41.
[124] Trebbi, F., P. Aghion and A. Alesina (2008). Electoral Rules and Minority Representation in U.S. Cities. *Quarterly Journal of Economics 123(1)*: 325-357.
[125] United Nations (2012). Human Rights in India: Status Report. Technical Report, WGHR, New Delhi.
[126] Unnithan, N. P. (2013). textitCrime and Justice in India. Fort Collins: SAGE Law.