CSAE Working Paper WPS/2016-36

Imposing institutions:

Evidence from cash transfer reform in Pakistan

Muhammad Haseeb* Kate Vyborny†‡

November 18, 2016

Abstract

Institutions are recognized as critical for development, but there is limited evidence on whether policies designed to improve them can be effective. In this paper, we quantify the impact of an outside agency imposing on a developing country a new policy designed to improve a key government function: the allocation of public spending. As a condition of aid to the Pakistasn, international donor agencies imposed a new system for the country to select recipients for cash transfers. Before this reform, households in winning politicians' villages were 200-400% more likely to receive cash transfers than those in rivals' villages. The reform reduced favoritism for the best connected households: politicians' clan members in their villages. Some connected households continued to receive more transfers, likely because politicians assisted them in overcoming administrative hurdles. However, the reform improved targeting substantially in politicians' own villages and province-wide. As a result of this imposed policy, the public legitimacy of government transfer programs increased by about 40%.

*Ph.D. candidate, University of Warwick; former Research Fellow, Lahore School of Economics.

†Postdoctoral Associate, Department of Economics, Duke University; Visiting Research Fellow, Lahore School of Economics. Corresponding author: katherine.vyborny@duke.edu.

‡We are grateful for guidance from Marcel Fafchamps. We have benefited from feedback on earlier drafts from Madiha Afzal, Sami Bazzi, Alan de Brauw, Mike Callen, James Fenske, Jenny Guardado, Kaivan Munshi, Jake Shapiro, Maximo Torero, and Laura Zimmerman, and helpful conversations with Kathleen Beegle, Erlend Berg, Azam Chaudhry, Ali Cheema, Jacobus Cilliers, Michael Clemens, Julie Berry Cullen, Clement de Chaise Martin, Mirko Draca, Pascaline Dupas, Erica Field, Claudio Ferraz, Haris Gazdar, Matthew Gentzkow, Mike Geruso, Naved Hamid, Clement Imbert, Herbert Kitschelt, Chris Ksoll, Julien Labonne, Clare Leaver, Steve Lyon, Nicolas Martin, Ted Miguel, Ijaz Nabi, Suresh Naidu, Matthew Nelson, Ben Olken, Simon Quinn, Jake Shapiro, Bilal Siddiqi, Duncan Thomas, Milan Vaishnav, Xiao Yu Wang, Xiaoxue Zhao, and participants in seminars at Duke, Oxford, CGD, IFPRI, LUMS, AIMS, and the CSAE, RECODE, DIAL, and MWIEDC conferences. We thank Julien Labonne for providing guidance on and carrying out the third-party data split, and Jake Shapiro for sharing data on national elections in Pakistan. We appreciate help from Naveed Akbar, Saleem Baloch, Ali Cheema, Ijaz Nabi, Cem Mete, Muhammad Tahir Noor, Khurram Shahzad, and Khaleel Tetlay in understanding the institutional context of BISP, and the Punjab Bureau of Statistics for feedback during questionnaire development. We thank Misha Saleem, Amber Nasir, and Abbas Raza for research assistance, Tamiah Nasir, Sayaf Naseem, Hamid Tiwana, Fahad Manzoor, Mahniya Zafar and Sila Aqsa for help with data collection and cleaning, and the Centre for the Study of African Economies at Oxford, particularly Rose Page, Richard Payne and Gail Wilkins, and the Lahore School of Economics, particularly Zenab Naseem, for organizational support. We gratefully acknowledge funding from the British Academy International Partnerships Initiative and the Lahore School of Economics.

1 Introduction

One of the key functions of public institutions is to allocate public spending. For some types of public expenditure, this includes deciding how much will be targeted to different geographic areas or even individuals (such as social safety nets). In contexts with weak institutions, public officials may direct spending towards groups or individuals they prefer for personal or political reasons: their friends, relatives, ethnic groups, home regions, or political supporters. A growing body of literature has established that this kind of favoritism by officials in this allocation can be substantial.

International agencies (such as the World Bank) often try to impose policy changes to improve institutions in developing countries, in part to reduce elite capture and favoritism in public spending. But micro evidence from a range of contexts has demonstrated remarkable persistence in political and social institutions Acemoglu *et al.* (2001, 2014, 2013); Dell (2010); Banerjee and Iyer (2005), and there is an active debate on whether deliberate attempts to reform institutions as a policy, especially by outsiders, can actually be effective (Acemoglu and Robinson, 2008; Banerjee and Duflo, 2014).

In this paper, we provide quantitative evidence on the effectiveness of institutional reform imposed from outside. We study a case in which international donor agencies imposed a new system for selecting recipients of public funds in Pakistan as a condition for funding. Previous mechanisms for selecting recipients for targeted public funds were discretionary and allowed room for officials to choose their preferred recipients. The new system comprised collection of nationwide data from field surveys, gathering it into a centralized database, the formation of a body at the federal government level responsible for managing the data and identifying recipients, and the imposition of a strict formula for selection of recipients. We study the effect of this institutional reform.

There were delays during the negotiation and setup of this new system. In a hurry to start the transfer program, the government of Pakistan used an alternative targeting system: it designated powerful national politicians to select recipients from among their constituents. This interim process was similar to pre-existing processes used for targeting resources in Pakistan in that the choice of recipients was largely at the discretion of a government official. The process of implementing the new scorecard began one year later in selected pilot districts, and was then phased in to the rest of the country. The program was scaled up, but all other features of its design and administration stayed the same during this period. We compare outcomes before and after the implementation of the reform to identify its effects on favoritism in distribution of the cash transfer, effective targeting of the poor, and the perceived legitimacy of the government social safety net program.

To identify the effect of targeting on favoritism, we collect a primary survey in the origin villages of Members of the National Assembly, and their rivals: immediate runners-up, or the winners and runners-up in the last election cycle. We identify the origin villages where these politicians were born or have ancestral land; thus politicians

cannot select their origin villages. This sampling strategy allows us to compare households who have a connection to winning and rival politicians, before and after the reform.

We find that before the reform, favoritism was large. Households from the origin villages of winning politicians were between 200 - 400% more likely to receive the cash transfer than those in the other politicians' villages. After the reform, we find some evidence suggesting that winning politicians' villages continued to experience an advantage in receiving BISP. However, it did not increase, even as the BISP program was scaled up dramatically. We find a significant reduction in favoritism for the best-connected group - the politician's own clan in his origin village.

One possible interpretation of an observed advantage for households connected to a politician is that he selects the neediest recipients from among those for whom he has information: those in his social network. The politician might even have better information on household poverty for those in his network than that measured through the proxy means test (Conning and Kevane, 2000; Bardhan and Mookherjee, 2000, 2005; Niehaus *et al.*, 2013; Alatas *et al.*, 2012; Basurto *et al.*, 2016; Alderman, 2002).[1] We rule out these possibilities by documenting that (a) the reform decreased transfers to wealthier households, even as the overall level of BISP increased and even when measuring wealth based on indicators not included in the scorecard; (b) the government eliminated half of the households that politicians nominated because they were ineligible even under a very minimal set of criteria imposed pre-reform, and 75% of those that were accepted were then eliminated as ineligible under the new criteria after the reform; and (c) even in the politicians' villages many observably poor households were not included before the reform.

In contexts with weak institutions, voters may see capture and/or clientelism as legitimate, because ethnic or caste networks help to hold politicians to account to deliver any goods or assistance, or because capture is seen as the price paid for elites bringing in resources that would not come otherwise (Keefer and Vlaicu, 2007; Wade, 1985; Platteau, 2004; Anderson *et al.*, 2015). It is not obvious that the BISP reform would necessarily improve the perception of the program or government spending. We use the phase-in of the reform to test for this. Controlling for the scale-up of the program, we find that the reform substantially improved the perceived legitimacy of the government's transfer programs across the province as a whole. Thus a reform imposed by outside donors had a substantial impact on the political legitimacy of government activity.[2]

We discuss possible mechanisms for the continued advantage that winners' villages appear to experience even after the reform. We rule out "grandfathering in" of households who had previously received the program; most pre-reform recipients were disenrolled, and effects persist when excluding all pre-reform recipients. In contrast, we find that assistance in getting ID cards (a prerequisite for receiving the cash transfer) played a significant role.

[1] Another possibility is that the effect represents politicians elected from among a disadvantaged group successfully representing their interests (cf. Besley *et al.* (2004, 2007)).

[2] In this respect, our results are more in line with Banerjee *et al.* (2014), who use a survey vignette experiment to identify voter preferences for politician clientelism, holding other politician attributes constant.

Information or assistance about the appeals process may also play a role. This demonstrates the importance of administrative barriers to takeup (cf. Duclos (1995); Currie (2004); Coady *et al.* (2004); Besley (1990)). These have sometimes been used as "ordeal mechanisms" to elicit self-selection of the most needy or those most likely to make use of a transfer or subsidized good (Alatas *et al.*, 2013b; Dupas *et al.*, 2016). However, our results highlight the fact that these may lead to elite capture, as elites help their connections overcome administrative barriers, even when eligibility itself is not discretionary.

One weakness of targeting systems based on survey data is the potential for strategic misreporting. Such a system may be more vulnerable to misreporting than targeting by officials in each area, because such officials may have independent information about households' poverty status, or even be able to punish misreporting. Positing that households are likely to misreport consistently on different surveys to avoid any risk of being caught in a discrepancy, we compare reported observable and unobservable household assets to test for misreporting in the survey data we study. We do not find any evidence of misreporting, although we cannot rule out that there may have been misreporting on the original scorecard survey, or that it could occur in future updates to the survey.

The identification strategy we use in our politician village sample is based on comparing the home villages of competing politicians; thus the sampling excluded areas that have been consistently won by the same candidate in a landslide. The estimates are a Local Average Treatment Effect for politicians' villages in competitive constituencies. How this compares to the average treatment effect could depend on whether the favoritism effect we estimate reflects clientelism, i.e. politicians using BISP to win votes, or capture, i.e. politicians using BISP to benefit themselves and their connections for personal reasons (cf. Bardhan and Mookherjee (2012)). We test whether the effects generalize to nearby villages sampled in a subset of constituencies, and find a very similar pattern of results. In addition, the results from our analysis of targeting and legitimacy are based on a provincially representative survey. Here data are province-wide identification is based on the phase-in of the reform in pilot districts which appear similar to the province as a whole. We find significant and large improvements in both targeting and legitimacy.

While the Members of the National Assembly may have been able to assist households in their villages to receive BISP after the reform, we do find evidence that even these powerful politicians were constrained. The reform reduced their ability to provide benefits to their closest connections. Overall, the reform substantially improved both the program's targeting of the poor and perceptions of its fairness among the population.

Collecting data for a proxy means test can be costly, especially in contexts where there is no central administrative source of data such as tax returns; in some circumstances this could exceed the benefits of improved targeting (Besley, 1990). BISP officials estimate the cost of implementing the reform, including collecting and checking the national survey and processing it to identify recipients, at PKR 5.21 billion ($52 million USD). This comes to less than 2% of the nearly $4 billion paid out to recipients until 2016, at which point the government plans to repeat the

exercise. While winners' villages may still experience some advantage, the large improvement in targeting quality and perceived legitimacy are a substantial value for this investment.

The international aid donors led the technical development of the new system and have advocated strongly for its continuation and the continuation of the cash transfer program. They supported the transfer program itself financially, but it is mostly funded by the government of Pakistan, and administered by government agencies. As of 2013, only 10% of the program's funding came from donors (Asian Development Bank, 2013). The government has extended the BISP targeting system into other programs, such as a conditional cash transfer program (Nabi, 2013). In addition, in 2013, a different party was elected to the national government in Pakistan; the new government has increased funding to the cash transfer program, and is working with the donors to carry out a new survey to update the poverty scorecard data. This is an unusual occurence in Pakistan, where governments tend to de-fund programs established under previous governments and replace them. Some of this continued investment in the program may be due to the donor agencies exerting overall pressure on the government - using a larger portfolio of aid as leverage, rather than only the funding for the program itself.[3] However, donors often have difficulty enforcing conditions of economic policy or good governance on recipients because of administrative pressure to continue the flow of funding or because of political considerations (particularly when the recipient state is a geostrategic priority) (Kanbur, 2000; Montinola, 2010; Molenaers *et al.*, 2015; Bourguignon and Platteau, 2014; Santiso, 2001; Killick, 1998; Kilby, 2009). Thus the resilience of this new institution for allocating public funds is impressive.

Our study contributes to two main strands of literature. The first includes studies that quantify favoritism (Fafchamps and Labonne, 2016a; Caeyers and Dercon, 2012; Burgess *et al.*, 2015; Hsieh *et al.*, 2011; Besley *et al.*, 2004, 2012; Carozzi and Repetto, 2014; Hodler and Raschky, 2014; Do *et al.*, 2016; Mu and Zhang, 2011), and assess what kinds of institutional features are associated with its prevalance, such as periods of greater democratic institution, better institutions overall, or higher education levels (Kitschelt and Wilkinson, 2007; Weitz-Shapiro, 2012; Stokes *et al.*, 2013; Burgess *et al.*, 2015; Hodler and Raschky, 2014). In this literature, however, it is difficult to identify the effect of a specific policy change that policymakers could intentionally try to replicate.

The second strand of literature examines whether deliberate policy efforts can effectively improve institutions. Banerjee and Duflo (2014) provide an in-depth review of this issue. Of particular interest in our context are studies that examine whether international agencies can leverage foreign aid to effect institutional change in a country. This is particularly important given the concern that overall, aid flows may weaken institutions in recipient states (Moss *et al.*, 2006; Djankov *et al.*, 2008; Bräutigam and Knack, 2004). The programs the donors themselves are funding may be affected by clientelism and capture (Briggs, 2014; Jablonski, 2014; Jayne *et al.*, 2001; Öhler and Nunnenkamp, 2014; Hodler and Raschky, 2014). Donors often seek to use their financial leverage as well as technical

[3]The International Monetary Fund has continued to incorporate BISP spending targets as a part of its performance criteria in lending to Pakistan (International Monetary Fund, 2016)

input on institutional design to reduce this influence - not only over their own funds, but also on the allocation of the government's own resources. A number of studies using country-level data examine the difficulty donors have had in enforcing such conditions (Kanbur, 2000; Montinola, 2010; Molenaers *et al.*, 2015; Bourguignon and Platteau, 2014; Santiso, 2001; Killick, 1998; Kilby, 2009).

A well-developed micro experimental literature tests the impact of specific interventions in governance to reduce capture and corruption, or improve the allocation of public resources. These interventions range from systematic release of information to new technologies such as biometric ID cards, for example Casey *et al.* (2012); Humphreys *et al.* (2015); Banerjee *et al.* (2015, 2016); Björkman-Nyqvist *et al.* (2014); Ravallion *et al.* (2013); Callen *et al.* (2016); Pande (2011); Muralidharan *et al.* (2016); Fujiwara and Wantchekon (2013); Ferraz and Finan (2008); Wantchekon (2003)). A number of experimental and non-experimental studies have looked specifically at variations in targeting regimes for social safety net programs (Alatas *et al.*, 2012; Bardhan and Mookherjee, 2006; Galasso and Ravallion, 2005; Kilic *et al.*, 2013). In some sense all these studies could be seen as testing an effort to improve institutions. In some cases the interventions are carried out by a donor (or researcher), in others they are done with government participation or ownership. Some of these studies have involved new institutions set up under pressure from donors in aid-dependent states (e.g. Beath *et al.* (2013)). However, these studies generally vary constraints on officials at a local level; they do not allow us to examine whether high-level officials can be constrained by reforms imposed from outside.

Our study focuses on favoritism by national politicians. This may be quite different from the behavior of local elites in decentralized programs. Powerful national politicians may have power to get around formalized systems built into programs to reduce their manipulability. They are based in the capital, unlike local elites in recipient villages, which may give them more access to interfere with centralized institutions. In many contexts, including South Asia, politicians have significant power over appointment and reassignment of bureaucrats (e.g. Iyer and Mani (2012); Callen *et al.* (2016); Nath (2016); Gulzar and Pasquale (2016)). They may be able to use this to get around a targeting system implemented by bureaucrats. Studies of local elites have often shown relatively small welfare effects (Alatas *et al.*, 2013a; Bardhan and Mookherjee, 2006; Basurto *et al.*, 2016). In contrast, work on capture and favoritism by high level officials have found large effect sizes, particularly in contexts where institutions are weak (Hodler and Raschky, 2014; Burgess *et al.*, 2015; Fafchamps and Labonne, 2016a). However, these studies often do not have the kind of detailed household-level information which we use in our survey to measure the impacts on patterns of distribution between different groups and on the accuracy of targeting, and to our knowledge none of them have variation in deliberate policies intended to constrain these officials.

Our paper contributes to this literature in three ways. First, we study the causal impact of a specific institutional reform that was imposed by an outside party (international aid donors). Thus the results directly speak to the

question of whether deliberate policies can improve institutions (Banerjee and Duflo, 2014). Second, the external validity of our study is high: the variation we study occurred as a kind of natural experiment as a result of negotiation between the donors and the government of Pakistan, and the implementation of both the pre and post reform schemes occurred on a national scale. Third, we are able to test the impact of this reform on high-level national officials, who are likely to have greater power to capture resources and to find ways around institutions designed to constrain them.

To our knowledge, this is the first paper studying the effectiveness of a specific policy change in a natural setting designed to reduce favoritism by high-level officials in the distribution of public funds. Politician favoritism has large effects worldwide and in particular in weak institutional contexts; public agencies have devoted substantial resources to try to set up new institutional mechanisms that can constrain it.

The rest of the paper proceeds as follows. Section 2 describes the context; Section 3 details the primary and secondary data. Section 4 lays out our empirical strategy. Section 5 presents the results, and Section 6 concludes.

2 Context

Pakistan is a major recipient of foreign aid, receiving $3.1 billion USD in official development assistance in 2014, or $20 per capita (OECD DAC). However, with this amount making up 1.4% of gross national income, it is not among the most heavily aid dependent states, 48 of which receive more than 5% of GNI in aid.[4] Thus it may have more bargaining power with aid donors than these most aid dependent states. Pakistan also suffers from weak institutions, with state resources often subject to capture by politicians and the elite (cf. Khwaja and Mian (2005); Cheema *et al.* (2009, 2012); Cheema and Mohmand (2009)); this is a persistent concern for aid agencies as well as for civil society within the country.

In 2008, a group of international aid donors including the World Bank, UK Department for International Development, US Agency for International Development, and the Asian Development Bank worked with the newly elected national government in Pakistan to set up a major cash transfer program. The government, led by the Pakistan People's Party (PPP), dubbed the program the Benazir Income Support Program after their recently assassinated leader, Benazir Bhutto. The BISP is an unconditional transfer; once selected, recipients are supposed to receive 1000 PKR per month (USD 10) indefinitely, or until actively identified as no longer eligible.[5] The program has 7 million recipient households, and has paid out a total of almost $4 billion as of 2016, making it by far the

[4]These figures do not include military aid, which has been substantial in recent years. However, military aid is typically conditioned on security cooperation, not institutional reforms of the kind we study in this paper.

[5]The monthly payment was later increased to 1200 PKR to adjust for inflation. Cheema *et al.* (2014, 2015) find that household reports and BISP's administrative data reflect that households did miss some installments. Because our data is recall based, we focus on receiving the transfer as a binary variable, which is a salient event and less likely to be subject to recall bias.

largest social protection program ever implemented in Pakistan, and one of the largest in the world.[6]

The launch of the BISP is part of an international trend. Over the past two decades, new targeted transfers have been introduced in many developing countries, with an estimated 150 million recipient households worldwide as of 2008 (Barrientos and Hulme, 2009). Yet programs that are targeted to households or individuals may be particularly vulnerable to elite capture or political influence on targeting, more so than other forms of public spending.[7] In the case of the BISP, opponents of the governing PPP party immediately objected that it was politicized; the program's branding with the name of Benazir Bhutto exacerbated this concern.

As a condition for financing a substantial portion of the program's costs, the donor agencies required the government of Pakistan to set up a completely new institutional mechanism to formalize the selection of recipients for BISP. This new mechanism would include the collection of a short questionnaire on key indicators of poverty ("Poverty Scorecard") to all potentially eligible households nationwide, the collection and management of the data in a new government agency in Islamabad, and the use of the data to calculate a proxy means test as the sole determinant of eligibility for the transfer. The World Bank played a lead role in designing this system (cf. Hou (2009)). However, despite the donors' leverage, it is not obvious that they would be able to impose this system by fiat. Overall, 90% of the funds for BISP have come from domestic sources (Asian Development Bank, 2013); and historically donors have had difficulty enforcing aid conditionality. In this case, the government agreed to the new system, but initiated the program using an alternative system. For the first two years of the program, national politicians, including Members of the National Assembly (MNAs), were officially responsible for nominating recipients. MNAs are directly elected, each representing a constituency of approximately 300,000 registered voters. Each MNA was given 8,000 nomination forms to sign up beneficiaries from their constituencies. They were given a few criteria based on what could be verified in the national ID database: the recipients should not have a machine readable passport, an ID card for emigrants (NICOP); an account with a foreign-owned bank; or have any household member who is a government employee. However, given that the vast majority of Pakistani households would qualify under these guidelines, this effectively gave the politicians a great deal of discretion to select recipients. This level of discretion was similar to previous social support programs in Pakistan; for example, recipients for the *zakat* program were only required to be "needy" and were selected at the discretion of local committees (Clark, 2001). In fact, 50% of the politicians' original nominees were disqualified based on these minimal criteria and never received BISP even before the reform. Nayab and Farooq (2012) find in an independent household survey that a substantial proportion of those who did receive BISP in this period also reported characteristics that would make them ineligible under the minimal criteria. Independent researchers as well as the donor agencies expressed concern that politicians were

[6] Nabi (2013); Leary *et al.* (2011); Cheema *et al.* (2014, 2015); World Bank (2013) provide a more in-depth discussion of the BISP program and its institutional features.

[7] Keefer and Vlaicu (2007); Keefer (2007), for example, have argued that high levels of expenditure on targeted transfers are in fact symptomatic of clientelism.

targeting their connections for the transfer (Gazdar, 2011; Khan and Qutub, 2010; World Bank, 2013).

One year after the inception of the program, the government began the rollout of the field-based data collection required as a part of the institutional reform agreed with the donors. The national census organization collected data for part of the country, and the government contracted out the rest to two major quasi-government NGOS, the Pakistan Poverty Alleviation Fund andd the Rural Support Program Network, and for a small sample in the Federally Administered Tribal Areas, to a private firm. These bodies fielded the survey and passed the hard copy questionnaires to a central office in the capital for data entry. The indicators collected in the survey were: the number of household members under 18 or over 65; the household head's education level; number of children currently attending school; number of rooms in the household's dwelling; the type of toilet used; ownership of land, livestock, and durable assets including a refrigerator, freezer, washing machine, air conditioner, heater, stove, television, microwave, car, tractor, or motorcycle. The field exercise included an audit process to spot check the results. In order to address the possibility of missed households, the organizations carrying out the survey carried a "mop-up" exercise to identify households that had been missed. Local influentials were contacted to help identify these households.

The agencies collecting the data submitted the forms directly to the agency responsible for national ID cards, NADRA. This agency calculated a "poverty score" for each household, a weighted sum of the indicators collected. The aid donors developed and proposed this methodology and a set of indicators based on best statistical prediction of household consumption. The exact weights used to calculate the score were kept secret in an effort to avoid gaming; BISP officials report that the exact weights were kept a secret even from the BISP agency's head. Households with a score below 16.17 were deemed eligible for the BISP cash transfer, while those above this score were ineligible.[8]

The government administration then initiated transfers to eligible recipients via the post office. Those who had received BISP received a letter informing them that they were no longer eligible to receive the transfer.

The entire process was initiated in 2009 in 15 pilot districts (out of 106 total in the country), including four districts in the province we study, Punjab. These districts were selected by BISP agency officials in order to cover a range of geographic areas, more urban and more rural areas, and areas with different poverty levels; MNAs and other politicians were not involved in the district selection process. Table 3 shows a comparison of observable characteristics of the pilot districts and the rest of the province from the year before the BISP cash transfer was introduced in 2008. Panel A shows household characteristics from the 2007-8 MICS survey, while Panel B shows constituency-level characteristics from Election Commission of Pakistan data. The pilot districts are very similar

[8]This cutoff was selected to achieve an approximate national number of recipients based on the program's budget. Cheema *et al.* (2014, 2015) use this cutoff to estimate a Regression Discontinuity Design and find a positive impact of the program on some outcome measures, including consumption, heath expenditure, and women's control over finances.

to the rest of the province on political variables. Households in these districts are similar on most observables to those in other districts. They are more likely to live in completely rudimentary houses, and less likely to be female-headed. One of these variables is positively correlated with poverty and BISP eligibility, while the other is negatively correlated; there is no comprehensive pattern of households in these pilot areas being more or less likely to be eligible.

The scorecard data collection was completed in these pilot districts in June 2010; from July, BISP transfers were initiated to the newly selected recipients and discontinued to ineligible recipients. In the remaining districts, the scorecard data collection was completed in June 2011 and transfers were initiated from July onwards.

Recipients selected under the old system, who did not qualify under the new criteria, were supposed to be removed from the list and their payments stopped; this amounted to 75% of the previous recipients. BISP officials report some pushback from politicians on this, but they resolved these concerns in a series of briefings with the politicians, emphasizing that the program was being scaled up dramatically to many more constituents. Without this scale-up, the reform might have been blocked.

It is possible that politicians would have anticipated the change in targeting mechanisms and taken this into account in their nomination decisions. This could have caused them to choose more individuals they preferred as recipients but anticipated would be disenrolled later, in order to ensure that these individuals received at least some funds. However, it seems unlikely that the politicians believed that these kinds of rules would be consistently applied, given that 50% of even their initial nominations were rejected based on the minimal criteria applied at that stage.

As of 2016, the government and donors are preparing for another field survey to update the poverty scorecard data, allowing for new households to be systematically enrolled and others which no longer meet the criteria to be "graduated". To date, the only other way for a recipient to be disenrolled from the program is through the system's "grievance redressal mechanism," which functions through a national hotline as well as case management offices at the level of the *tehsil*. Any party can report a case of a household who may have received the transfer inappropriately, and these offices are responsible for a household visit to re-verify assets. In practice, however, reports of this kind are nonexistent; reports to the hotline and the tehsil offices consist of households who want to be included, others calling on their behalf, or problems with the payment mechanism. Therefore, when we observe households in our sample who stop receiving BISP, this can be attributed to changes in targeting, rather than the program's effectiveness in lifting them out of poverty.

After the change in targeting system, the government also implemented a change in the delivery mechanism for funds. At the inception of the program, cash was sent to recipients through the post office. In part because of complaints of corruption by postal workers, the government introduced a "smart card" to be issued to beneficiaries,

which could be used like an ATM card to take out cash from bank branches. This was piloted in a different set of pilot districts in 2012 and introduced to the rest of the country in 2013. These reforms appear to have reduced petty corruption (Cheema *et al.*, 2014, 2015); there have also been some issues with the new system, notably difficulty for female recipients in conservative parts of the country to travel to the bank to retrieve their cash. This reform did not overlap with our period and locations of study in our politician village data, but it could be one of the mechanisms for the change we see in the MICS data.

During this time period, the intended target group, unconditional nature of the transfer, and the amount of benefits provided did not change. The agencies managing the program, the political environment and politicians in office were all the same before and after the reform. This allows us to compare the outcomes over that time period and isolate the effect of the reform. However, the program was scaled up to many more beneficiaries. In our main sample, the proportion of households receiving BISP is more than three times as high after the reform. This could lead to concerns that changes we observe before and after the reform are due to the scale up. In addition to controlling for the overall increase in BISP after the reform, we also construct all our tests so that scale-up would if anything bias us away from a finding of interest, rather than driving our findings. For example, we test for a reduction in targeting inclusion errors (transfers to the wealthy), which would be mechanically increased by the scale-up; we discuss this further in Section 4. We also show that saturation of politicians' preferred recipients does not drive our effects of interest.

Even with a formal beneficiary selection process and the other controls introduced as a part of the reform, there could still be scope for continued favoritism in the targeting of BISP. Politicians could influence the agencies carrying out the poverty score card survey to manipulate their data (cf. Litschig (2012); Niehaus *et al.* (2013)), or pressure bureaucrats to interfere with the formula or data (Banful, 2011; Camacho and Conover, 2011). There is also scope for influence on distribution after initial beneficiary selection. Alatas *et al.* (2013a), for example, find that elite influence on assistance targeting in Indonesia occurs during distribution rather than beneficiary selection. Khan and Qutub (2010) report that before the reform, in some cases politicians and influential people collected the money intended for the beneficiaries, then redistributed it to their preferred beneficiaries. This might still have taken place after the reform to beneficiary selection, particularly before the shift to the use of biometric smart cards for withdrawing the cash. Another possibility is that politicians could assist those in their villages to overcome administrative hurdles in getting the transfer. Households are not required to make any application for BISP, but a female household member must be named as the recipient of the transfer, and must get a National Identity Card issued if she does not already have one. Politicians could assist potential recipients in the process of getting an ID card, or could help them resolve problems with the delivery of the stipend.

For those households whose poverty score was above the cutoff of 16.17 but below 20, this letter informed them

that they could file an appeal, by sending a letter, calling the BISP hotline, or visiting the *tehsil* level offices. In practice, households that filed an appeal were automatically re-enrolled as long as they met a second, less stringent, set of criteria. The authorities reviewed the data again and re-enrolled an appealing household if they were between 16.17 and 20 points on the scorecard, and either (a) had a household size less than 3; (b) had 4 or more children under 12 years of age, (c) had any member above 65 years, or (d) had any member with a disability. As of 2014, approximately 600,000 households had filed appeals and approximately 35% of these were then enrolled in BISP.

3 Data

We use the delayed introduction of the BISP reforms to study their effects on transfer targeting, favoritism, and perceptions of fairness using two household survey datasets. Both surveys were conducted in Punjab, the most populous province of Pakistan, with 80 million of the country's 180 million people.

The first dataset is a unique household survey of politicians' origin villages, used to identify favoritism for the village of the winning politician. We collected this survey data in 2013 in rural Punjab, in collaboration with the Lahore School of Economics.

We took a random sample of constituencies that either had a close outcome in 2008 (5% vote margin), and/or a two-way switch, i.e. the top two candidates exchanged places between 2002 and 2008. We eliminated constituencies in which politicians of the same clan won in both 2002 and 2008, in order to identify the effect of winning on the politician's clan. Multiple candidates contest every MNA seat, but because they are elected on a first-past-the-post basis, there are usually two or at most three major candidates. For each constituency, we identified the 2008 MNA winner and runner-up, and the 2002 winner and runner-up. We designate all these alternative politicians as "rivals."[9] In many cases, one or both of the 2002 top finishers ran again in both 2002 and 2008, so each constituency has between one and three rival politicians.

We used public source information and informed contacts to identify the origin villages where these politicians were either born or have ancestral land; thus politicians cannot select their origin villages. The village population comprises less than 1% of the whole constituency.

We were able to identify and survey an origin village for 36 out of a total of 53 candidates in all 19 constituencies. We surveyed just under 8,000 households across these 36 villages. Cases in which we did not identify an origin village for a candidate could occur because the candidate had no rural base (i.e. came from a family that has been based in an urban area for the last few generations). It could also occur if there was such a village but we were

[9]While it is possible that one of the 2002 candidates could be an ally with the 2008 winner, it is difficult to establish this due to party switching and changing alliances within parties and clans; members of the same party and/or clan sometimes contest against each other for the same seat. These politicians' villages are all well-connected communities, making them a good comparison group to the 2008 winner.

unable to find the information through our sources. This could bias our results if it was driven by less prominent politicians, who might be less likely to attract resources to their village. This would bias the estimates of favoritism, but should not affect the change in favoritism before and after the BISP reform. The candidates for whom we did not identify a village were roughly split between the 2008 winners and rival politicians (10 winners and 7 rivals), which does not suggest a pattern of systematic under-representation of less powerful politicians. As a robustness check, we present constituency fixed-effects estimations in all the main results tables.

Of the 36 total villages, 19 villages correspond to the 2008 winner and runner-up in 11 constituencies that had a close vote margin in 2008, and thus can be used to estimate a close-elections regression discontinuity (Lee, 2008); we include results with this subset as a robustness check in all tables. For a subset of constituencies, we also identified villages neighboring the winners' and rivals' villages, i.e. those in the same *patwari* circle, an administrative group of about 3-5 villages. We surveyed approximately 5,000 households in 17 such nearby villages.

The survey fieldwork took place in February 2013, two months before Pakistan's 2013 election; this allows us to measure cash transfers to households throughout the 2008-2013 term of the National Assembly. The information on cash transfers received before and after the reform was collected through a retrospective question. Because this question is retrospective, recall error and recall biases such as telescoping (Neter and Waksberg, 1964) could be a concern. However, BISP is a well-known program, receiving it is a salient event; in pilot surveys we found that respondents readily recalled it. We also took two measures to help to prevent and address recall bias. Each respondent was first asked whether his/her household has ever received BISP. If the answer was "yes," the enumerator used a series of time reference points to identify the years in which they received the program (Loftus and Marburger, 1983). In addition, we define the pre- and post-reform BISP variable based on a restricted set of years. A household is considered to have received BISP pre-reform if they reported receiving it in 2008 and/or 2009. A household is considered to have received BISP post-reform if they reported receiving it in 2012 and/or 2013. We present further evidence that recall bias does not drive our results in Section 5.6.2.

We identify households who are of the same clan as the MNA or his/her rivals by matching the clan of the these politicians as reported by local village officials with the household's self-reported clan ("zaat" or "biradari"). The question to households was posed before any questions about government assistance or politicians to avoid biasing the responses. Table A2 show descriptives for politicians' clans and other households; politicians' clans tend to be wealthier. As a robustness check, we test whether households in a politician's village are more likely to report the same clan as the politician if he won, and find no evidence of this.

The survey also included data on the household's key assets at the time of the survey (2013) and in 2007, including agricultural and residential land, and whether the household lived in a *kacha* (rudimentary) or *pakka* (solid) dwelling - universally understood local categories for the type of building material used. It also covered receipt of other kinds

of government benefits and assistance, including all the government social protection programs, and a short module at the end covering connections to the MNA and other key officials. The full questionnaire is available as a web appendix.

We identify the party of the winner and rivals based on their party allegience at the time they ran for office, as listed in Election Commission documents; we divide these politicians into the majority party in the national government (People's Party of Pakistan) and all others. We also used public documents to identify which of the politicians serve as ministers, advisors to the Prime Minister, or parliamentary secretaries, and identify these as "senior" politicians.

The second dataset is the Punjab Multiple Indicator Cluster Survey, or MICS. This survey is conducted by the provincial government in collaboration with UNICEF. We use four rounds of the MICS survey data, which are repeat cross sections from 2003, 2007, 2011, and 2014. Each round covers a representative sample of the population of the province.

After the introduction of the BISP, from 2011 onwards, respondents are asked about whether the household received BISP within the last year. At the time of the 2011 survey, households in the pilot districts would have observed BISP transfers under the new system for over a year, and the households in those districts who report receiving BISP would have received it under the new system. Households in other districts would still have received it under the old system. By the time of the 2014 survey wave, BISP had been operating under the new system for several years in all districts, so we consider all districts as "treated" with the reform at this point.

All rounds of the MICS surveys include questions on other government cash and in-kind transfer programs, as well as detailed measures of household assets and income, and a wealth index.[10] We also use the MICS data to calculate a proxy for official eligibility for the BISP transfer post reform, based on the scorecard.[11]

The MICS surveys from 2007 onwards also include a yes/no question on the respondent's perceptions, asked immediately after the respondent reports on the BISP and other targeted benefits: "Do you agree that the government schemes are beneficial to the common man?" Because of its placement in the questionnaire, this question captures the perception of government targeted benefits programs, rather than the government as a whole. We use this to capture the effect of the BISP reform on the perceived legitimacy of government programs.

Table 1 shows summary statistics for the variables used in our analysis from both datasets. Table A1 shows

[10] These variables include characteristics of the house, durable assets such as TV, and refrigerator; animals, vehicles, and access to public services such as electricity and drinking water. The index was calculated using principal components analysis, based on Rutstein and Johnson (2004) and Filmer and Pritchett (2001). It is described in more detail in Punjab Bureau of Statistics (2008). This index has some overlap with the BISP poverty score, but includes a more detailed list of assets and different weights.

[11] We do not have access to the exact formula used by the government, because the indicator weights used are closely guarded to help discourage possible manipulation. However, the partners who worked with the government on data collection shared an approximation of the scorecard formula with us. We map this to the MICS variables. All of the variables from the scorecard have an exact equivalent in the MICS survey except for two variables: the number of buffaloes (the MICS survey does not distinguish buffaloes and cows) and the number of rooms in the household (the MICS survey includes only bedrooms, while the scorecard includes living spaces). Because assets may have changed between the scorecard survey and the MICS survey, in part due to the BISP transfer itself, we also compare results using this variable to analysis using major assets such as land, that would not be affected by the transfer.

the household-level correlates of receiving BISP at any time from the program's inception until the survey date in early 2013. Column 1 shows our basic set of covariates, specified in the pre-analysis plan because they were used for identifying BISP recipients; column 2 shows a full set of additional covariates.

There is a growing awareness of the potential for over-rejection of null hypotheses in empirical work due to specification search (e.g. Brodeur *et al.* (2016); Gerber and Malhotra (2008)). Some have proposed registration of detailed pre-analysis plans (e.g. Casey *et al.* (2012)), but others have raised concerns that this limits researchers' ability to learn from the data (e.g. Deaton (2012); Humphreys *et al.* (2013); Olken (2015)). To address this, we use a novel combination of approaches in the analysis of our primary data (the politician village sample).

We registered an initial pre-analysis plan with the Experiments in Governance and Politics registry. We indicated that our pre-analysis plan was non-binding, following Humphreys *et al.* (2013). The original pre-analysis plan is available on the EGAP registry website. We made changes in our specification from the pre-analysis plan in order to incorporate feedback and improve the analysis. A description of the changes and results of the originally proposed specification are presented in Appendix B; the results are consistent with the results we present here.

Second, we adapted the "sample split" method proposed by Fafchamps and Labonne (2016b), reserving half of the data for out-of-sample testing. A third party researcher used randomization code we wrote to partition the data, then released only the selected half to us as the "training" sample, and archived the other half (the "testing" sample). We conducted all analysis on the training subsample. We present the results from the "training" sample in the paper. Effectively, all our results from the politician village sample presented here form a second pre-analysis plan, which can also be registered in a public registry. After that point, we will send our code to the third-party researcher to replicate our estimations (a) on the pooled sample, for increased precision, and (b) on the "testing" sample only, as a robustness check. Any spurious results would not be replicated in the reserved portion. The results of this robustness test will be reported in the final paper. This combination approach allows more flexibility in departures from our pre-analysis plan. Thus we are working with half the data in all the analysis shown here; our main working sample is a total of just under 4,000 households in all 36 villages.

4 Empirical Strategy

4.1 Favoritism

We use the primary household survey data from politician origin villages to examine the effect of the reform on politician favoritism. We compare distribution of the BISP cash transfer in winner and rival politician villages before and after the reform. In this dataset we do not use pilot districts as a source of identification because of limited overlap between the sample and the pilot districts, and because doing so would rely heavily on the precise

Table 1: Means of key variables

	Politician village sample		Representative sample (MICS)	
	Full sample	RD subset	Full sample	Rural
Household level				
Received BISP: pre reform[a]	0.017	0.016	0.019	0.021
Received BISP: post reform[a]	0.048	0.051	0.026	0.033
Received BISP: ever	0.072	0.080	-	-
Received any pre-08 needs-based transfer[b]	0.001	0.002	0.038	0.035
2008 winner's clan — winner's village	0.169	0.181	-	-
2008 winner's clan — rival's village	0.044	0.066	-	-
2008 rival's clan — rival's village	0.135	0.106	-	-
2008 rival's clan — winner's village	0.073	0.088	-	-
Related to elected or local official	0.039	0.032	-	-
Female head	0.029	0.026	0.074	0.071
Any daughters currently aged 18-25	0.392	0.394	0.202	0.187
Rudimentary house[c]	0.327	0.352	0.107	0.161
No agricultural land	0.785	0.791	0.682	0.557
Less than 12.5 acres agricultural land	0.979	0.974	0.967	0.959
No cattle[c]	0.713	0.740	0.628	0.453
No residential land	0.170	0.178	0.139	0.103
HH head 5th grade or higher	0.447	0.446	0.589	0.514
HH head 8th grade or higher	0.321	0.306	0.478	0.402
HH head 10th grade or higher	0.215	0.210	0.340	0.264
Years HH has lived in village	77.672	74.918	-	-
HH member received ID (02-07)	0.521	0.498	-	-
Constituency level				
Winning party: PPP (majority party)	0.263	0.364	0.306	-
Winning party: PML-N	0.579	0.455	0.426	-
Winning party: Independent	0.158	0.182	0.08	-
"Senior" politician	0.474	0.556	-	-
Vote margin	0.064	0.022	0.269	-
Observations				
Households	3693	2267	255483	159070
Villages	36	19	5545	3063
Constituencies	19	11	150	-

Notes: Politician village sample values reported in 2013; MICS sample pooled cross-sections 2003-4, 2007-8, 2011, and 2014. a: MICS sample values are from 2011 and 2014 rounds only. b: MICS sample - 2003 round only. c: Politician village sample - recall of 2007 value.

recall of years in which the household received the transfer.

To quantify favoritism, we need a causal estimate of the effect of a politician's election on assistance to his village. However, there are many reasons that households in politicians' villages may differ from other households. Politicians in Pakistan tend to come from more socially and economically privileged groups, so households connected to them are likely to be privileged ex ante. This could bias our estimate upwards, because these households know how to "work the system" and access government assistance effectively. Alternatively, our estimate could be biased downwards, because these households are less likely to need certain programs, or choose not to apply for them to avoid stigma (Gille, 2013).

Second, it is possible that villages are chosen strategically. This may happen in one of two ways. First, politicians may choose their residence strategically. This could lead to simultaneity bias. For example, they could move to be in areas where households are well-served by the government, to maximize access to receptive voters. This could bias our estimate upwards. However, note that all contesting candidates could potentially move villages strategically. Thus comparing winners and rivals would address this bias, unless politicians move *after* the election.

To address these concerns, we compare a winning politician's origin village with the villages of his rivals. As described in section 3, we identified and sampled the origin villages of winning politicians in the 2008 election, and their rivals, i.e. the 2008 runner-up and the 2002 winner and runner-up. A politician's origin village is defined as that in which he was born or has ancestral land; politicians cannot select their origin villages.

As a robustness check, we estimate the same specifications for the subsample of winners and runners-up from the 2008 elections with a 5% margin, i.e. a close-election Regression Discontinuity Design (Lee, 2008).[12]

To estimate the effect of the BISP reform on favoritism, we incorporate this approach into a difference-in-differences estimate. We aggregate the data into two periods, pre and post reform, and estimate the following linear probability model:

$$BISP_{ict} = \alpha + \beta_1 WIN_i PRE_t + \beta_2 WIN_i POST_t + \gamma POST_t + \delta X_i + \eta Z_c + u_i + \epsilon_{ict} \quad (1)$$

$BISP_{ict}$ is a dummy defined as 1 if household i in constituency c received the BISP cash transfer at any time during time period t. WIN is a dummy variable for the winner's origin village. Recall that the sample is composed of households in the origin villages of the 2008 winner and his rivals. X is a vector of household observables; Z is a

[12]Recent work on U.S. elections, such as Sekhon and Caughey (2011); Grimmer *et al.* (2011) shows that just-winners are more likely to be bigger campaign spenders and to control key state administrative bodies, which calls into question the identifying assumption that winners and rivals are similar. Note that this result does not necessarily imply actual rigging of the election, nor does rigging necessarily create this kind of correlation (cf. Simpser (2013); Gehlbach and Simpser (2015)). For either campaigning or rigging to cause problems for our empirical strategy, politicians in power would have to closely monitor the projected voting outcomes and allocate resources to achieve a vote count just above that of his competitors and no more. This seems unlikely given the information available that this could take place to the level of precision required. A recent systematic study of elections from a range of other developed and developing countries by Eggers *et al.* (2015) shows no such pattern in any other context studied apart from the United States.

vector of constituency-level observables, including the 2008 vote margin. PRE and $POST$ are dummies for before and after the BISP reform, respectively. u_i is a household-level idiosyncratic term; we estimate a household-level random effects model to increase efficiency. β_1 and β_2 are the effect of a winning politician on BISP for households in his village, before and after the reform respectively. We will test for a decrease in favoritism, $H_0 : \beta_1 = \beta_2$.

The identifying assumption is that households in origin villages of winners and rivals do not differ in any unobserved way that causes them to have different *trends* in receiving BISP before and after the reform. We test whether households in our sample are similar on observables between the origin villages of winners and rivals. Table 2 shows the results. The two groups are generally similar, although households in the winners' villages are slightly more likely to own land (in our main estimation sample, this is significant at the 10% level in the main sample, and is the only significant result of 15 variables tested). If anything, this suggests that they are slightly *less* likely to be eligible for the cash transfer. We include these observable characteristics as controls and test robustness to interacting the household controls with a post-reform term.

The entire period we study occurred during one electoral term, so the same politicians were in office throughout. In addition, most aspects of the BISP cash transfer program besides the reform stayed constant. The main change over this period was the scale up of the program to more beneficiaries. This means that β_1 and β_2 represent marginal effects compared to different base levels. For example, an estimate of $\beta = .04$ would represent a much greater distortion in targeting if the sample mean is .02 than if it is 0.1. Since the mean increases in the rivals' villages, our tests for a decrease in the level of favoritism represent a conservative approach. The scale-up of BISP could also have affected the composition of recipients, for example if polticians' preferred set of recipients is covered and they do not influence the selection of additional recipients. In Section 5.6.1 we present detailed evidence that this sort of saturation effect does not drive our results.

Needy households could migrate to be closer to a winning politician, to gain access to services. To address this concern, we collect migration history for each household, and drop all 342 households who have moved to the village at any time since the 2002 election from all analysis with the politician village sample. Because our data is retrospective, we cannot identify households that might have left the village during this period. However, it is unlikely that differential out-migration could drive our results. For differential out-migration to drive our difference-in-differences results, households who are more likely to receive BISP when the winner selects recipients *and* then be disqualified from BISP after the reform would have to be more likely to leave rival politicians' villages. There is no straightforward explanation for the latter to occur.

We cluster standard errors at the village level. As an additional robustness check, in all tables based on the politician village sample, we present bootstrapped standard errors clustered at the village level and wild-cluster

bootstrapped standard errors clustered by the constituency, as proposed by Cameron *et al.* (2008).[13]

To test for differences in favoritism by clan and other groups, we estimate a version of (1) interacted with indicator variables for each group.

4.2 Targeting

It is possible that when given discretion to choose recipients, politicians use their village- and clan-based networks to identify poor recipients. Thus the BISP reform might re-allocate funds to households that are not connected to politicians, but not necessarily poorer than the previous recipients. We use the reform to test for this.

Targeted transfers can be subject to both inclusion errors (transfers received by wealthier households) and exclusion errors (poorer households who do not receive transfers). The scale-up of the BISP program which occurred along with the reform could mechanically reduce exclusion errors. However, the scale-up would tend to *increase* inclusion errors. Therefore, we focus on testing the effect of the reform on inclusion errors. If the reform decreased these errors, the reform must have improved targeting.

We assess inclusion errors based on both indicators that were and were not a part of the poverty scorecard. Indicators that formed a part of the scorecard can be used to assess whether the new targeting rules were followed in practice, and whether they were binding. (If all pre-reform BISP recipients were poor enough to be eligible by the new criteria, we would see no effect on targeting based on these indicators.)

Indicators not included in the scorecard can be used to assess whether the proxy means test effectively improved targeting more broadly. Before the reform, politicians could have used local knowledge to select recipients who were poor, but who might not appear to be poor based on the limited set of scorecard indicators. If this was the case, we might see a reduction in inclusion error based on the scorecard indicators but no similar effect in the other indicators.

We include indicators that are directly observed by the survey enumerator (the physical state of the walls, roof and floor of the respondent's dwelling) and so not subject to respondent misreporting.[14]

Simultaneity bias could also arise in these estimates, since households could use the cash to buy assets. To check for this, we test wealth indicators that are pre-determined, such as education, or costly, long-lasting assets such as land. In the politician village sample, we use lagged values of the house quality and cattle variables. Table A4 shows the full list of wealth proxies we use in each of the two data sets, and their characteristics: whether they are part of the scorecard, directly observed by enumerators, and unaffected by the BISP.

[13]For brevity, we report p-values for post-estimation tests based on the bootstrap method only as a robustness check when the standard test is significant.

[14]Since neither the MICS survey nor our politician village survey was connected in any way with the BISP, there would also be no incentive for enumerators to misreport wealth indicators, e.g. to assist respondents in qualifying for the transfer.

To test how the reform changed targeting in our politician village sample and how favoritism affected targeting, we estimate a version of (1) interacted with each of the key wealth proxies available in the data.

To test how the reform affected targeting in the province as a whole, we use the MICS repeated cross-section data and exploit the rollout of the BISP reform district by district. We define a dummy for "post reform" that is equal to one if the reform had been implemented in the district-year. All specifications also include district and survey wave fixed effects. Thus the identifying assumption is that trends in BISP allocation do not systematically differ between pilot and other districts for reasons other than the reform. We estimate:

$$BISP_{idt} = \beta_0 + \beta_1 WLTH_{idt} + \beta_2 POST_{dt} + \beta_3 WLTH_{idt} POST_{idt} + \alpha_d + \mu_t + \epsilon_{idt} \tag{2}$$

Where $BISP$ is a dummy for household i receiving BISP, and $WLTH$ is a binary wealth proxy. $POST$ is a dummy which takes value 1 if the reform has been implemented in district d at time t; note that it takes on different values for different district in the same year, thus capturing the district-by-district phase-in. α and μ are district and survey wave fixed effects. Our coefficient of interest is β_3: whether inclusion error decreased, i.e. wealthy households were less likely to receive BISP after the reform.

4.3 Legitimacy

To identify effects on the perceived legitimacy of government social safety net programs, we again use the phase-in of the reform for identification. We estimate:

$$\begin{aligned} APPROVE_{it} =& \beta_1 + \beta_2 POST_{dt} + \beta_3 BISP_{it} + \beta_4 T_{it} + \delta \widetilde{WLTH}_i + \beta_5 BISP_{it} WLTH_{H,i} + \beta_6 T_{it} WLTH_{H,i} \\ &+ \beta_7 BISP_{-i,dt} + \alpha_d + u_t + \epsilon_{idt} \end{aligned} \tag{3}$$

Where $APPROVE$ is a dummy taking value 1 if the respondent agrees with the statement "government schemes benefit the common man," $\widetilde{WLTH}_i$ is a vector of wealth quintile dummies, with $WLTH_H$ designating the richest two quintiles; $BISP_{it}$ and T_{it} are dummy variables for HH i receiving BISP and other government cash transfers, α_d district fixed effects, and u_t round fixed effects. $POST_{dt}$ is again a dummy which takes value 1 if the reform has been implemented in district d at time t. Our coefficient of interest is β_2: the effect of the reform on public perceptions of government programs.

The identifying assumption is that there is no differential trend in approval between pilot districts and others that is not driven by the reform itself. Recall that these districts are fairly representative of the province as a whole on political as well as household characteristics (Table 3). The scale-up of the program to more recipients

could change approval in these districts; respondents would likely express more approval for government programs simply because they see them in greater quantity. To address this, we control for the prevalence of BISP in the district at time t, $BISP_{-i,dt}$, defined as the proportion of households in district d who received BISP in year t, excluding household i. We also estimate a version of (3) controlling more flexibly for the scale-up by incorporating fixed effects for each decile of $BISP_{-i,dt}$.

5 Results

5.1 BISP reform and politician favoritism

Table 4 shows the results of Equation 1, comparing favoritism for winning politician's villages before and after the BISP reform. At the inception of the program, politicians had the official responsibility for selecting households to receive the transfer. The villages of winning politicians were substantially more likely to receive BISP before the reform. Given the politicians' role, this is not surprising. The magnitude of this effect is dramatic, however. Before the reform, households in the winner's village are 1.5 to 2 percentage points more likely to receive the BISP transfer. Comparing them to the "control group" mean - the proportion of recipients in the rival politicians' villages - these households are 200 - 400% more likely to receive BISP.

With the reform, the transfer was scaled up to more recipients. The "control group" mean reflects this increase. In the full sample, the winner's village term is similar in magnitude to the pre-reform term, and is significant in the specifications with covariates and constituency fixed effects. This pattern does not persist in the RDD subset of close winners and runners-up. However, we cannot reject the null that winners' villages receive similar levels of favoritism before and after the reform in any of the specifications.

Overall, there is tentative evidence that the winners' villages may have continued to have some advantage in receiving BISP after the reform. We discuss possible mechanisms for this in Section 5.7. However, this is not proportional to the massive increase in BISP. Before the reform, households in the winner's villages are two to four times more likely to receive BISP. After the reform, the point estimates suggest they are 50% more likely to receive the transfer than those in other politicians' villages.

We now break down the estimates by clan. The winner's clan is a substantial minority (17%) of his village who have a closer connection to him. We divide households into two groups: those who are members of the winning politician's clan and all others, and interact this variable with our main specification 1. Table 5 shows the results. Comparing the first and third terms, we see that before the reform, all clans benefited from having a politician from their village elected. The point estimates of favoritism for the politician's clan are larger than those for other clans before the reform, although the difference is not robustly significant. After the reform, favoritism *decreased*

Table 2: Balance between winner and rival villages

	(1)	(2)	(3)	(4)
	Rival's village	Winner's village	Difference	SE difference
Full sample				
Same clan as origin-village politician	0.169	0.113	0.056	(0.073)
Related to elected or local official	0.037	0.031	0.006	(0.019)
Female head	0.029	0.026	0.003	(0.008)
Any daughters currently aged 18-25	0.376	0.414	-0.039	(0.039)
Rudimentary house (lag - 2007)	0.325	0.330	-0.005	(0.093)
No agricultural land	0.750	0.836	-0.086*	(0.044)
Less than 12.5 acres agricultural land	0.975	0.982	-0.008	(0.009)
No cattle (lag - 2007)	0.705	0.714	-0.009	(0.053)
No residential land	0.151	0.208	-0.057	(0.093)
HH head 5th grade or higher	0.448	0.444	0.003	(0.056)
HH head 8th grade or higher	0.312	0.322	-0.010	(0.048)
HH head 10th grade or higher	0.206	0.221	-0.016	(0.036)
Years HH has lived in village	78.170	77.355	0.815	(5.361)
Years squared	6650.941	6518.008	132.933	(692.200)
HH member received ID (02-07)	0.492	0.535	-0.043	(0.061)
RD subset				
Same clan as home-village politician	0.181	0.106	0.075	(0.088)
Related to elected or local official	0.042	0.010	0.032	(0.020)
Female head	0.029	0.020	0.009	(0.009)
Any daughters currently aged 18-25	0.369	0.450	-0.081*	(0.044)
Rudimentary house (lag - 2007)	0.336	0.389	-0.053	(0.148)
No agricultural land	0.749	0.885	-0.137**	(0.050)
Less than 12.5 acres agricultural land	0.971	0.982	-0.010	(0.013)
No cattle (lag - 2007)	0.713	0.799	-0.086	(0.067)
No residential land	0.118	0.310	-0.192	(0.141)
HH head 5th grade or higher	0.435	0.471	-0.036	(0.079)
HH head 8th grade or higher	0.299	0.321	-0.022	(0.062)
HH head 10th grade or higher	0.208	0.215	-0.007	(0.052)
Years HH has lived in village	76.033	72.457	3.576	(8.137)
Years squared	6386.510	5858.895	527.615	(1068.079)
HH member received ID (02-07)	0.508	0.475	0.032	(0.065)

Notes: Politician village sample. Where a characteristic could easily change because of election outcomes over recent years (cattle, rudimentary house), we collected recall data from 2007. The other variables (land, head's education, female head, number of adolescent girls) are unlikely to be affected by the election outcome. Standard errors are clustered at the village level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Characteristics of reform pilot districts

	Pilot	Other districts Mean	Difference	SE difference
HH level (MICS)				
Female headed HH	0.04	0.05	-0.0191	(0.00799)**
No agricultural land	0.68	0.65	0.0238	(0.0727)
Less than 12.5 acres ag land	0.95	0.96	-0.0105	(0.0165)
No cattle	0.60	0.59	0.00829	(0.0874)
HH head 5th grade or higher	0.47	0.49	-0.0191	(0.0317)
HH head 8th grade or higher	0.47	0.49	-0.0191	(0.0317)
HH head 10th grade or higher	0.31	0.33	-0.0211	(0.0298)
Rudimentary dwelling (partial / full)	0.52	0.44	0.0738	(0.0706)
Rudimentary dwelling (full)	0.18	0.11	0.0714	(0.0330)**
N	7980	83095		
Constituency level (Election Commission)				
PPP vote share	0.26	0.30	-0.0363	(0.0804)
PML-N vote share	0.30	0.31	-0.0156	(0.0727)
PPP won	0.23	0.32	-0.0877	(0.156)
PML-N won	0.46	0.41	0.0467	(0.173)
N	13	135		

Notes: Household level data are MICS 2007-8 data. Constituency level data are 2008 election results, Election Commission of Pakistan. Standard errors are estimated with a regression of the variable listed on a dummy for pilot district, clustered at the district level. * $p < 0.1$, ** $p < .05$, *** $p < .01$.

significantly for the winner's clan in his own village. The overall level of BISP in this group actually decreased significantly after the reform, even as the scaleup led to an increase in BISP for households in the rivals' villages and for other clans in the winners' villages. (Households from the winning politician's clan were a substantial proportion of those who stopped receiving BISP after the reform, as shown in Table A14.) The reform successfully limited favoritism for the best-connected groups.

Table 4: Impact of BISP reform on favoritism for politicians' villages

	(1)	(2)	(3)	(4)	(5)
	HH received BISP cash transfer				
Winner's village, pre-reform	0.016	0.017	0.021	0.018	0.019
	(0.008)**	(0.006)***	(0.010)**	(0.009)*	(0.008)**
	[0.008]**	[0.007]***		[0.008]**	[0.009]**
	{0.041}**	{0.006}***		{0.099}*	{0.051}**
Winner's village, post-reform	0.015	0.025	0.021	-0.000	0.011
	(0.016)	(0.011)**	(0.012)*	(0.025)	(0.014)
	[0.017]	[0.011]**		[0.028]	[0.020]
	{0.253}	{0.053}*		{0.983}	{0.431}
Post reform	X	X	X	X	X
Vote margin		X			X
Vote margin x winner's village			X		
HH control variables		X	X		X
Additional HH control variables		X			X
HH control variables x post reform		X	X		X
Party dummies		X			X
Party dummies x post reform		X	X		X
Constituency dummies			X		
RDD subset				X	X
Rival village mean, pre-reform		0.008		0.004	
Rival village mean, post-reform		0.040		0.051	
N	7386	7386	7386	4534	4534
P-value: Winner's village, pre-reform = winner's village, post-reform					
Robust SE clustered by village	0.957	0.504	0.978	0.488	0.612
Bootstrapped SE clustered by village	0.957	0.527		0.520	0.687
Wild-cluster bootstrap, clustered by constituency	0.959	0.521		0.333	0.641

Notes: Politician village sample. (Parentheses: robust standard errors clustered at the village level.) [Brackets: bootstrapped standard errors clustered at the village level.] {Braces: P-value for cluster wild bootstrap test of $H_0 : \beta = 0$.} Basic and additional household controls are those shown in Column 1 and 2 of Table A1, respectively. * p < 0.1; ** p < .05; *** p < .01.

Table 5: BISP reform and favoritism by clan

	(1)	(2)	(3)	(4)	(5)
	HH received BISP cash transfer				
Winner's village, winner's clan - pre reform	0.047	0.043	0.031	0.046	0.035
	(0.023)**	(0.015)***	(0.028)	(0.026)*	(0.020)*
	[0.022]**	[0.019]**		[0.025]*	[0.026]
	{0.133}	{0.065*}		{0.258}	{0.255}
Winner's village, winner's clan - post reform	0.003	-0.006	-0.021	0.018	-0.004
	(0.017)	(0.019)	(0.026)	(0.013)	(0.020)
	[0.019]	[0.025]		[0.012]	[0.028]
	{0.890}	{0.806}		{0.347}	{0.846}
Winner's village, other clan - pre reform	0.011	0.015	0.016	0.012	0.018
	(0.006)*	(0.006)***	(0.009)*	(0.006)**	(0.008)**
	[0.006]*	[0.007]**		[0.005]**	[0.010]*
	{0.067}*	{0.009}***		{0.100}	{0.103}
Winner's village, other clan - post reform	0.022	0.028	0.027	0.003	0.012
	(0.019)	(0.012)**	(0.012)**	(0.028)	(0.016)
	[0.022]	[0.014]*		[0.025]	[0.016]
	{0.161}	{0.059}*		{0.860}	{0.469}
Winner's clan, post, winner's clan x post	X	X	X	X	X
Vote margin		X			X
Vote margin x winner's village			X		
HH control variables		X	X		X
Additional HH control variables		X			X
HH control variables x post reform		X	X		X
Additional HH control variables x post reform		X			X
Party FE, party FE x post reform		X	X		X
Constituency FE			X		
RDD subsample				X	X
Observations	7386	7386	7386	4534	4534
Mean dependent variable:					
Rival village, other clan, pre reform		0.008		0.005	
Rival village, other clan, post reform		0.041		0.055	
P-values:					
Winner's village winner's clan, pre = post					
Robust SE clustered by village	0.011**	0.041**	0.044**	0.037**	0.089*
Bootstrapped SE clustered by village	0.004***	0.045**		0.026**	0.211
Wild-cluster bootstrap, clustered by constituency	0.034**	0.125		0.222	0.321
Winner's village other clan, pre = post					
Robust SE clustered by village	0.532	0.333	0.376	0.746	0.740
Bootstrapped SE clustered by village	0.598	0.388		0.722	0.709
Wild-cluster bootstrap, clustered by constituency	0.389	0.321		0.600	0.734

Notes: Politician village sample. (Parentheses: robust standard errors clustered at the village level.) [Brackets: bootstrapped standard errors clustered at the village level.] {Braces: P-value for cluster wild bootstrap test of $H_0 : \beta = 0$.} Basic and additional household controls are those shown in Column 1 and 2 of Table A1, respectively. * p < 0.1; ** p < .05; *** p < .01.

5.2 Targeting the poor

The BISP reform clearly reduced favoritism towards the households best connected to politicians in power. But did it have any effect on whether the cash transfers reached the poor? If politicians select the neediest recipients about whom they have reliable information, this could produce a similar pattern of favoritism for the politician's connections. In this case, the reform might reduce favoritism without improving targeting. It is possible that politicians even have better information on poverty than that captured by the scorecard; in that case the reform could make targeting worse.

To address this, we first test how the reform affected targeting performance based on a range of wealth indicators. As outlined in Section 4, we focus on inclusion errors - BISP received by relatively wealthy households. A reduction in these errors cannot be a mechanical result of the scale-up of the BISP program.

Table 6 shows Equation 1 for the politician village sample, interacted with a vector of wealth proxy indicators. Before the reform, both poor and relatively wealthy households benefitted from having a politician from their village in office; the favoritism terms are not significantly different from each other. For two key assets, owning agricultural land and living in a solid house, inclusion errors dropped significantly after the reform in the politician village sample. While agricultural land was included in the new BISP scorecard, house material was not. The coefficient for wealthy alone is never significant and much closer to zero for all wealth proxies, suggesting that there was little targeting towards poorer households in this sample pre-reform. The result of improved targeting holds when we restrict the sample only to the villages of winning politicians (Table A7).

To assess the impact of the reform on targeting accuracy in the province as a whole, we use the representative MICS survey data; Table 7 shows the results. The estimates in Panel A use as wealth proxies our approximation of BISP eligibility and related indicators which are covered in the scorecard; Panel B includes only indicators which are *not* included in the scorecard. All estimates include district and round fixed effects. The results show a decrease in inclusion error both for indicators that were and were not targeted. This pattern holds for all but two of the 15 indicators, ownership of land and dwelling. By one measure, ownership of dwelling, inclusion errors appear to have increased. However, this measure is also correlated with more rural areas with limited property markets; this may simply reflect greater distribution of the cash transfer in more remote rural areas (Panel B, Column 1). For the other other wealth proxies, the coefficient wealthy x reform is two to ten times the magnitude of the coefficient on wealthy. Province-wide, the reform reduced inclusion errors substantially.

Receiving the BISP cash transfer could have affected the wealth indicators used, biasing the estimates of inclusion error upward. However, there is no reason that this bias should interact with the reform. In addition, this is only a potential concern for some of the variables. Education of the household head and spouse could not have been affected by the cash transfer (adult continuing education is non-existent in this context). The consistent results

with these indicators demonstrate that this is not driving our results.

It is also possible that the introduction of the scorecard could have caused BISP recipient households to under-report their assets on later surveys. However, the dwelling characteristics are directly observed by the survey enumerator rather than reported by the respondent; we find the same pattern of effects on these variables.

Fifty percent of the households nominated by politicians before the reform were disqualified based on not meeting even the minimal criteria in place at that time. The politicians approached the BISP agency to request additional nomination slots, but this request was declined. This suggests that politicians either knew very little about the status of these households, or that they expected the rules not to be applied to them. Nayab and Farooq (2012) also find that a substantial number of households who report that they *did* receive the transfer pre-reform did not qualify under all the criteria. Of the politicians' nominees who were accepted, 75% were then disqualified once the scorecard data collection was complete. Given the scale of this discrepancy, politicians could only be selecting the poorest households they have information about if either they observe a quantitatively important dimension poverty that was not measured by the scorecard (or the other measures we use); or if these politicians only know very wealthy households.

The first explanation is implausible given that we observe an improvement in targeting even when we measure inclusion errors based on wealth proxies that were not part of the targeting process. In addition, these indicators represent major assets for households in this context. Using survey data from a pre-BISP survey year (MICS 2003-4), we show that the characteristics on the poverty scorecard and the other characteristics on which we observe an improvement in targeting predict 40% of the household variation in income and consumption (Table A8). At the same time, many of the pre-reform recipients do have these same assets. It is implausible that there is a substantial unobserved characteristic of poor households that is both uncorrelated with these assets, observable to the politician and unobservable to us.

We also rule out the second explanation, because the winning politicians passed over a substantial number of households who *were* observably poor. Table A9 shows the characteristics of households in these villages who did not receive BISP before the reform; over 30% of the households in their villages who did not receive BISP before the reform lived in rudimentary houses, and three quarters do not own agricultural land. It is implausible that this pattern of results reflects politicians selecting households they know to be needy based on private information but who do not appear needy based on the scorecard. In addition, Table A9 shows that the majority of these households report that they know the winning politician through some kind of personal interaction, and these include many of the observably poor households. After the reform, there were new recipients selected even within the winners' own villages. Yet wealthier households were less likely to receive BISP even within these villages.

Thus the favoritism we observe cannot be attributed to politicians using reliable information through their social

network to choose the poorest households (or using a different definition of poverty). Rather, it reflects favoritism for either personal or electoral objectives (capture or clientelism).

5.3 Misreporting

One potential weakness of a survey-based targeting system such as the one implemented in the BISP reform is vulnerability to misreporting. It may be easier for a household to misreport its assets to a survey team than to a local individual (such as a politician or his local agents), who may have better information about the household.

The BISP survey teams were encouraged to conduct the survey inside the house whenever possible, but typically the interviews took place at the doorstep (due in part to gender issues, with female respondents reluctant to admit enumerators). Thus many of the household assets could potentially be misreported. Any party can report a case of a household who may have misreported their assets to a hotline or the offices of the program.[15] These offices are responsible for investigating such cases. In practice, however, reports of this kind are nonexistent; reports to the hotline and the tehsil offices consist entirely of households who want to be included, others calling on their behalf, or problems with the payment mechanism. This is not surprising since there is no direct benefit of reporting on one's neighbor. In addition, households found to have misreported are simply disenrolled from the program. Thus, there is no downside risk to misreporting. In contrast, if a powerful member of the local community selects recipients, he may be able to exert some informal punishment, deterring misreporting.

Misreporting could also be a mechanism for winner villages' continued advantage after the reform. If politicians shared information about the scorecard survey with connected households in advance, these households may have been more likely to under-report their assets. Even without public knowledge of the poverty score formula, which was closely guarded, it would have been obvious to those informed about the survey that under-reporting assets could benefit them.

We do not have access to the original BISP dataset. However, we test for misreporting on both our primary data and the secondary data. Household survey respondents may not distinguish between the BISP survey and other government or private surveys, or that they would under-report their assets everywhere for consistency (cf. Hurst *et al.* (2014)). This is consistent with anecdotal evidence from field survey teams in Pakistan, who report that respondents often indicate expectations of government assistance after completing a survey, despite efforts to inform them to the contrary. Since there is no benefit to the respondent of reporting assets correctly on a household survey, even a small chance that a survey might be linked with a government payment, or that the data would be cross-checked with previous reports to the government, would likely lead a respondent who has misreported assets to the BISP survey enumerators to repeat this behavior in later surveys.

[15] There is one office per *tehsil*; in Punjab, this means approximately 160 offices covering a population of 80 million households

Table 6: BISP reform, village favoritism, and inclusion errors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	HH received BISP cash transfer							
Winner x pre ref. x poor	0.027	0.020	0.023	0.018	0.045	0.055	0.020	0.018
	(0.015)*	(0.017)	(0.008)***	(0.009)*	(0.026)*	(0.029)*	(0.008)**	(0.010)*
	[0.015]*	[0.021]	[0.008]***	[0.013]	[0.025]*	[0.035]	[0.008]**	[0.011]
	{ 0.051}*	{ 0.249}	{ 0.004}***	{ 0.088}**	{ 0.280}	{ 0.370}	{ 0.014}**	{ 0.090}*
Winner x pre ref. x rich	0.015	0.015	0.001	0.009	0.014	0.014	0.015	0.015
	(0.005)***	(0.006)***	(0.007)	(0.007)	(0.005)***	(0.005)***	(0.007)**	(0.006)**
	[0.006]***	[0.009]*	[0.006]	[0.008]	[0.006]**	[0.005]***	[0.008]**	[0.008]*
	{0.007}***	{0.075}*	{0.906}	{0.129}	{0.013}**	{0.017}**	{0.060}*	{0.085}*
Winner x post ref. x poor	0.025	-0.012	0.025	0.005	0.010	-0.013	0.016	-0.009
	(0.025)	(0.025)	(0.016)	(0.018)	(0.022)	(0.027)	(0.015)	(0.019)
	[0.022]	[0.037]	[0.015]	[0.017]	[0.027]	[0.039]	[0.017]	[0.021]
	{0.362}	{0.664}	{0.153}	{0.798}	{0.650}	{0.584}	{0.301}	{0.645}
Winner x post ref. x rich	0.016	0.013	-0.004	-0.002	0.021	0.007	0.025	0.030
	(0.011)	(0.012)	(0.008)	(0.010)	(0.013)	(0.016)	(0.014)*	(0.016)*
	[0.012]	[0.016]	[0.009]	[0.011]	[0.014]	[0.018]	[0.015]*	[0.019]
	{0.190}	{0.353}	{0.584}	{0.881}	{0.145}	{0.642}	{0.126}	{0.087}*
Rich	-0.006	0.001	0.004	-0.003	-0.007	-0.016	0.001	-0.003
	(0.007)	(0.011)	(0.005)	(0.006)	(0.010)	(0.013)	(0.005)	(0.006)
	[0.007]	[0.012]	[0.005]	[0.007]	[0.010]	[0.015]	[0.006]	[0.005]
Rich x post reform	-0.039	-0.058	-0.017	-0.024	-0.009	0.008	-0.014	-0.023
	(0.016)**	(0.021)***	(0.011)*	(0.013)*	(0.017)	(0.024)	(0.010)	(0.016)
	[0.016]**	[0.022]***	[0.009]**	[0.010]**	[0.019]	[0.032]	[0.011]	[0.019]
Post reform	0.014	0.033	-0.004	0.006	-0.019	-0.029	-0.018	-0.011
	(0.025)	(0.037)	(0.020)	(0.028)	(0.024)	(0.031)	(0.023)	(0.033)
	[0.025]	[0.043]	[0.018]	[0.028]	[0.030]	[0.031]	[0.021]	[0.036]
Wealth proxy	Solid house		Ag land		Res land		Head ed 8+	
Sample	Full	RD	Full	RD	Full	RD	Full	RD
N	7386	4534	7386	4534	7386	4534	7386	4534
Mean BISP, rival village, poor								
Pre reform	0.012	0.007	0.007	0.048	0.015	0.014	0.010	0.006
Post reform	0.071	0.095	0.043	0.056	0.068	0.064	0.049	0.065
P-values:								
Winner x wealthy, pre = post	0.886	0.878	0.597	0.350	0.566	0.692	0.472	0.390
Winner x poor, pre = post	0.956	0.390	0.924	0.556	0.413	0.138	0.812	0.265
Winner x pre, wealthy = poor	0.393	0.771	0.038**	0.422	0.244	0.155	0.545	0.767
Winner x post, wealthy = poor	0.698	0.246	0.105	0.726	0.577	0.418	0.584	0.068*

Notes: Politician village sample. All specifications include controls for vote margin, HH controls listed in Column 1 of Table A1, HH controls x post reform, party dummies, party dummies x post reform. (Parentheses: robust standard errors clustered at the village level.) [Brackets: bootstrapped standard errors clustered at the village level.] {Braces: cluster wild bootstrap p-value for $H_0 : \beta = 0$.} * p < 0.1; ** p < .05; *** p < .01.

Our strategy for testing for misreporting relies on the fact that some assets can be more readily verified by simple observation during the household visit. We consider the physical status of the house to be observable. In the MICS survey, enumerators are explicitly instructed to observe this variable directly. Other variables, such as ownership of land and cattle, are less easy to observe (nearby plots or livestock could belong to landlords or neighbors).

In the MICS sample, we test whether the introduction of the scorecard decreased reported unobserved assets. Again, we exploit the phase-in of the reform, which allows district and year fixed effects. Table A10 shows the results. Columns 1-3 show specifications with observable assets as the dependent variable, while columns 4-6 test the effect on unobservable assets. If the scorecard led to misreporting, we should the reform to decreased reports of unobservable assets only. While there were strong time trends in these asset measures overall, the introduction of the scorecard has no significant impact on either observable or unobservable assets.

In our politician village sample, the balance tests (Table 2) indicate that households in winning politicians' villages do not report lower asset levels than those in the rivals' villages. In addition, we test whether the *correlation* between the observable and unobservable variables is weaker in the winner's village. If politically connected households misreport more often than others do, we should see a weaker correlation between observable and unobservable wealth proxies in the winner's village and clan. Table A11 shows the results. There is no significant evidence of greater misreporting by these groups.

We find no evidence of misreporting in general or by politically connected households. However, with the available data, we cannot rule out that respondents misreported assets on the BISP survey and then reported them correctly in these later surveys. It is also important to note that because of greater awareness of the BISP targeting system (and of the lack of consequences for misreporting), households might be more likely to misreport their assets on subsequent rounds of the poverty scorecard survey; the donor agencies and BISP administration should design the exercise to minimize these risks.

5.4 Legitimacy

In contexts with weak institutions, voters may see capture and/or clientelism as legitimate, because ethnic or caste networks help to hold politicians to account to deliver any goods or assistance, or because capture is seen as the price paid for elites bringing in resources that would not come otherwise. It is not obvious that the BISP reform would necessarily improve the perception of the program or government spending. We use the district-wise phase-in of the BISP reform to identify effects on perceived legitimacy of the government's social safety net program.

Table 8 shows the results. The dependent variable is an indicator for whether the respondent agrees that "government schemes benefit the common man". Implementation of the reform increases this approval by 10 percentage points. This is approximately a 40% increase over the pre-implementation mean. Column 2 includes

flexible controls for the scale-up effect, with fixed effects for each decile of BISP prevalance; the results are unchanged.

5.5 Local Average Treatment Effect and geographic spread of favoritism

Politicians' favoritism for their own origin villages was substantial before the reform. But each village represents only about 1% of the population of the constituency. Does this effect represent a larger pattern of favoritism or just assistance to a small group of neighbors and friends?

Our estimates from our politician village sample are based on comparing the home villages of politicians in competitive constituencies; thus the sampling excluded areas that have been consistently won by the same candidate in a landslide. The estimates are a Local Average Treatment Effect for politicians' villages in competitive constituencies. How this compares to the average treatment effect could depend on whether the favoritism effect we estimate reflects clientelism, i.e. politicians using BISP to win votes, or capture, i.e. politicians using BISP to benefit themselves and their connections for personal reasons (cf. Bardhan and Mookherjee (2012)). If politicians allocate BISP to gain votes, the size of the effect may vary between competitive and stronghold constituencies, although it is ambiguous where it would be greater (Dixit and Londregan, 1996; Cox and Mccubbins, 1986). In either case, a political motivation suggests that favoritism should extend beyond the home village, given that its size (1% of the constituency) would make it a minor part of any electoral coalition. If favoritism is motivated by personal objectives, it could be larger in less competitive constituencies where politicians can allocate resources to their chosen recipients without concern for winning over marginal voters. However, this effect could be localized to their immediate circle and have little impact on the broader population.

We cannot distinguish between capture and clientelism objectives in the effect we estimate in our politician village sample. However, for a subset of constituencies, we surveyed not only politicians' villages but also neighboring villages - those in the same *patwari* circle, an administrative group of about 3-5 villages, or about 3-5% of the constituency. Table 9 shows a version of (1) on the sample of winners' villages and nearby villages, with the effects for the two groups. The results for these villages are similar, suggesting that favoritism spread beyond the politician's immediate surroundings. In addition, the results on targeting and legitimacy are based on a representative sample, demonstrating that the reform had impacts province-wide.

5.6 Robustness checks

5.6.1 Saturation

As described in Section 2, during the period we study, the BISP program was scaled up to more beneficiaries. This could lead to concerns that the difference in estimated favoritism before and after the reform simply reflects the fact that the politician's preferred recipients were already covered, so there was no need for them to do anything

further to ensure these households benefited. However, the effect we estimate cannot be attributed to this kind of "saturation effect".

First, despite the concentration of BISP beneficiaries in the winners' villages pre-reform, even more households in the winners' villages received BISP after the reform than before it. This suggests that there were potential beneficiaries in the winners' villages before the reform who had not yet been covered, who would likely have been additional potential beneficiaries that the politicians might have wanted to select.

More importantly, 75% of the households who were nominated for the transfer under the initial system were disenrolled after the transition to the new system. We find similar results in our politician village sample: 63% of the households who received BISP transfers before the reform - those who were directly nominated by the politicians - stopped receiving BISP after the reform. So the politicians' preferred beneficiaries were clearly not all covered after the reform, ruling out saturation as the explanation for our findings.

5.6.2 Recall bias

Our data on whether a household received BISP before or after the reform is based on retrospective questions. A potential concern is that recall bias drives our results. This would be a problem if politicians emphasize their involvement in BISP more clearly to households in their own villages, and as a result these households remember more clearly than others that they started receiving the transfer shortly after the election. They thus report an earlier start date for receiving BISP. This could result in a correlation between our measure of the pre-reform period of BISP and the winning politician's village.

In addition to the survey methodology designed to minimize recall bias described in Section 3, we also address this by showing that our results hold even when using a different construction of the dependent variable which requires no detailed recall. We estimate a cross-section version of our main specification, using as the dependent variable whether the household said "yes" to the question "Has anyone in your household ever received BISP?". This does not rely on a detailed report of when the transfer was received. Since the transfer is well known by name, substantial recall bias seems implausible in this question. We then repeat this estimation with the respondent's report of receiving BISP in the last one year.

Table A12 shows the results. There is a significant effect on ever receiving BISP for the winner's village, and for both the winner's clan and other clans. This effect is still positive and significant for receiving BISP in the last year, but the coefficient is also signficantly lower for all groups. The patterns exactly mirror those in the panel estimations (Tables 4 - 5), demonstrating that our results are not driven by recall bias.

5.6.3 Substitution

After the BISP reform, politicians could use their influence over other programs to direct more assistance to their preferred recipients. As a result, the reform could decrease favoritism in BISP without improving the targeting of assistance programs as a whole. To address this, in our survey of politicians' villages we gathered information on all national and provincial targeted assistance programs the household received. Households reported in which years they received each program. We use this information and details of each program to construct an index of the approximate total value of all targeted assistance received (other than BISP). We then estimate Equation 1 with this index as the dependent variable. We also estimate the same specification as a negative binomial with the count of programs as the dependent variable. Table A15 shows the results. The winner's village term is positive but not significant; the magnitudes are similar before and after the BISP reform and are not significantly different from each other. There is no evidence of a pattern of substitution of favoritism into other benefits.

5.6.4 Endogenous reporting of clan

Our "same clan" variable is based on a question at the beginning of the survey, which is not framed with reference to the politician and comes up before any such topics are discussed. However, it is possible that households change their reported clan in our survey based on which politician won the election. Cassan (2015) demonstrates that households in British India changed their reported caste identity between census rounds in response to legal restrictions imposed on the roles of different castes.

We test whether in our politician village sample, households are more likely to report the same clan as the local politician if he won. A13 shows the results. There is no significant evidence of this type of endogenous clan reporting in our sample.

In addition, the reported clan is constant for the pre- and post-reform periods in the data. Thus it could not drive the pattern of post-reform reduction in favoritism for the winner's clan, shown in Table 5.

5.7 Mechanisms

Before the BISP reform, politicians had an official role in selecting recipients. However, the new beneficiary selection mechanism was designed to avoid this. The evidence suggesting persistence of favoritism shown in Table 4 begs the question: how did politicians influence targeting after the reform? In this section, we consider several possible mechanisms for the post-reform effect.

5.8 Politician power

More powerful politicians might effectively find a way to influence the distribution of BISP despite the reform. To test this, we interact the winner's village term with two proxies for politician power: the politician's membership in the governing (majority) party, and his individual level of seniority, as measured by membership in key parliamentary committees. Tables A5 and A6 show the results of the main specification interacted with the politician's party and seniority levels, respectively. The results do *not* indicate that more powerful politicians exert greater favoritism. The point estimates do suggest stronger favoritism for the winner's village in these areas before the reform; however, these differences are not significant. Across these specifications, there is no consistent pattern of a greater reduction in favoritism for less powerful politicians. We do not find evidence that the BISP reform constrained only the weak.

5.8.1 "Grandfathering" in previous recipients

During the rollout of the new targeting mechanism, previous BISP recipients still continued to receive the transfers. By 2011, the transition to the new system was complete, with ineligible households removed from the recipient rolls. For this reason, we only include transfers from 2008 and 2009 as pre-reform, and 2012 and 2013 as post-reform, as detailed in section 4. However, it might still be possible that previous recipients stayed on the rolls due to politicians' interference.

Overall, seventy-five percent of the recipients selected by politicians in the first phase of the program were found ineligible and disenrolled after the scorecard exercise. If politicians intervened to keep some well-connected households on the rolls, we would expect to see this reflected in our politician village sample. However, in our sample, sixty-three percent of the households in our sample who received BISP transfers before the reform stopped receiving BISP after the reform. To test for grandfathering formally, we also estimate the cross-sectional version of our main specification on only the subsample of households who had not received BISP in the pre-reform period. Those who received BISP previously are dropped from the sample. Table 10 shows the results. The results are similar when restricting the sample to households who had not received the transfer before, and the coefficients of interest are not significantly different between the estimates on the full sample and subsample. Clearly "grandfathering" does not explain any post-reform favoritism.

5.8.2 Survey coverage and data manipulation

The data used for targeting BISP could be subject to errors or manipulation at several stages. Households could misreport their assets to the survey enumerators. Politicians could help households connected to them to anticipate the survey and how to respond strategically. As discussed in section 5.3, we do not have access to the BISP data, but we find no evidence of differential patterns of misreporting in our household survey data.

Households could be missed during the national survey to collect data for the poverty scorecard, either due to error, enumerator shirking, or because they were not present on the day the survey team came to the area. Local influentials who met with the survey team could ensure coverage of the areas of a village considered highest status or most important, under the influence of politicians from the area. In addition, during the "mop-up exercise" the field teams relied on local officials to ensure coverage of households missed during the main fieldwork. Well-connected households might have been more likely to be informed and identified for the followup exercise.

Finally, it is possible that politicians could exert pressure on the agencies that enter and manage the data to manipulate results for selected households. However, the BISP reform incorporated separation of agencies involved in this process, with the national identity card agency NADRA responsible for managing the data and external agencies hired to enter it, as well as an extensive audit process in which the scorecard variables were compared between the hard copy forms and database. The World Bank maintained involvement in this audit process. Interfering in this process might be possible, but would likely be difficult due to the number of actors involved and the level of scrutiny.

5.8.3 Targeted troubleshooting and appeals

An alternative mechanism for the effect we estimate after the reform is targeted troubleshooting. Politicians or their local agents may assist households who are having problems with a program. For example, if the cash transfers are being stolen by low-level officials involved in their distribution, households may approach a politician for help in addressing the issue, and the politician may intervene with the bureaucracy.

Alternatively, officials may ensure the politician's close contacts get good service, in the hopes of currying favor or to avoid problems. There were widespread problems with the cash delivery to households under the original distribution system through the post office (Khan and Qutub, 2010), so this seems plausible. The appeals process also provides scope for politicians to play a role post reform. Although the post-appeal selection was still not discretionary, politicians may have helped households in their villages understand and complete the appeals process.

5.8.4 Assistance with identity cards

Female members of recipient households had to get a National Identity Card issued if they did not already have one. This requirement was constant over the pre and post reform period. It is possible that politicians or their agents provided assistance with this part of the process, to help potential recipients get enrolled in BISP; Chaudhry and Vyborny (2013) find that assistance getting an ID issued is by far the most common task on which households report assistance from local influentials in rural Punjab. These ID cards are used for a range of official purposes

including legal transactions and voting, so there are many reasons why a politician might assist constituents in obtaining these cards. However, these factors did not change over the period of the BISP reform.

In the politician village survey, we collected data on new ID cards issued to household members over the past ten years. Table 11 presents two tests using this data. Columns 1-2 show our basic estimations with receipt of a new ID card as the dependent variable. The results are remarkably similar to the pattern of results for BISP in the main estimations: household members in winners' villages are significantly more likely to receive a new ID card, and this effect drops significantly for members of the same clan after the reform.

Columns 3-6 show the same specifications with BISP transfer as the dependent variable; in columns 4 and 6 we include receipt of a new ID card as a control variable, and test whether the coefficient of interest changes. Introducing the ID card as a mechanism variable reduces estimated favoritism significantly both pre and post reform, again in a pattern that is consistent with the pattern of favoritism results overall.

These results suggest that winning politicians assisted individuals in their villages to get ID cards, and that this played a significant role in the remaining favoritism after the reform. Despite the transition to a rules-based system for selecting recipients, some potential recipients still faced administrative hurdles in getting the transfer; assistance from a winning politician helped to clear these hurdles.

6 Discussion

While the existing literature on politician favoritism is extensive, few studies examine the causal impact of specific institutions or policies which might constrain favoritism. In this paper, we have demonstrated that an institutional reform imposed from outside substantially reduced favoritism for high level politicians' closest connections in allocating public spending.

The evidence is not consistent with a politician selecting the poorest among his network: pre-reform, politicians selected wealthier households at the expense of poorer households even within their own villages, many of whom they knew personally. It is only consistent with politicians - and this was effectively reduced by the institutional reform. Province-wide, the reform substantially improved the quality of targeting of the program, at an administrative cost of less than 2% of the budget paid out to beneficiares.

International donors designed and advocated the new system; but the government has kept it in place, and covers 90% of the BISP program's budget. Although donors have made BISP spending a continued condition of donor support overall, the mixed history of compliance with such conditions makes the government's continued investment in this program far from guaranteed. The new party in government from 2013 kept the program's features, simply placing the faces of its leading politicians in prominent placement on the program materials in place of the previous

party's deceased leader, Benazir Bhutto. The program's budget was in fact increased after an election and change in power to a new political party - breaking with Pakistan's historical trend of each national government defunding social protection programs instituted by previous governments. Finally, the reform also dramatically increased the perceived legitimacy of the government's transfer programs. The BISP reform, designed and imposed from outside Pakistan, made a substantial positive impact on the country's institutions for social assistance.

Table 7: BISP scorecard reform and inclusion errors: provincial MICS sample

Panel A: Inclusion error based on indicators targeted on new BISP scorecard					
	(1)	(2)	(3)	(4)	(5)
	HH received BISP cash transfer				
Wealthy	-0.0244*	-0.0217***	-0.0198***	-0.0201***	-0.0360***
	(0.0128)	(0.00460)	(0.00453)	(0.00223)	(0.00318)
Post reform	0.191***	0.0523**	0.0348	0.0395*	0.0246
	(0.0354)	(0.0209)	(0.0207)	(0.0198)	(0.0209)
Wealthy x post reform	-0.167***	-0.0689***	-0.0542***	-0.0422***	0.00408
	(0.0293)	(0.0133)	(0.0117)	(0.00845)	(0.00576)
Round = 2014	0.00926	0.00953	0.0101	0.0116	0.00808
	(0.0176)	(0.0152)	(0.0167)	(0.0164)	(0.0181)
Wealth proxy used	BISP eligibility (est.)	MICS wealth index Q4-5	MICS wealth index Q5	Head ed grade 8+	Ag land
District FE	YES	YES	YES	YES	YES
Mean BISP, poor, pre-reform	0.066	0.040	0.036	0.041	0.037
N	133562	133562	133562	125127	133512

Panel B: Inclusion error based on indicators not targeted on new BISP scorecard						
	(1)	(2)	(3)	(4)	(5)	(6)
	HH received BISP cash transfer					
Wealthy	-0.00380	-0.0108***	-0.0122***	0.00658	-0.0162***	-0.0133***
	(0.00262)	(0.00377)	(0.00444)	(0.00629)	(0.00307)	(0.00252)
Post reform	0.0438*	0.0688***	0.0634**	0.0962***	0.0289	0.0128
	(0.0223)	(0.0240)	(0.0240)	(0.0261)	(0.0216)	(0.0198)
Wealthy x post reform	-0.0443***	-0.0747***	-0.0480***	-0.0927***	-0.0498***	0.0151**
	(0.0103)	(0.0138)	(0.0110)	(0.0173)	(0.00919)	(0.00622)
Round = 2014	0.00773	0.0149	0.0114	0.0202	0.0146	0.00870
	(0.0184)	(0.0175)	(0.0189)	(0.0187)	(0.0188)	(0.0181)
Wealth proxy used	Urban	Solid roof	Solid walls	Solid floor	Spouse ed 8+	Own dwelling
District FE	YES	YES	YES	YES	YES	YES
Mean BISP, poor, pre-reform	0.032	0.041	0.043	0.038	0.034	0.038
N	133336	133336	132161	132901	122031	133543

Notes: Provincially representative Multiple Indicator Cluster Survey, 2011 and 2014 rounds. Post reform is an indicator variable defined at the district-year level as described in Section 4. Standard errors in parentheses, clustered at the district level. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 8: BISP reform and legitimacy of public programs

	(1)	(2)
	Respondent agrees "government schemes are beneficial to the common man"	
Post reform	0.102***	0.0894***
	(0.0298)	(0.0310)
HH received BISP	0.392***	0.387***
	(0.0183)	(0.0199)
HH received other government transfers	0.307***	0.318***
	(0.0247)	(0.0295)
HH in top 2 wealth quintiles and received BISP	-0.0832***	-0.0772***
	(0.0249)	(0.0259)
HH in top 2 wealth quintiles and received other gov't transfers	-0.209***	-0.206***
	(0.0415)	(0.0421)
BISP prevalance in district in survey year	-0.0600	
	(0.312)	
Survey year = 2014	-0.0475*	-0.0354
	(0.0285)	(0.0263)
Decile dummies of BISP prevalance in district in survey year		X
Wealth quintile dummies	X	X
District dummies	X	X
Mean Y - pre reform	0.242	
N	130132	130132

Notes: Multiple Indicator Cluster Survey, 2011 and 2014 rounds. Post reform is an indicator variable defined at the district-year level as described in Section 4. BISP prevalance is defined as the fraction of sampled households in the district-year who received BISP, excluding the respondent. Standard errors in parentheses, clustered at the district level. * p < .1, ** p < .05, *** p < .01.

Table 9: Geographic spread of favoritism

	(1)	(2)	(3)
	HH received BISP cash transfer		
Winner's village, pre reform	0.018	0.019	0.018
	(0.008)**	(0.006)***	(0.010)*
	[0.008]**	[0.006]***	
	{0.019}**	{0.000}***	
Winner's village, post reform	0.012	0.024	0.015
	(0.017)	(0.013)*	(0.013)
	[0.015]	[0.016]	
	{0.415}	{0.112}	
Nearby village, pre reform	0.022	0.024	0.014
	(0.011)**	(0.010)**	(0.009)
	[0.014]	[0.010]**	
	{0.030}**	{0.002}***	
Nearby village, post reform	0.008	0.012	0.002
	(0.016)	(0.012)	(0.011)
	[0.017]	[0.014]	
	{0.528}	{0.140}	
Post reform	0.036	0.021	-0.024
	(0.009)***	(0.032)	
	[0.010]***	[0.035]	
Vote margin		X	X
Vote margin x winner's village			X
HH control variables		X	X
Additional HH control variables		X	
HH control variables x post reform		X	X
Party dummies		X	
Party dummies x post reform		X	X
Constituency dummies			X
P-values (robust SE clustered at village level):			
Winner's village = nearby village, pre reform	0.755	0.660	0.702
Winner's village = nearby village, post reform	0.811	0.348	0.284
Nearby village pre reform = nearby village post reform	0.420	0.422	0.473
N	9736	9736	9736

Notes: Politician village sample including additional nearby (same *patwari* circle) villages as described in section 3. (Parentheses: robust standard errors clustered at the village level.) [Brackets: bootstrapped standard errors clustered at the village level.] {Braces: P-value for cluster wild bootstrap test of $H_0 : \beta = 0$.} Basic and additional household controls are those shown in Column 1 and 2 of Table A1, respectively. * p < 0.1; ** p < .05; *** p < .01.

Table 10: Grandfathering recipients

	(1)	(2)	(3)	(4)
	Received BISP cash transfer (post reform)			
	All	HHs with no pre-reform BISP	All	HHs with no pre-reform BISP
Winner's village	0.023	0.018		
	(0.012)*	(0.010)*		
Winner's village x winner's clan			-0.004	-0.000
			(0.020)	(0.020)
Winner's village x other clans			0.025	0.020
			(0.013)*	(0.011)*
HH control variables	X	X	X	X
Additional HH control variables	X	X	X	X
N	3693	3631	3693	3631
P-values, coefficients equal:				
Winner's village, models 1 and 2		0.157		
Winner's village x other clan, models 3 and 4				0.127

Notes: Politician village sample, post-reform observations only. Columns 1 and 3 are estimated on the full sample; columns 2 and 4 are estimated on the sample of households who did not receive BISP before the reform (2008-9). Standard errors in parentheses, clustered at village level. * $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$.

Table 11: Mechanism test: ID cards

	(1)	(2)	(3)	(4)	(5)	(6)
	HH received new ID card		HH received BISP cash transfer			
Winner's village, pre-reform	0.020		0.019	0.002		
	(0.007)***		(0.007)***	(0.001)*		
	[0.007]***		[0.007]***	[0.002]		
	0.036**		0.036**	0.073*		
Winner's village, post-reform	0.026		0.019	-0.003		
	(0.015)*		(0.013)	(0.004)		
	[0.016]		[0.017]	[0.004]		
	0.175		0.225	0.463		
Winner's village, winner's clan, pre-reform		0.039			0.039	0.006
		(0.014)***			(0.015)***	(0.004)
		[0.024]*			[0.019]**	[0.004]
		0.145			0.130	0.129
Winner's village, winner's clan, post-reform		0.008			-0.005	-0.011
		(0.017)			(0.016)	(0.004)***
		[0.023]			[0.019]	[0.005]**
		0.625			0.878	0.091*
Winner's village, other clan, pre-reform		0.016			0.015	0.002
		(0.007)**			(0.007)**	(0.001)
		[0.010]*			[0.008]*	[0.001]*
		0.067*			0.054*	0.097*
Winner's village, other clan, post-reform		0.033			0.026	-0.002
		(0.017)**			(0.015)*	(0.004)
		[0.018]*			[0.015]*	[0.004]
		0.109			0.141	0.630
HH received new ID card				0.849		0.849
				(0.020)***		(0.020)***
				[0.022]***		[0.024]***
Observations	7386	7386	7386	7386	7386	7386
Mean Y:						
Rival village, pre reform	0.008			0.008		
Rival village, post reform	0.048			0.040		
Rival village, other clan, pre reform		0.009				0.008
Rival village, other clan, post reform		0.050				0.041
P-values (robust SE clustered at village level):						
Winner's village, pre = post						
Robust clustered by village	0.759					
Bootstrap clustered by village	0.740					
Wild-bootstrap clustered by constituency	0.663					
Winner's village, winner's clan, pre = post						
Robust clustered by village		0.040**				
Bootstrap clustered by village		0.077*				
Wild-bootstrap clustered by constituency		0.063*				
Winner's village, other clan, pre = post						
Robust clustered by village		0.366				
Bootstrap clustered by village		0.339				
Wild-bootstrap clustered by constituency		0.253				
Winner's village, pre equal model 3 v. 4				0.007***		
Winner's village, post equal model 3 v. 4				0.081*		
Winner's village, winner's clan, pre equal model 5 v. 6						0.004***
Winner's village, winner's clan, post equal model 5 v. 6						0.650
Winner's village, other clan, pre equal model 5 v. 6						0.029**
Winner's village, other clan, post equal model 5 v. 6						0.049**

Notes: Politician village sample. All specifications include an indicator for post reform, vote margin, and household controls are those shown in Column 1 of Table A1. Columns 2, 5 and 6 include indicators for winner's clan and winner's clan x post reform. (Parentheses: robust standard errors clustered at the village level.) [Brackets: bootstrapped standard errors clustered at the village level.] {Braces: P-value for cluster wild bootstrap test of $H_0 : \beta = 0$.} * p < 0.1; ** p < .05; *** p < .01.

References

Acemoglu, D., Johnson, S. and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, **91** (5), 1369–1401.

—, Reed, T. and Robinson, J. a. (2014). Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone. *Journal of Political Economy*, **122** (2), 319–368.

— and Robinson, J. (2008). The Role of Institutions in Growth and Development. *World Bank, Washington DC*, **1**, 1–33.

—, Robinson, J. A. and Santos, R. J. (2013). The monopoly of violence: Evidence from colombia. *Journal of the European Economic Association*, **11** (SUPPL. 1), 5–44.

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., Purnamasari, R. and Wai-poi, M. (2013a). Does Elite Capture Matter? Local Elites and Targeted Welfare Programs in Indonesia. *NBER Working Paper No. 19798*.

—, —, —, — and Tobias, J. (2012). Targeting the Poor: Evidence from a Field Experiment in Indonesia. *American Economic Review*, **102** (4), 1206–1240.

—, Hanna, R., Olken, B. A. and Wai-poi, M. (2013b). Ordeal Mechanisms in Targeting: Theory and evidence from a field experiment in Indonesia. *NBER Working Paper No. 19127*.

Alderman, H. (2002). Do local officials know something we don't? Decentralization of targeted transfers in Albania. *Journal of Public Economics*, **83** (3), 375–404.

Anderson, S., Francois, P. and Kotwal, A. (2015). Clientelism in Indian Villages. *American Economic Review*, **105** (6), 1780–1816.

Asian Development Bank (2013). Pakistan: Social Protection Development Project. *Report RRP PAK 45233*.

Banerjee, A., Duflo, E., Imbert, C., Mathew, S. and Pande, R. (2016). Can E-Governance Reduce Capture of Public Programs? Experimental Evidence from a Financial Reform of India's Employment Guarantee. *Mimeo, MIT*.

—, Hanna, R., Kyle, J. and Olken, B. A. (2015). Information is Power: Identification Cards and Food Subsidy Programs in Indonesia. *NBER Working Paper No. 20923*.

— and Iyer, L. (2005). History, institutions, and economic performance: The legacy of colonial land tenure systems in India. *American Economic Review*, **95** (4), 1190–1213.

Banerjee, A. V. and Duflo, E. (2014). Under the Thumb of History? Political Institutions and the Scope for Action. *Annual Review of Economics*, **6**, 951–971.

—, Green, D., McManus, J. and Pande, R. (2014). Are Poor Voters indifferent to whether elected leaders are criminal or corrupt? *Political Communication*, **31** (April 2015), 37–41.

Banful, A. B. (2011). Do formula-based intergovernmental transfer mechanisms eliminate politically motivated targeting? Evidence from Ghana. *Journal of Development Economics*, **96** (2), 380–390.

Bardhan, P. and Mookherjee, D. (2000). Capture and Governance at Local and National Levels. *American Economic Review*, **90** (2), 135–139.

— and — (2005). Decentralizing antipoverty program delivery in developing countries. *Journal of Public Economics*, **89** (4 SPEC. ISS.), 675–704.

— and — (2006). Pro-poor targeting and accountability of local governments in West Bengal. *Journal of Development Economics*, **79** (2), 303–327.

— and — (2012). Political Clientelism and Capture: Theory and Evidence from West Bengal, India.

Barrientos, A. and Hulme, D. (2009). Social protection for the poor and poorest in developing countries: reflections on a quiet revolution: commentary. *Oxford Development Studies*, **37** (4), 439–456.

Basurto, P., Dupas, P. and Robinson, J. (2016). Decentralization and Efficiency of Subsidy Targeting: Evidence from Chiefs in Rural Malawi. *Mimeo, Stanford University*.

Beath, A., Christia, F. and Enikolopov, R. (2013). Empowering women through development aid: Evidence from a field experiment in Afghanistan. *American Political Science Review*, **107** (03), 540557.

Besley, T. (1990). Means Testing versus Universal Provision in Poverty Alleviation Programmes. **57** (225), 119–129.

—, Pande, R., Rahman, L. and Rao, V. (2004). The Politics of Public Good Provision : Evidence from Indian Local Governments . *Journal of the European Economic Association*, **2** (2-3), 416–426.

—, — and Rao, V. (2007). Political Economy of Panchayats in South India. *Economic and Political Weekly*, (IFCO 2190), 661–666.

—, — and — (2012). Just Rewards? Local Politics and Public Resource Allocation in South India. *The World Bank Economic Review*, **26** (2), 191–216.

Björkman-Nyqvist, M., Walque, D. D. and Svensson, J. (2014). Information is Power: Experimental Evidence of the Long Run Impact of Community Based Monitoring. *World Bank Policy Research Working Paper No. 7015*.

Bourguignon, F. and Platteau, J.-P. (2014). The Hard Challenge of Aid Coordination. *World Development*.

Bräutigam, D. A. and Knack, S. (2004). Foreign Aid, Institutions, and Governance in Sub-Saharan Africa. *Economic Development and Cultural Change*, **52** (2), 255–285.

Briggs, R. C. (2014). Aiding and abetting: Project aid and ethnic politics in kenya. *World Development*, **64**, 194–205.

Brodeur, A., Le, M., Sangnier, M. and Zylberberg, Y. (2016). Star Wars: The Empirics Strike Back. *American Economic Journal: Applied Economics*.

Burgess, R., Miguel, E., Morjaria, A., Jedwab, R. and Padró i Miquel, G. (2015). The Value of Democracy: Evidence from Road Building In Kenya. *American Economic Review*, **105** (6), 1817–1851.

Caeyers, B. and Dercon, S. (2012). Political Connections and Social Networks in Targeted Transfer Programs: Evidence from Rural Ethiopia. *Economic Development and Cultural Change*, **60** (4), 639–675.

Callen, M., Gulzar, S., Hasanain, A. and Khan, M. Y. (2016). *The Political Economy of Public Sector Absence: Experimental Evidence from Pakistan*. CEPR Discussion Papers 11321, C.E.P.R. Discussion Papers.

Camacho, A. and Conover, E. (2011). Manipulation of Social Program Eligibility. *American Economic Journal: Economic Policy*, **3** (2), 41–65.

Cameron, A. C., Gelbach, J. B. and Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics*, **90** (August), 414–427.

Carozzi, F. and Repetto, L. (2014). Sending the Pork Home: Birth Town Bias in Transfers to Italian Municipalities. *Journal of Public Economics*, **134**, 42–52.

Casey, K., Glennerster, R. and Miguel, E. (2012). Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan. *Quarterly Journal of Economics*, pp. 1755–1812.

Cassan, G. (2015). Identity-Based Policies and Identity Manipulation: Evidence from Colonial Punjab. *American Economic Journal: Economic Policy*, **7** (4), 103–131.

Chaudhry, A. and Vyborny, K. (2013). Patronage in Rural Punjab: Evidence from a New Household Survey Dataset. *Lahore Journal of Economics*, **18** (Special Edition), 183–209.

Cheema, A. and Mohmand, S. (2009). Bringing Electoral Politics to the Doorstep: Who Gains, Who Loses? *Mimeo, Lahore University of Management Sciences*.

—, Mohmand, S. K. and Patnam, M. (2009). Colonial Proprietary Elites and Institutions: Persistence of De Facto Political Dominance. *Mimeo, Lahore University of Management Sciences*.

—, Naqvi, A. A., Naseer, F. and Siddiqi, B. (2012). Land Rights and Long-Run Development: Evidence from Pakistan. *Mimeo, Stanford University*.

Cheema, I., Farhat, M., Hunt, S., Javeed, S., Keck, K. and O'Leary, S. (2015). Benazir Income Support Programme: Second Impact Evaluation Report. *Oxford Policy Management report*, (December).

—, —, —, —, Pellerano, L. and Leary, S. O. (2014). Benazir Income Support Programme:: First follow-up impact evaluation report. *Oxford Policy Management report*.

Clark, G. (2001). Pakistan's Zakat System: A Policy Model for Developing Countries as a Means of Redistributing Income to the Elderly Poor. *Social Thought*, **20** (3), 47–75.

Coady, D., Grosh, M. and Hoddinott, J. (2004). Targeting of Transfers in Developing Countries: Review of Lessons and Experience. *World Bank - IFPRI report*.

Conning, J. and Kevane, M. (2000). Community Based Targeting Mechanisms for Social Safety Nets.

Cox, G. W. and Mccubbins, M. D. (1986). Electoral Politics as a Redistributive Game. *Journal of Politics*, **48** (2), 370–389.

Currie, J. (2004). The Take Up of Social Benefits. *NBER Working Paper No. 10488*.

Deaton, A. S. (2012). Your wolf is interfering with my t-value! *Royal Economic Society Newsletter*, **159** (4).

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica*, **78** (6), 1863–1903.

Dixit, A. and Londregan, J. (1996). The Determinants of Success of Special Interests in Redistributive Politics. *The Journal of Politics*, **58** (04), 1132.

Djankov, S., Montalvo, J. G. and Reynal-Querol, M. (2008). The curse of aid. *Journal of Economic Growth*, **13** (3), 169–194.

Do, Q.-a., Nguyen, K.-T. and Tran, A. (2016). One Mandarin Benefits the Whole Clan : Hometown Favoritism in an Authoritarian Regime. *CEP Discussion Paper, No 1409*.

Duclos, J.-Y. (1995). Assessing the performance of an income tax. *Bulletin of Economic Research*, **47** (2), 115–126.

Dupas, P., Hoffmann, V., Kremer, M. and Zwane, A. P. (2016). Targeting health subsidies through a nonprice mechanism: A randomized controlled trial in Kenya. *Science*, **353** (6302), 889–895.

Eggers, A. C., Fowler, A., Hainmueller, J., Hall, A. B. and Snyder, J. M. (2015). On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races. *American Journal of Political Science*, **59** (1), 259–274.

Fafchamps, M. and Labonne, J. (2016a). Do Politicians' Relatives Get Better Jobs? Evidence from Municipal Elections in the Philippines. *Mimeo, Stanford University.*

— and — (2016b). Using Split Samples to Improve Inference on Causal Effects. *NBER Working Paper No. 21842.*

Ferraz, C. and Finan, F. (2008). Exposing Corrupt Politicians: the Effects of Brazil's Publicly Released Audits on Electoral Outcomes. *Quarterly Journal of Economics*, **123** (2), 703–745.

Filmer, D. and Pritchett, L. H. (2001). Estimating Wealth Effects Without Expenditure Data—Or Tears: An Application To Educational Enrollments In States Of India. *Demography*, **38** (1), 115–132.

Fujiwara, T. and Wantchekon, L. (2013). Can informed public deliberation overcome clientelism? Experimental evidence from Benin. *American Economic Journal: Applied Economics*, **5** (4), 241–255.

Galasso, E. and Ravallion, M. (2005). Decentralized targeting of an antipoverty program. *Journal of Public Economics*, **89** (4), 705–727.

Gazdar, H. (2011). Social protection in Pakistan: in the midst of a paradigm shift? *CSP Research Report, Institute for Development Studies, Sussex, UK.*

Gehlbach, S. and Simpser, A. (2015). Electoral Manipulation as Bureaucratic Control. *American Journal of Political Science*, **59** (1), 212–224.

Gerber, A. and Malhotra, N. (2008). Do statistical reporting standards affect what is published? Publication bias in two leading political science journals. *Quarterly Journal of Political Science*, **3**, 313–326.

Gille, V. (2013). Stigma in positive discrimination application? Evidence from quotas in education in India. *Mimeo, Universite Paris Sorbonne.*

Grimmer, J., Hersh, E., Feinstein, B. D. and Carpenter, D. (2011). Are Close Elections Random? *Mimeo, Stanford University.*

Gulzar, S. and Pasquale, B. J. (2016). Politicians , Bureaucrats , and Development: Evidence from India. *Mimeo, NYU.*

Hodler, R. and Raschky, P. A. (2014). Regional Favoritism. *Quarterly Journal of Economics*, pp. 995–1033.

Hou, X. (2009). Poverty Scorecard for Pakistan: An Update Using the PSLM20072008 Data 1 November 9, 2009. *Mimeo, World Bank*.

Hsieh, B. C.-t., Miguel, E., Ortega, D. and Rodriguez, F. (2011). The Price of Political Opposition: Evidence from Venezuela's Maisanta. *American Economic Journal: Applied Economics*, **3** (April), 196–214.

Humphreys, M., Sanchez de la Sierra, R. and van der Windt, P. (2013). Fishing. *Political Analysis*.

—, — and Windt, P. V. D. (2015). Social engineering in the tropics: a grassroots democratization experiment in the Congo. *Mimeo, Columbia University*.

Hurst, E., Li, G. and Pugsley, B. (2014). Are Household Surveys Like Tax Forms: Evidence from Income Underreporting of the Self Employed. *The Review of Economics and Statistics*, **96** (1), 19–33.

International Monetary Fund (2016). Tenth Review Under the Extended Arrangement and Request for Modification of Performance Criteria. *IMF Pakistan Country Report No. 16/94*.

Iyer, L. and Mani, A. (2012). Traveling agents: political change and bureaucratic turnover in India. *Review of Economics and Statistics*, **94**, 723–739.

Jablonski, R. S. (2014). How Aid Targets Votes: The Impact of Electoral Incentives on Foreign Aid Distribution. *World Politics*, **66** (2), 1–39.

Jayne, T. S., Strauss, J., Yamano, T. and Molla, D. (2001). Giving to the poor? Targeting of food aid in rural Ethiopia. *World Development*, **29** (5), 887–910.

Kanbur, R. (2000). Aid, Conditionality and Debt in Africa. In *Foreign Aid and Development: Lessons Learnt and Directions for the Future*, p. 10.

Keefer, P. (2007). Clientelism, Credibility, and the Policy Choices of Young Democracies. *American Journal of Political Science*, **51** (4), 804–821.

— and Vlaicu, R. (2007). Democracy, Credibility, and Clientelism. *Journal of Law, Economics, and Organization*, **24** (2), 371–406.

Khan, S. N. and Qutub, S. (2010). The Benazir Income Support programme and the Zakat programme: A political economy analysis of gender. *Working Paper, Overseas Development Institute*.

Khwaja, A. I. and Mian, A. (2005). Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market. *The Quarterly Journal of Economics*, **120** (4), 1371–1411.

Kilby, C. (2009). The political economy of conditionality: An empirical analysis of World Bank loan disbursements. *Journal of Development Economics*, **89** (1), 51–61.

Kilic, T., Whitney, E. and Winters, P. (2013). Decentralized Beneficiary Targeting in Large-Scale Development Programs Insights from the Malawi Farm Input Subsidy Program. *World Bank Policy Research Working Paper 6713*, (November).

Killick, T. (1998). Donor Conditionality and Policy Reform. In *The Political Dimension of Economic Growth*, pp. 278–293.

Kitschelt, H. and Wilkinson, S. (2007). *Patrons, Clients and Policies: Patterns of Democratic Accountability and Political Competition*.

Leary, S. O., Cheema, I., Hunt, S. and Pellerano, L. (2011). Benazir Income Support Programme Impact Evaluation Baseline Survey Report. *Oxford Policy Management report*.

Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, **142** (2), 675–697.

Litschig, S. (2012). Are rules-based government programs shielded from special-interest politics? Evidence from revenue-sharing transfers in Brazil. *Journal of Public Economics*, **96** (11-12), 1047–1060.

Loftus, E. F. and Marburger, W. (1983). Since the eruption of Mt. St. Helens, has anyone beaten you up? Improving the accuracy of retrospective reports with landmark events. *Memory & cognition*, **11** (2), 114–20.

Molenaers, N., Dellepiane, S. and Faust, J. (2015). Political Conditionality and Foreign Aid. *World Development*, **75**, 2–12.

Montinola, G. R. (2010). When does aid conditionality work? *Studies in Comparative International Development*, **45** (3), 358–382.

Moss, T., Pettersson, G. and van de Walle, N. (2006). An Aid-Institutions Paradox? A Review Essay on Aid Dependency and State Building in Sub-Saharan Africa. *Center for Global Development Working Papers*, (74), 1–28.

Mu, R. and Zhang, X. (2011). The Role of Elected and Appointed Village Leaders in the Allocation of Public Resources Evidence from a Low-Income Region in China. *IFPRI Discussion Papers*, (January).

Muralidharan, K., Niehaus, P. and Sukhtankar, S. (2016). Building State Capacity: Evidence from Biometric Smartcards in India. *American Economic Review*, **106** (10), 2895–2929.

Nabi, I. (2013). Two Social Protection Programs in Pakistan. *Lahore Journal of Economics*, **18** (September), 283–304.

Nath, A. (2016). Bureaucrats and Politicians. *Boston University IED Working Paper 269*.

Nayab, D. and Farooq, S. (2012). Effectiveness of cash transfer programmes for household welfare in Pakistan: The case of the Benazir income support program. *Poverty and Social Dynamics Paper Series No. 4*.

Neter, J. and Waksberg, J. (1964). A study of response errors in expen- ditures data from household interview. *Journal of the American Statistical Association,*, **59**, 18–55.

Niehaus, P., Atanassova, A., Bertrand, M. and Mullainathan, S. (2013). Targeting with Agents. *American Economic Journal: Economic Policy*, **5** (1), 206–238.

Öhler, H. and Nunnenkamp, P. (2014). Needs-based targeting or favoritism? The regional allocation of multilateral aid within recipient countries. *Kyklos*, **67** (3), 420–446.

Olken, B. A. (2015). Promises and Perils of Pre-Analysis Plans. *Journal of Economic Perspectives*, **29** (3), 61–80.

Pande, R. (2011). Can Informed Voters Enforce Better Governance? Experiments in Low Income Democracies. *Annual Review of Economics*.

Platteau, J.-p. (2004). Monitoring Elite Capture in Community-Driven Development. *Development and Change*, **35** (2), 223–246.

Punjab Bureau of Statistics (2008). Multiple Indicator Cluster Survey (MICS) Punjab Provincial Report.

Ravallion, M., Van De Walle, D., Dutta, P. and Murgai, R. (2013). Testing Information Constraints on India's Largest Antipoverty Program. *World Bank Policy Research Working Paper 6598*.

Rutstein, S. O. and Johnson, K. (2004). The DHS Wealth Index. *USAID DHS Comparative Reports No. 6*.

Santiso, C. (2001). Good Governance and Aid Effectiveness : The World Bank and Conditionality. *The Georgetown Public Policy Review*, **7** (1), 1–22.

Sekhon, J. S. and Caughey, D. (2011). Elections and the Regression Discontinuity Design: Lessons from Close US House Races, 1942-2008. *Political Analysis*, **19** (4).

Simpser, A. (2013). *Why Governments and Parties Manipulate Elections: Theory, Practice, and Implications.* Cambridge University Press.

Stokes, S. C., Dunning, T., Nazareno, M. and Brusco, V. (2013). *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics.* Cambridge University Press.

Wade, R. (1985). The market for public office: Why the Indian state is not better at development. *World Development*, **13** (4), 467–497.

Wantchekon, L. (2003). Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin. *World Politics*, **55** (03), 399–422.

Weitz-Shapiro, R. (2012). What Wins Votes: Why Some Politicians Opt Out of Clientelism. *American Journal of Political Science*, **56** (3), 568–583.

World Bank (2013). Pakistan: Towards an Integrated National Safety Net System. *World Bank Report No. 66421-PK.*

A Supplementary tables

Table A1: HH correlates of receiving BISP cash transfer

	(1)	(2)	(3)
	HH ever received BISP cash transfer		
Female head	0.161	0.149	0.005
	(0.050)***	(0.047)***	(0.002)**
Any daughters currently aged 18-25		0.050	
		(0.011)***	
Rudimentary house[a]	0.070	0.060	0.038
	(0.012)***	(0.014)***	(0.004)***
No agricultural land	0.040	0.027	0.030
	(0.012)***	(0.012)**	(0.002)***
Less than 12.5 acres agricultural land	-0.004	-0.012	0.015
	(0.021)	(0.019)	(0.002)***
No cattle[a]		0.013	-0.012
		(0.010)	(0.002)***
No residential land		0.023	0.002
		(0.021)	(0.002)
HH head 5th grade or higher		-0.024	-0.014
		(0.013)*	(0.002)***
HH head 8th grade or higher		0.008	-0.010
		(0.015)	(0.002)***
HH head 10th grade or higher		-0.004	-0.018
		(0.016)	(0.002)***
Years HH has lived in village		-0.001	
		(0.002)	
Years squared		0.000	
		(0.000)	
Related to elected or local official		-0.001	
		(0.013)	
Any HH member had a national ID card issued 2002-2007		-0.001	
		(0.014)	
2008 winner's clan		0.003	
		(0.021)	
2008 rival's clan		-0.030	
		(0.012)**	
Rural			0.017
			(0.0017)***
Constant	0.017	0.053	0.025
	(0.019)	(0.049)	(0.003)***
Sample	Politician villages		MICS
Observations	3693	3693	123041

Notes: Columns 1 and 2: Politician village sample, 2013; (a) recall from 2007. Column 3: MICS 2011 and 2014. Standard errors in parentheses, clustered at village level. * $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$.

Table A2: Summary statistics by clan

	Politician clans	Others	Difference	SE difference
Female head	0.03478	0.0277	0.00708	(0.00930)
Any daughters currently aged 18-25	0.3526	0.399	-0.0464	(0.0342)
Rudimentary house (lag - 07)	0.2833	0.335	-0.0517	(0.0906)
No ag land	0.608	0.815	-0.207	(0.0691)***
Less than 12.5 acres ag land	0.949	0.984	-0.0350	(0.0143)**
No cattle (lag - 07)	0.594	0.734	-0.140	(0.0462)***
No residential land	0.1282	0.177	-0.0488	(0.0633)
HH head 5th grade or higher	0.5096	0.436	0.0736	(0.0721)
HH head 8th grade or higher	0.38	0.310	0.0700	(0.0634)
HH head 10th grade or higher	0.2612	0.207	0.0542	(0.0558)
Years in village	79.584	77.34	2.244	(3.078)
Years squared	6817.6	6547.7	269.9	(385.6)
HH member received ID (02-07)	0.4902	0.527	-0.0368	(0.0571)
N	547	3146		

Notes: Politician village sample. Same clan is defined as 1 for any household i that is the same clan as the politician from i's village, whether that politician won or lost. Standard errors are clustered at the village level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3: Balance in electoral outcomes in close races

Variable	N	Proportion	P-value H_0: proportion = 0.5
	10% margin		
PML-Q (governing party)	61	0.49	0.90
PPPP	63	0.54	0.53
PML-N	35	0.46	0.61
Independent	34	0.59	0.30
Individual incumbent re-elected	42	0.55	0.54
	5% margin		
PML-Q (governing party)	33	0.48	0.86
PPPP	30	0.7	0.03**
PML-N	21	0.43	0.51
Independent	24	0.54	0.68
Individual incumbent re-elected	26	0.50	1.00
	2.5% margin		
PML-Q (governing party)	22	0.64	0.20
PPPP	12	0.42	0.56
PML-N	12	0.42	0.56
Independent	11	0.55	0.76
Individual incumbent re-elected	17	0.53	0.81

Notes: 2008 election results, Election Commission of Pakistan. Each observation is one MNA constituency.

Table A4: Wealth proxies

	Politician village sample	MICS sample	Overlap with BISP scorecard PMT	Directly observed by enumerator	Unaffected by BISP
Solid house	X	X		X	
Ag land	X	X	X		X
Own dwelling / residential land	X	X			
Head ed 8+	X	X	X		X
BISP eligibility (estimated)		X	X		
MICS wealth index top quintiles		X	X		
Urban		X		X	X
Spouse ed 8+		X			X

Table A5: BISP reform and village favoritism by party

	(1)	(2)	(3)	(4)	(5)
	HH received BISP cash transfer				
Winner's village, majority party: pre-reform	0.005	0.009	0.004	0.005	0.002
	(0.006)	(0.007)	(0.010)	(0.006)	(0.013)
	(0.007)	(0.007)		(0.007)	(0.016)
	0.529	0.392		0.568	0.905
Winner's village, majority party: post-reform	-0.003	0.020	0.007	-0.010	0.012
	(0.016)	(0.016)	(0.010)	(0.018)	(0.026)
	(0.017)	(0.018)		(0.019)	(0.031)
	0.793	0.317		0.823	0.732
Winner's village, other party: pre-reform	0.021	0.019	0.022	0.024	0.027
	(0.010)**	(0.008)**	(0.012)*	(0.013)*	(0.012)**
	(0.011)**	(0.009)**		(0.012)*	(0.014)*
	0.043**	0.015**		0.122	0.090*
Winner's village, other party: post-reform	0.023	0.027	0.022	0.009	0.016
	(0.020)	(0.015)*	(0.015)	(0.033)	(0.022)
	(0.021)	(0.016)*		(0.034)	(0.025)
	0.159	0.066*		0.772	0.475
Majority party, post reform, majority party x post reform	X	X	X	X	X
Vote margin		X			X
Vote margin x winner's village			X		
HH control variables		X	X		X
Additional HH control variables		X			X
HH control variables x post reform		X	X		X
Additional HH control variables x post reform		X			X
Party FE		X			X
Party FE x post reform		X	X		X
Constituency FE			X		
RDD subsample				X	X
Observations	7386	7386	7386	4534	4534
Mean BISP in rival politician villages:					
Majority party winner in office, pre-reform		0.006		0.006	
Minority party winner in office, pre-reform		0.01		0.004	
Majority party winner in office, post-reform		0.03		0.037	
Minority party winner in office, post-reform		0.044		0.055	
P-values:					
Winner's village, majority party = other party, pre reform	0.196	0.250	0.137	0.180	0.209
Winner's village, majority party = other party, post reform	0.316	0.746	0.275	0.620	0.900
Winner's village majority party, pre = post reform	0.505	0.349	0.796	0.260	0.493
Winner's village other party, pre = post reform	0.932	0.676	0.993	0.663	0.670

Notes: Politician village sample. (Parentheses: robust standard errors clustered at the village level.) [Brackets: bootstrapped standard errors clustered at the village level.] {Braces: P-value for cluster wild bootstrap test of $H_0 : \beta = 0$.} Basic and additional household controls are those shown in Column 1 and 2 of Table A1, respectively. * p < 0.1; ** p < .05; *** p < .01.

Table A6: BISP reform and village favoritism by politician seniority

	(1)	(2)	(3)	(4)	(5)
	HH received BISP cash transfer				
Winner's village, senior politician - pre reform	0.007	0.010	0.004	0.010	0.006
	(0.005)	(0.006)*	(0.012)	(0.005)**	(0.008)
	[0.006]	[0.006]		[0.007]	[0.015]
	{0.352}	{0.191}		{0.276}	{0.556}
Winner's village, senior politician - post reform	0.015	0.032	0.023	0.020	0.033
	(0.030)	(0.017)*	(0.012)*	(0.042)	(0.019)*
	[0.032]	[0.023]		[0.038]	[0.030]
	{0.548}	{0.161}		{0.562}	{0.253}
Winner's village, junior politician - pre reform	0.023	0.023	0.038	0.022	0.025
	(0.013)*	(0.009)**	(0.012)***	(0.016)	(0.018)
	[0.012]**	[0.012]**		[0.016]	[0.025]
	{0.090}*	{0.009}***		{0.269}	{0.408}
Winner's village, junior politician - post reform	0.016	0.019	0.024	-0.020	-0.009
	(0.016)	(0.014)	(0.017)	(0.026)	(0.020)
	[0.018]	[0.018]		[0.025]	[0.031]
	{0.405}	{0.344}		{0.331}	{0.524}
Senior politician, post reform, senior x post	X	X	X	X	X
Vote margin		X			X
Vote margin x winner's village			X		
HH control variables		X	X		X
Additional HH control variables		X			X
HH control variables x post reform		X	X		X
Additional HH control variables x post reform		X			X
Party FE		X			X
Party FE x post reform		X	X		X
Constituency FE			X		
RDD subsample				X	X
Observations	7386	7386	7386	4534	4534
Mean BISP in rival politician villages:					
Senior politician in office, pre-reform		0.007			0.002
Junior politician in office, pre-reform		0.010			0.007
Senior politician in office, post-reform		0.040			0.046
Junior politician in office, post-reform		0.039			0.059
P-values:					
Winner's village, senior = junior politician, pre	0.219	0.233	0.033	0.468	0.285
Winner's village, senior = junior politician, post	0.987	0.539	0.962	0.414	0.098*
Winner's village, senior politician, pre = post	0.762	0.208	0.392	0.792	0.173
Winner's village, junior politician, pre = post	0.661	0.782	0.445	0.103	0.181

Notes: Politician village sample. Senior and junior politicians are indicators described in Section 3. (Parentheses: robust standard errors clustered at the village level.) [Brackets: bootstrapped standard errors clustered at the village level.] {Braces: P-value for cluster wild bootstrap test of $H_0 : \beta = 0$.} Basic and additional household controls are those shown in Column 1 and 2 of Table A1, respectively. * p < 0.1; ** p < .05; *** p < .01.

Table A7: Impact of reform on targeting: winner villages only

	(1)
	HH received BISP cash transfer
Male head x post reform	0.094
	(0.073)
	[0.077]
	{0.336}
Solid house x post reform	-0.048
	(0.023)**
	[0.025]*
	{0.171}
Ag land x post reform	-0.032
	(0.015)**
	[0.013]**
	{0.059}
Res land x post reform	0.032
	(0.033)
	[0.035]
	{0.637}
HH head ed 5+ x post reform	-0.012
	(0.020)
	[0.017]
	{0.514}
HH head ed 8+ x post reform	-0.002
	(0.018)
	(0.017]
	{0.942}
HH head ed 10+ x post reform	0.014
	(0.018)
	[0.016]
	{0.497}
Observations	3812

Notes: Difference-in-difference estimates on sample of 2008 winners' villages only. Specification includes controls for all household wealth proxies (male head, education levels, land ownership, and house type), post reform, and party dummies. (Parentheses: robust standard errors clustered at the village level.) [Brackets: bootstrapped standard errors clustered at the village level.] {Braces: P-value for cluster wild bootstrap test of $H_0 : \beta = 0$.}

Table A8: Proxy wealth measures, income and consumption

	(1)	(2)
	Ln(income)	Ln(expenditure)
Eligible per poverty scorecard (approx.)	-0.00493	0.00459
	(0.0177)	(0.0102)
Solid house	0.131***	0.0929***
	(0.0160)	(0.00871)
Residential land	0.00577	-0.00286
	(0.0199)	(0.0124)
Agricultural land	0.539***	0.194***
	(0.0431)	(0.00957)
Rural	-0.0315	-0.0553***
	(0.0245)	(0.0180)
Female head	-0.319***	-0.0600***
	(0.0398)	(0.0143)
HH head ed 5+	-0.148***	-0.112***
	(0.0174)	(0.0141)
HH head ed 8+	-0.166***	-0.124***
	(0.0229)	(0.0166)
HH ed 10+	-0.0717**	-0.0654***
	(0.0286)	(0.0225)
Spouse ed 5+	-0.0175	-0.00977
	(0.0193)	(0.0122)
Spouse ed 8+	0.0574**	0.0482**
	(0.0266)	(0.0202)
Spouse ed 10+	0.178***	0.111***
	(0.0233)	(0.0208)
Electric connection	0.0545	0.0936***
	(0.0357)	(0.0252)
Buffalos and cows/sheep	0.728***	0.309***
	(0.0370)	(0.0145)
Buffalos only	0.555***	0.201***
	(0.0309)	(0.0111)
Cows/sheep only	0.181***	0.0778***
	(0.0311)	(0.0148)
No Buffalos/Cows/Sheep	0.0483	0.0133
	(0.0426)	(0.0226)
Stove	0.774***	0.712***
	(0.0265)	(0.0293)
A/C	-0.00464	0.0350*
	(0.0286)	(0.0190)
Fridge	0.365***	0.349***
	(0.0137)	(0.0127)
Improved toilet	0.177***	0.177***
	(0.0154)	(0.0119)
Basic toilet	0.0584**	0.0879***
	(0.0218)	(0.0197)
Room ratio	-0.00229	-0.0463*
	(0.0316)	(0.0230)
Number of members age 5-16	-0.0122**	-0.000550
	(0.00484)	(0.00303)
Head's education	0.175***	0.142***
	(0.0168)	(0.0140)
Number of members age 18-65	0.0561***	0.0686***
	(0.00394)	(0.00297)
Constant	7.850***	7.687***
	(0.0604)	(0.0258)
Observations	28789	28751
R-squared	0.421	0.414
Adjusted R-squared	0.420	0.414

Notes: Provincially represntative MICS 2003-4 sample only. Standard errors in parentheses clustered at the district level. * p < .1, ** p < .05, *** p < .01.

Table A9: Characteristics of households in winners' villages who did not receive BISP pre-reform

	(1)	(2)
	HHs in winner's village who did not receive BISP pre-reform	
	All	Knows winning politician
Female head	0.025	0.026
Rudimentary house	0.319	0.331
No ag land	0.746	0.721
HH head ed 5+	0.450	0.502
No cattle	0.701	0.697
No residential land	0.144	0.101
No ag land, res land or cattle	0.105	0.062
Knows winning politician	0.635	1.000
Observations	1859	1180

Notes: Politician village sample. Sample means. Column 1 is the sample of households in winners' villages who did not receive BISP before the reform, one observation per household. Column 2 is the subset of the Column 1 sample who also report that they know the winning politician through personal interaction.

Table A10: Misreporting: MICS province-wide data

	(1)	(2)	(3)	(4)	(5)	(6)
	Directly observable by enumerator			Unobserved by enumerator		
	Any part of house solid	House fully solid	Electric connection	Owns ag land	Number rooms	Received remittances
Post reform	0.0132*	-0.0198	0.0319	0.00768	0.0169	0.00978
	(0.00667)	(0.0275)	(0.0810)	(0.0209)	(0.0840)	(0.0115)
Survey wave 2003	-0.0838***	-0.0797***	-0.217***	-0.0322***	0.347***	-0.0124
	(0.0109)	(0.0124)	(0.0380)	(0.00704)	(0.0307)	(0.00898)
Survey wave 2011	0.0119*	-0.0296***	0.407***	-0.0128**	-0.120***	0.0228***
	(0.00592)	(0.0105)	(0.0225)	(0.00491)	(0.0252)	(0.00622)
Survey wave 2014	0.0523***	0.0316	-0.0621	-0.0414*	-0.244**	-0.00742
	(0.0138)	(0.0358)	(0.0810)	(0.0230)	(0.0903)	(0.0129)
Rural	-0.129***	-0.408***	0.0774***	0.311***	-0.129***	0.0327***
	(0.0175)	(0.0190)	(0.00864)	(0.0104)	(0.0170)	(0.00814)
Constant	0.969***	0.821***	1.025***	0.138***	2.160***	0.0781***
	(0.00934)	(0.00996)	(0.00777)	(0.00664)	(0.0187)	(0.00632)
District FE	X	X	X	X	X	X
N	252234	252234	255309	255106	254582	254614

Notes: Provincially representative Multiple Indicator Cluster Survey, 2003, 2007, 2011 and 2014 rounds. Post reform is an indicator variable defined at the district-year level as described in Section 4. Standard errors in parentheses, clustered at the district level. * p < .1, ** p < .05, *** p < .01.

Table A11: Misreporting: politician village sample

	(1)	(2)	(3)	(4)	(5)	(6)
	No res land	No ag land	No cattle	No res land	No ag land	No cattle
Rudimentary house	0.295	0.156	0.108	0.327	0.137	0.096
	(0.108)***	(0.044)***	(0.060)*	(0.103)***	(0.045)***	(0.064)
Rudimentary house x winner's village	0.025	0.036	-0.103			
	(0.171)	(0.058)	(0.088)			
Winner's village	-0.096	-0.071	0.004			
	(0.069)	(0.053)	(0.064)			
Winner's village x winner's clan				-0.069	-0.319	-0.130
				(0.068)	(0.118)**	(0.155)
Winner's village x winner's clan x rud. house				0.193	0.146	-0.188
				(0.152)	(0.113)	(0.142)
Winner's village x other clan				-0.024	-0.040	0.017
				(0.058)	(0.053)	(0.065)
Winner's village x other clan x rud. house				-0.039	0.044	-0.087
				(0.148)	(0.053)	(0.099)
Winner's clan				0.029	-0.039	-0.042
				(0.062)	(0.099)	(0.124)
Rudimentary house x winner's clan				-0.204	0.095	0.163
				(0.166)	(0.091)	(0.112)
N	3693	3693	3693	3693	3693	3693

Notes: Politician village sample, one observation per household. (Parentheses: robust standard errors clustered at the village level.) * p < 0.1; ** p < .05; *** p < .01.

Table A12: Robustness check for recall bias

	(1)	(2)	(3)	(4)
	HH received BISP cash transfer:			
	Ever	Last year	Ever	Last year
Winner's village	0.047	0.025		
	(0.012)***	(0.012)**		
	(0.013)***	(0.014)*		
	0.003***	0.058*		
Winner's village x winner's clan			0.051	-0.005
			(0.024)**	(0.019)
			(0.027)*	(0.020)
			0.157	0.821
Winner's village x other clan			0.047	0.028
			(0.013)***	(0.013)**
			(0.016)***	(0.014)**
			0.001***	0.076*
P-values:				
Winner's village equal, model 1 v. 2	0.005***			
Winner's village, other clan equal, model 3 v. 4			0.016**	
Winner's village, winner's clan equal, model 3 v. 4			0.000***	
Winner's clan	X	X	X	X
HH control variables	X	X	X	X
Additional HH control vvariables	X	X	X	X
Party FE	X	X	X	X
Observations	3693	3693	3693	3693

Notes: Politician village sample, one observation per household. (Parentheses: robust standard errors clustered at the village level.) Basic and additional household controls are those shown in Column 1 and 2 of Table A1, respectively. * p < 0.1; ** p < .05; *** p < .01.

Table A13: Robustness check for endogenous clan reporting

	(1)	(2)	(3)	(4)
	HH reports same clan as politician from village			
Winner's village	0.072	0.057	0.075	0.036
	(0.071)	(0.066)	(0.088)	(0.086)
HH control variables		X		X
RD subset			X	X
Rival village mean			0.097	
Observations	3693	3693	2267	2267

Notes: Politician village sample, one observation per household. (Parentheses: robust standard errors clustered at the village level.) Household controls are those shown in Column 1 of Table A1. * p < 0.1; ** p < .05; *** p < .01.

Table A14: Characteristics of households receiving BISP in one period only

	Received BISP cash transfer: Pre reform only	Post reform only	Difference	SE difference
Female head	0.153	0.072	-0.0813	(0.0536)
Any daughters currently aged 18-25	0.514	0.542	0.0278	(0.0929)
Rudimentary house (lag - 2007)	0.486	0.613	0.127	(0.119)
No agricultural land	0.889	0.935	0.0456	(0.0430)
Less than 12.5 acres agricultural land	0.986	1.000	0.0139	(0.0145)
No cattle (lag - 2007)	0.792	0.786	-0.00595	(0.0728)
No residential land	0.347	0.262	-0.0853	(0.148)
HH head 5th grade or higher	0.278	0.280	0.00198	(0.0887)
HH head 8th grade or higher	0.181	0.197	0.0159	(0.0616)
HH head 10th grade or higher	0.0833	0.131	0.0476	(0.0384)
Years HH has lived in village	67.76	74.883	7.123	(7.620)
Years squared	5463.2	6248.200	785.0	(900.1)
HH member received ID card (02-07)	0.431	0.471	0.0397	(0.122)
Related to elected or local official	0.0278	0.012	-0.0159	(0.0217)
Clan of 2008 winning politician	0.236	0.042	-0.194	(0.0861)**
Clan of 2008 runner-up	0.0417	0.042	7.49e-17	(0.0310)
N	72	168		

Notes: Politician village sample, one observation per household. (Parentheses: robust standard errors from a regression of the household characteristic on a group indicator (pre-reform recipient only or post-reform recipient only), clustered at the village level.) * p < 0.1; ** p < .05; *** p < .01.

Table A15: Substitution with other programs

	(1)	(2)
	Total value of benefits - excluding BISP	
Winner's village, pre-reform	26.144	
	(22.119)	
Winner's village, post-reform	36.641	
	(28.436)	
Winner's village, winner's clan - pre reform		59.731
		(43.255)
Winner's village, winner's clan - post reform		38.128
		(24.789)
Winner's village, other clan - pre reform		23.485
		(22.480)
Winner's village, other clan - post reform		27.783
		(22.651)
Post reform	X	X
Vote margin	X	X
HH control variables	X	X
Additional HH control variables	X	X
HH control variables x post reform	X	X
Party dummies	X	X
Party dummies x post reform	X	X
Rival village mean	191.7	
Rival village, winner clan mean		159.3
Rival village, other clan mean		193.2
N	7386	7386
P-value: Winner's village, pre = post	0.928	
Winner village same clan = other clan, pre		0.393
Winner village same clan = other clan, post		0.707
Winner village same clan, pre = post		0.541
Winner village other clan, pre = post		0.881

Notes: Politician village sample. Value variable is constructed as the approximate total value of all non-BISP targeted assistance programs received, including the girls' stipend, free textbooks, zakat, bait-ul-mal, and sasta rashaan programs. Value is calculated based on administrative data on each program as described in the Pre-Analysis Plan. (Parentheses: robust standard errors clustered at the village level.) Household controls are those shown in Column 2 of Table A1. * $p < 0.1$; ** $p < .05$; *** $p < .01$.

B Changes from pre-analysis plan and robustness to original specifications

In 2013, we filed a non-binding Pre-Analysis Plan on the EGAP pre-registration database. The original plan can be found at http://egap.org/wp-content/uploads/2013/06/20130921-Vyborny-Haseeb-Punjab.pdf

We made the following changes from the PAP:

- Disaggregation showed that the BISP program was driving the aggregate results, because it is much more common than other unconditional cash transfers and by far the largest value program. In addition, the timing of the policy experiment of the switch to the scorecard allowed for a cleaner and better-identified test of differences in program implementation than the comparison between different programs outlined in the PAP. Therefore we focused our analysis on BISP and accordingly on the 2008-2013 period.

- Based on feedback from seminar audiences and reviewers, we eliminate the "two-way switches" from our main analysis sample to focus on the cleaner causal identification provided by the close-election Regression Discontinuity estimates.

- Because of overlapping categories of caste and sub-caste in the Pakistani context and in our data, we combined these variables into a single definition of clan as described in the paper. Based on feedback from seminar audiences and reviewers, we focused on the interaction of clan with the origin village variable rather than analyzing it independently, because the sample is already a sample of same-clan households from *within politicians' origin villages*, so any estimates of the effect of a clan connection with a politician would be specific to the effect within these villages.

- Most of the variation in the control variables was between categories, so we did not use the multivariable fractional polynomial procedure during the covariate selection process. Otherwise the covariate selection was the same as described in the plan.

- We decided to keep analysis of government employment for a separate analysis, because the issues affecting patronage for employment differ substantially from those affecting the allocation of cash and in-kind transfers.

- The PAP omitted one low-incidence government program included in the questionnaire: the Nawaz Sharif fund.

The results of the estimation as originally outlined are consistent with the findings presented in this paper. Table B1 shows these results.

Table B1: PAP specification

	(1)	(2)	(3)	(4)
	OLS Value	Negative binomial Count	OLS Value	Negative binomial Count
Winner's village	913.539 (318.852)***	0.058 (0.072)		
Village of a politician who ever won	203.844 (224.087)	0.003 (0.068)		
Winner's clan			140.024 (260.963)	-0.020 (0.077)
Clan of a politician who ever won			122.809 (400.525)	-0.143 (0.081)*
Round 2 (2008-2013)	1748.897 (415.321)***	0.449 (0.078)***	1306.117 (251.953)***	0.368 (0.111)***
Sample	Villages of winners and rivals		Clans of winners and rivals	
HH control variables	X	X	X	X
Winner's village as control variable			X	X
Winner's clan as control variable	X	X		
Party FE	X	X	X	X
Observations	7426	7426	1590	1590

Notes: Politician village sample; Columns 3-4 estimated for the subsample who share the clan of either the winner or the rival politicians. Value variable is constructed as the approximate total value of all targeted assistance received, including the BISP, new ID cards, girls' stipend, free textbooks, zakat, bait-ul-mal, and sasta rashaan programs. Value is calculated based on administrative data on each program as described in the Pre-Analysis Plan. (Parentheses: robust standard errors clustered at the village level.) Household controls are those shown in Column 2 of Table A1. * $p < 0.1$; ** $p < .05$; *** $p < .01$.