Data, discretion and institutional capacity:

Evidence from cash transfers in Pakistan

Muhammad Haseeb* and Kate Vyborny†‡

October 22, 2020

Abstract

Administrative data is key to many government functions; but generating and maintaining it is costly and challenging in low-income countries. We study an overhaul of public assistance in Pakistan that created a national database of household assets and used the data to means-test cash transfers, eliminating discretion in their allocation. We use difference-in-differences and regression discontinuity approaches to quantify the effect of this reform. Favoritism and transfers to wealthy households dropped; we estimate that the welfare benefits of the reform were seven times as large as its costs. The reform improved public perceptions of social assistance and helped create a robust institution that survived political transitions.

JEL codes: D73, I38, H53

*Post-doctoral fellow, Institute of Economics and Econometrics, Geneva School of Economics and Management, University of Geneva; former Research Fellow, Lahore School of Economics.

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

‡This study was approved by Oxford Central University Research Committee. Analysis was pre-registered with the EGAP registry (20130921AA). We are grateful for guidance from Marcel Fafchamps. We appreciate feedback from Madiha Afzal, Sami Bazzi, Alan de Brauw, Mike Callen, Erica Field, James Fenske, Jenny Guardado, Kaivan Munshi, Julien Labonne, Ben Olken, Simon Quinn, Jake Shapiro, Maximo Torero, Xiao Yu Wang, Xiaoxue Zhao, and Laura Zimmerman, and comments from participants in seminars at Duke, Oxford, CGD, IFPRI, LUMS, AIMS, the World Bank, and the CSAE, RECODE, DIAL, and MWIEDC conferences. We thank Julien Labonne for advising on and carrying out the data split and "testing" estimations, and Naveed Akbar, Saleem Baloch, Ali Cheema, Shujaat Farooq, Haris Gazdar, Ijaz Nabi, Cem Mete, Muhammad Tahir Noor, Sarah Saeed, Khurram Shahzad, and Khaleel Tetlay for help in understanding the program context. We thank Misha Saleem, Amber Nasir, and Abbas Raza for research assistance, and CSAE Oxford, Naved Hamid and the Center for Research on Economics and Business at the Lahore School of Economics, and Shamim Rafique, Sajid Rasul and colleagues at the Punjab Bureau of Statistics for institutional support for the project and data collection. We gratefully acknowledge funding from the British Academy and the Lahore School of Economics.

1 Introduction

Accurate administrative data enables institutional functions such as collection of taxes (Pomeranz, 2015), prevention of crime (Doleac, 2017; Anker *et al.*, 2019), enabling credit markets (Giné *et al.*, 2012), improvement of health care (World Bank, 2018a), enforcement of laws against child marriage (Hanmer and Elefante, 2016), and protection of property rights (Beg, 2020). However, collecting and maintaining comprehensive and accurate administrative data in low-income countries is difficult and costly. In the absence of administrative data, key decisions, such as assessing who owns a particular plot of land, whether a girl is of the age of consent for marriage, or whether a prospective borrower is creditworthy, are often left to the discretion of individual officials.

Administrative data may have particular importance in targeting social assistance programs. Governments have expanded these programs - such as cash transfers - to hundreds of millions of recipients worldwide over the past decades; these programs continue to increase in their size and scope (World Bank, 2018b). In developing countries, where administrative data on income and assets is limited, transfers are often targeted using data that are only proxies for poverty; collecting more detailed information to reduce noise in the proxy is expensive and may be more subject to gaming (Niehaus *et al.*, 2013). In some circumstances economists have argued that in some circumstances such administrative costs of targeting could exceed the benefits (Besley, 1990).

In the absence of centralized administrative data, many low-income countries leave the selection of recipients to the discretion of individual officials. An official with discretion to select recipients might be able to access more information about these households than a central bureaucracy would collect, e.g. through observation or through his social network. He might use this information to select recipients who are poor but would not qualify based on noisy eligibility proxies (Conning and Kevane, 2002; Bardhan and Mookherjee, 2000, 2005; Niehaus *et al.*, 2013; Alatas *et al.*, 2012; Basurto *et al.*, 2019; Alderman, 2002; Hussam *et al.*, 2018). However, a substantial body of evidence demonstrates favoritism by politicians and elites in public expenditure, whether for personal objectives (nepotism, corruption) or political ones (clientelism, vote buying) (Caeyers and Dercon, 2012; Hodler and Raschky, 2014; Burgess *et al.*, 2015; Fafchamps and Labonne, 2017a). Giving officials discretion to select recipients directly could reduce the barriers to favoritism. Thus, the overall effect of replacing targeting based on discretion with targeting based on data is uncertain.

In this study, we address this gap by quantifying the effects of replacing discretion with administrative data in a public assistance program. We study a nationwide overhaul of Pakistan's social assistance program, the Benazir Income Support Program, an unconditional cash transfer to millions of households across the country. Before the reform, elected officials were asked to identify poor households for the cash transfer program and pass a list of their names to the cash transfer agency, which managed the payments. The elected officials were asked to select poor households, but given wide discretion to identify the poor because there was no administrative data available to verify income or wealth.

The World Bank and other donor agencies worked with reformers in the government of Pakistan to create a new national administrative database of households and their assets, and to use it to replace this discretionary targeting system with data-based (proxy means test) targeting. Over this period, the target population, the value of the transfer per household, and payment mechanisms for the transfer remained the same, allowing us to isolate the effects of the introduction of administrative data for targeting the transfer.

We document three benefits of the new data-based system. First, the reform reduced favoritism towards households connected to elected officials. Second, it increased the proportion of the transfer targeted to poorer households. Third, it improved public perceptions of social assistance programs, which may have contributed to the program's continuation and expansion despite two changes in national political leadership.

We define favoritism as the causal effect of a connection between a household and an official on the probability of that household receiving the BISP transfer. Thus to quantify the first benefit, reduced favoritism, we first need to identify how much a connection to an elected official increases the chances of receiving the BISP transfer before the reform, then test whether the reform reduced this advantage. To achieve this, we conduct a novel survey of households in the origin villages of elected officials and a similar comparison group, the origin villages of their rivals (runners-up and previous candidates). These two groups of households are similar in terms of wealth, demographics and geography, and differ only in whether they are connected to an elected official currently holding office. We sample the villages where both groups were born or have ancestral land; this addresses the possibility of reverse causality due to politicians strategically relocating residence. The elected officials in office were the same throughout the period we study in this sample, so selection of officials is also not a concern for identification. We compare the cash transfers received by these two groups of households, before and after the reform.

We demonstrate that replacing discretion with data-based targeting reduced favoritism for households connected to elected officials. Before the reform, households from the origin villages of elected officials were far more likely to receive cash transfers than those from the villages of rival politicians: overall, these households are three times as likely to receive the transfer than those in the rival politicians' villages. Households in the winning politician's village who belong to his clan are over six times as likely to receive the transfer than households in the rival politician's village. After the reform, favoritism for the official's clan in his origin village disappears. The results are robust to alternative specifications including household fixed effects and a Regression Discontinuity design based on close elections (Lee, 2008). They are also robust to out-of-sample testing replicating pre-registered specifications on a reserved "testing" sample.

While the reform reduced favoritism, we find some evidence that favoritism was not completely eliminated. Despite the fact that the reform completely removed elected officials from the cash transfer selection and implementation process, households in their home villages from *other* clans (not the politician's own clan) may continue to have some advantage in receiving cash transfers after the reform. We explore the mechanisms through which favoritism might have continued after the reform. We use administrative data to test whether elected officials might have helped households misreport their assets during the development of the administrative database. We do not find evidence to support this: households in winners' villages have similar poverty scores in the administrative data to those in rivals' villages. Next, we test whether officials influenced the data collection process to ensure that coverage was more comprehensive for potentially eligible households in their home villages. Again, we find no evidence of this: the observed number of households in the administrative data is similar between winners' and rivals' villages, overall and across the wealth distribution. In contrast, we find suggestive evidence that elected officials may have assisted eligible households further downstream in the process, by helping previous recipients to keep their names from being removed from the rolls ("grandfathering"), possibly through the appeals process, or assisting them in meeting the administrative requirements of the program (getting identity cards). In targeting terminology, this suggests that after the reform elected officials may have assisted with resolving exclusion errors in signup, rather than introducing inclusion errors through favoritism as they did before the reform.

The second benefit of the reform was to make the targeting of BISP more pro-poor. To establish this, we use a difference-in-differences strategy with province-wide survey data. We establish that the areas selected for early and late implementation of the reform are similar on baseline observables and pre-trends. We then demonstrate that as the reform rolled out, targeting became more pro-poor across the province. We calculate that the improvements in targeting increase the welfare gains from the cash transfer program by 15%. This benefit was seven times as large as the reported administrative costs of the reform.

We explore whether the improvement in targeting is driven by a pure information mechanism or the removal of discretion. The reform both created new data on household poverty levels and used that data to remove discretion from targeting. The removal of discretion was only possible because of the new data. However, an alternative policy could have been to gather the data, release it to officials so they would have complete information on household poverty, yet still leave them discretion to select recipients. If our results were driven only by officials' lack of information about who was poor, such a policy would also improve targeting. To investigate this possibility, we examine an environment where the official has extensive information: his own home village. We show that the reform stopped transfers to observably wealthier households in the official's own home village. These households have assets that can be observed directly by any neighbor. Conversely, after the reform, transfers began to households in the official's own home village, who are observably poor and report that they know him. In his home village, the official can directly observe these dimensions wealth or poverty of his neighbors, so improved information cannot explain these changes. Thus we conclude that our results are driven by administrative data enabling the central government to eliminate discretion by individual officials, rather than a pure information effect.

The third benefit of the reform was to improve public perceptions of social assistance programs. We document this benefit using differences-in-differences estimates with province-wide data covering the phase-in of the reform. We find that it increased positive perceptions of social protection programs by 10 percentage points, a 40% increase over the baseline mean. Approval of social protection increased not only among households that might benefit from it, but also among the wealthiest households (who were *less* likely to receive benefits after the reform). This pattern is important because targeted social assistance programs typically benefit a small fraction of households; thus to be politically sustainable, they may require support from many non-beneficiaries. In addition, perceptions of the program improved across the political spectrum, with a more pronounced effect in opposition areas. This suggests that the reform helped to broaden support for social protection as a program, beyond households who might expect to benefit from favoritism or patronage based transfers.

This improvement in perceptions across wealth levels and across the political landscape may have contributed to the survival of the cash transfer program through two subsequent political transitions. Observers expected the program might be rolled back after the party that initiated the program lost power (Sarwar, 2018), as historically governments in Pakistan have often de-funded social assistance programs established by their predecessors, and added new programs alongside them. This allowed each successive government to add its political brand onto the new program, staff it from scratch, and potentially influence targeting more easily. Such "electoral cycles" of social protection programs are common in developing countries (Hickey, 2008). This kind of program instability can create a dead weight loss to society beyond mistargeting, because of higher administrative costs in setting up new programs and because income support transfers likely have lower benefits as a "safety net" against shocks if they are unreliable. Breaking with this pattern, each new party after the BISP reform kept the program's features and simply changed its branding. Not only this, they increased funding to the cash transfer program, initiated a round of updating the administrative dataset and expanded its use for other targeted programs. These changes suggest a strengthening of social assistance as an institution in Pakistan.

Our study makes three key contributions. First, we shed light on how concrete investments in government institutions can reduce misallocation of government funds due to favoritism. A large literature quantifies favoritism by politicians and elites in public expenditure, either for personal objectives (nepotism, corruption) or political ones (clientelism, vote buying) (Caeyers and Dercon, 2012; Hsieh *et al.*, 2011; Besley *et al.*, 2004, 2012; Carozzi and Repetto, 2014; Nguyen *et al.*, 2017; Mu and Zhang, 2011; Kitschelt and Wilkinson, 2007; Weitz-Shapiro, 2012; Stokes *et al.*, 2013; Briggs, 2014; Jablonski, 2014; Jayne *et al.*, 2001; Öhler and Nunnenkamp, 2014; Kilic *et al.*, 2015; Bardhan and Mookherjee, 2006; Galasso and Ravallion, 2005; Hodler and Raschky, 2014; Burgess *et al.*, 2015; Fafchamps and Labonne, 2017a). Some authors have argued that targeted programs are particularly vulnerable to capture or clientelism (Keefer and Vlaicu, 2007; Keefer, 2007). However, fewer studies have investigated how program design may reduce favoritism. While some interventions that reduce leakage (i.e. diversion of benefits to non-targeted recipients) have shown large benefits (Banerjee *et al.*, 2017a; Muralidharan *et al.*, 2016; Barnwal, 2019), existing research studying interventions intended to reduce favoritism in the *selection* of targeted recipients have found a limited impact on welfare, because favoritism is limited in scope or because well-connected households are not substantially wealthier than others (Alatas *et al.*, 2012, 2019; Basurto *et al.*, 2019). However, this could be context specific: for example, in a setting in which there is a high level of economic inequality between those who are connected and unconnected, favoritism may have a larger impact on welfare. In contrast to these studies, we find both evidence of substantial favoritism under discretion, as well as large benefits from eliminating discretion. In addition, we provide evidence on the potential mechanisms for the persistence of favoritism despite non-discretionary targeting. Previous work has documented the potential for reforms that eliminate discretion to fail due to manipulation of eligibility data or formulas (Banful, 2011; Camacho and Conover, 2011). In our setting, we rule out the possibility of data manipulation as a mechanism for the continued favoritism we observe for elected officials' villages. However, we find suggestive evidence that continued favoritism may have operated through assistance in overcoming administrative hurdles such as obtaining an ID card or launching an appeal. This demonstrates the importance of administrative barriers to takeup of social assistance programs. Previous work has explored the extent to which such barriers lead to exclusion error in social programs (Currie, 2006; Coady *et al.*, 2004; Besley, 1990; Kleven and Kopczuk, 2011), but has also highlighted that they can act as "ordeal mechanisms" which improve targeting through revealed preferences (Alatas *et al.*, 2016; Dupas *et al.*, 2016). However, our results suggest that administrative barriers may lead to advantages for the well-connected, even when eligibility itself is not discretionary. Unlike existing work, which focuses on how these barriers affect household self-selection, our results suggest that such administrative barriers can lead to effective favoritism despite the official elimination of discretion from targeting rules.

Second, we document the impact of investing in the institutional capacity of a social program on public perceptions of the program. Perceived fairness in distribution may be important in determining political support for social protection (Pritchett, 2005). However, some have argued that in contexts with weak institutions, discretionary, clientelistic transfers may be the political equilibrium; and voters may in fact see capture or clientelism as legitimate, helping to bring additional resources to a constituency (Diaz-Cayeros *et al.*, 2016; Keefer and Vlaicu, 2007; Wade, 1985; Platteau, 2004; Anderson *et al.*, 2015; Bardhan *et al.*, 2009; Bardhan and Mitra, 2014; Kitschelt and Wilkinson, 2007; Stokes *et al.*, 2013). A smaller strand of research has documented voter preferences against clientelism using survey experiments or variation across households in the type of assistance received (Banerjee *et al.*, 2014; Bardhan *et al.*, 2014). Several studies (De La O, 2013; Labonne, 2013; Imai *et al.*, 2020) test causally for political impacts of specific non-discretionary transfer programs. However, these estimates include the effect of support for government spending on such a program overall, as well as for the quality of its targeting and implementation. Kosec and Mo (2019) use a survey experiment to show that when surveyors make the respondents' poverty salient, poor households who are ineligible for transfers reduce their stated support for the government; this could be because the salience treatment leads respondents to conclude that targeting is unfair, or simply because it intensifies their overall frustration at not receiving support. Because we study a reform rolled out within a cash transfer program, we are able to separately identify the effect of a change in *program design* on citizen perceptions. To our knowledge, the only prior study to use within-program variation in design to study effects on perceptions of the fairness of targeting in a transfer program is Alatas *et al.* (2012), who find that households approve more of targeting outcomes under discretion than under a formula-based (proxy means test) approach. They attribute this to a difference between the indicators targeted in the formula and measures that community elites exerting discretion use to target transfers, which match community members' definition of poverty. Again our results contrast to theirs: we find that the BISP reform that generated administrative data for targeting and removed discretion, in addition to improving targeting and reducing favoritism, substantially improved public approval including among wealthy households and opposition party strongholds.

Third, we contribute to a broader literature exploring whether deliberate attempts can improve the performance of institutions (Acemoglu and Robinson, 2010; Banerjee and Duflo, 2014). There is growing evidence of persistence in institutions (Acemoglu *et al.*, 2001, 2014, 2013; Dell, 2010; Banerjee and Iyer, 2005). However, there is limited evidence on whether and how deliberate reform attempts can improve institutions (Banerjee and Duflo, 2014). We show that a deliberate effort by international aid donors and the national government to reform the mechanism for targeting public assistance in Pakistan improved not only program performance but also public perceptions of the program. Observers have expressed concern that aid may weaken institutions (Moss *et al.*, 2006; Djankov *et al.*, 2008; Bräutigam and Knack, 2004); we provide evidence that advocacy and technical assistance from international aid donors effectively strengthened the institution of social protection in Pakistan.

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

2 Context

Social assistance in Pakistan and BISP

Prior to 2008, Pakistan had a patchwork of small, overlapping social assistance programs (World Bank, 2007). These programs all used some form of discretionary or informal targeting. Changes in the party in power would result in these programs being disrupted completely, paused or defunded (Clark, 2001; Sarwar, 2018). In 2008, the newly elected government of the Pakistan People's Party (PPP) worked with aid donors including the UK aid agency (DfID) and the World Bank to set up a major cash transfer program. The government named the program the Benazir Income Support Program (BISP) in memory of the PPP's leader, who had recently been assassinated.

BISP is an unconditional transfer targeted at poor households. Once selected, recipients are supposed to receive a set amount quarterly; initially this was 3000 PKR per quarter, later increased periodically to account for inflation. The program has 7 million recipient households, and has paid out $4 billion, making it one of the largest social protection programs in the world (Nabi, 2013; Cheema *et al.*, 2015; World Bank, 2013).

BISP reform

The aid donors advocated that the government develop an administrative database to allow a formal system to select recipients for the transfer based on a non-discretionary formula (a proxy means test or "poverty scorecard"). Some key decisionmakers in the party's senior leadership were also anxious to ensure that this flagship program of the new government, branded with the name of the party's late leader, avoid scandals of corruption or favoritism, and agreed to implement this new system. The donors designed a short census of all potentially eligible households including a series of simple wealth proxies,[1] and a formula for converting these assets into a "poverty score" to predict household consumption (Hou, 2011).

The government agreed to use a system that the aid donors would design, but were in a hurry to start distributing funds. They initiated the cash transfer program before the new targeting mechanism was ready, and instead made elected officials responsible for nominating recipients in the interim. Each official was given 8,000 nomination forms to sign up beneficiaries from his or her constituency.

By fixing a quota for each elected official, including opposition party officials, the government aimed to avoid criticism of favoritism. However, officials still had a wide choice of recipients. They were asked to select the poor, but given minimal objective criteria based on what could be verified in the limited existing national administrative data: the recipients should not appear on the national list of individuals with a machine readable passport, an ID card for emigrants, an account with a foreign-owned bank, or any household member who is a government employee. Given that most Pakistani households meet these criteria, this gave the officials broad discretion, similar to previous social support programs. Even so, the officials nominated many recipients who were ineligible under these minimal criteria. Fifty percent of the officials' original nominees were disqualified without receiving the transfer, according to BISP officials. Even though these households were eliminated from the recipient list, Nayab and Farooq (2012) find that many who did receive BISP in this period were ineligible on the basis of these criteria. Researchers and donors expressed concern that officials were selecting 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 development of the administrative database to allow formula based targeting. The census organization, contracted NGOs and survey firms were all involved in gathering the data through a nationwide door-to-door listing with a short checklist of assets for each household. They then submitted the forms to the agency responsible for national ID cards, NADRA. NADRA calculated the "poverty score" for each household, using weights that were developed with the assistance of the aid donors. These weights were selected to predict household consumption from the simple asset checklist. The weights were not released even to most BISP officials. After the data were collected, a cutoff value was set based on a target proportion of households set based on the program's budget. Households with a score below the national cutoff value were deemed eligible for the BISP cash transfer.

[1] 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 detailed administrative data on poverty were kept in a national database in Islamabad, the National Socio-Economic Registry (NSER). Only the summary scores and eligibility status were communicated to field units of the cash transfer agency to implement the transfer. Officials estimate the cost of implementing the reform, including collecting and checking the national survey and processing it to identify recipients, at $52 million USD, or less than 2% of the nearly $4 billion paid out to recipients until 2016, when the government initiated an update of the administrative data. This is similar to the proportion of the total transfer budget spent on collecting administrative data for formula based targeting (proxy means tests) in other countries (Hanna and Olken, 2018).

This process started in 2009 in 15 pilot districts (out of 106 total in the country), including four districts in Punjab province. These districts were selected by BISP bureaucrats to cover a range of geographic areas, urban and rural, and areas with different poverty levels. The collection of information for the administrative database was completed in these pilot districts in June 2010; starting from July, BISP transfers were initiated to newly selected recipients and discontinued to ineligible recipients. In the remaining districts, the data collection was completed in June 2011 and transfers were initiated starting from July 2011.

Recipients selected under the old system who did not qualify under the new criteria were removed from the list, notified by a letter, and their payments stopped; this amounted to 75% of the pre-reform recipients.

Households may appeal to be added to the beneficiary list if not selected, and if their poverty score is above the cutoff of 16.17 but below 20 they are automatically re-enrolled as long as they met a second, less stringent, set of objective criteria.[2] According to BISP officials, as of 2014, approximately 600,000 households had filed appeals and approximately 35% of these were then enrolled.

Other than disqualification due to the reform, the only way for a recipient to be disenrolled from the program in the period we study was through the system's "grievance redressal mechanism," which functions through a national hotline and case management offices in the field.[3] In principle, anyone can report an ineligible household receiving BISP, and these offices are responsible for re-verifying assets through a household visit. In practice, BISP officials indicate that reports of this kind are extremely uncommon, so when households report they stopped receiving BISP, we attribute this to the reform. In 2016, after our data collection was complete, the government and donors began a new process to update the NSER database.

During the time period we study, the intended target group, size and unconditional nature of the transfer did not change. The agencies managing the program, the political environment and elected officials in office were all the same before and after the reform. This allows us to compare the outcomes over that time period

[2]If they were between 16.17 and 20 points on the BISP "poverty score" formula, and either (a) household size less than 3; (b) 4 or more children under 12 years old, (c) had any member above 65 years, or (d) any member with a disability.

[3]There is one office per *tehsil*; in Punjab, this means approximately 160 offices covering a population of 100 million households.

and isolate the effect of the reform.

Potential limitations of the reform

Even with a formal beneficiary selection process, there could still be scope for continued favoritism in BISP targeting. First, elected officials might be able to influence the administrative data itself (Litschig, 2012; Niehaus *et al.*, 2013; Banful, 2011; Camacho and Conover, 2011)). They might directly influence the organizations carrying out the poverty score card data collection. They might also influence bureaucrats in Islamabad to manipulate the data or formula, or to add recipients despite the fact that they do not meet the cutoff (Alatas *et al.*, 2019).[4] Second, elected officials could also give extra help to those in their home villages to overcome administrative hurdles in getting the transfer. Households are not required to apply for BISP, but once selected a female household member must get a National Identity Card to complete enrollment. Households in rural Punjab report assistance with getting an ID card as one of the most common forms of direct constituent assistance that politicians provide (Chaudhry and Vyborny, 2013). Elected officials might also assist households with the appeals process. This was not designed to be discretionary. However, the household had to initiate the appeal, leaving scope for officials to give households assistance in the process. In addition, BISP sent a letter to households who had been selected by the elected officials and were disenrolled under the reform; the letter informed them they had been disenrolled and also explained the appeals process, leaving scope for some households to be "grandfathered" in through successful appeals.

3 Data

We use the rollout of the BISP reform to study the effects of developing and using administrative data on transfer targeting, favoritism, and public perceptions of social assistance. We use three main data sources.

Home village survey

We conduct a household survey in the origin villages of Members of the National Assembly (MNAs), officials who are directly elected to represent a constituency of approximately 300,000 voters, and their opponents. We use election data compiled by Fair *et al.* (2017) to select a random sample of constituencies that are competitive in one of two ways: (1) they had a close outcome in 2008 (5% vote margin), and/or (2) the top two candidates exchanged places between 2002 and 2008. We eliminate constituencies in which the 2008 winner and runner-up were of the same clan, to identify effects on households in the same clan as the elected official. For each constituency, we identify the 2008 winner, who was in office and could select BISP

[4] 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.

recipients pre-reform. We also identify his/her rivals: the 2008 runner-up, the 2002 winner and the 2002 runner-up.[5] Since the top candidates often contest successive elections, each constituency has between one and three rivals. The survey fieldwork took place in February 2013, two months before a national election; thus the same elected officials were in office throughout the period covered in the survey data (2008-2013).

We use desk research and informed contacts to identify the villages where both groups were born or have ancestral land; this addresses the possibility of reverse causality due to politicians strategically relocating residence. We identify and survey an origin village for 36 out of a total of 53 candidates in all 19 constituencies.[6] We survey approximately 8,000 households across these 36 villages.

We use a detailed recall question to measure transfers received in each year from the inception of the program, through the reform, until the survey date, which was just before the next election. For estimation, we aggregate the data on benefits received in each year into two periods, pre and post reform, as outlined in Bertrand *et al.* (2004). We address potential recall bias in three ways. First, we focus on the binary outcome of receiving the transfer, not the amount or installments received. BISP is a well-known program, so receiving it is a salient event; respondents readily recalled it in pilot surveys. Second, each respondent is first asked whether his/her household has ever received BISP. If the answer was "yes," the enumerator uses a series of time reference points to identify the years in which they received the program to reduce recall bias (Loftus and Marburger, 1983). Third, 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- (post-) reform if they reported receiving it in 2008 or 2009 (2012 or 2013): in other words, we ignore BISP reported received in 2010 or 2011 in constructing these variables. [7] The survey also included similar questions on other transfer programs.

We also collect more detailed data on the household's connections to elected officials. First, we identify whether the household is of the same caste / clan ("biradari" / "zaat") as the elected official. Enumerators ask households to identify their clan at the beginning of the survey, in a question not framed with reference

[5]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.

[6]Cases in which we do not identify an origin village for a candidate could occur because the candidate came from a family that has been based in an urban area for several generations. They could also occur if there was such a village but we were unable to find the information through our sources. This could bias our results if it was driven by less prominent officials, who might be less likely to attract resources to their village or to find ways around the reform. However, this should bias our estimates towards zero: if anything, the least powerful and prominent rivals should bring fewer transfers into their villages, so if we do not observe them, this should decrease the gap between winners and rivals. The candidates for whom we did not identify a village were roughly split between the 2008 winners and rivals (10 winners and 7 rivals), which does not suggest under-representation of weaker candidates. We present constituency fixed-effects estimations in our main results tables, so that the identifying variation comes from constituencies where both a winner and rival village were identified.

[7]Because the BISP transfer is paid out quarterly to recipients from enrollment until disenrollment, respondents report receiving BISP for a continuous series of years. Thus, focusing on the early and late years in the recall period does not omit recipients, but rather prevents us from erroneously classifying households as receiving BISP pre-reform if they reported their start year with some error. Similarly, it prevent us from erroneously classifying households as receiving BISP post-reform if they reported their end year with some error, or they were disenrolled with some delay.

to the elected official and asked before politics, the BISP transfer or any related topics are discussed. We match the clan of the elected officials as reported by local village officials with the household's reported clan.[8] At the end of the survey, the enumerator also asks the respondent about whether he/she knows and/or has met the elected official and his/her rivals in that constituency.

To address the possibility of strategic migration by households seeking assistance from elected officials, we collect migration history for each household, and drop all 1,066 households who have moved to the village at any time since the 2002 election from our main estimation sample.[9]

To tie our hands against specification search in our primary data (Casey *et al.*, 2012), we use a novel combination of approaches in the analysis of our home village data. We registered an initial *non-binding* pre-analysis plan with the Experiments in Governance and Politics registry (Humphreys *et al.*, 2013). Second, we adapted the "sample split" method proposed by Fafchamps and Labonne (2017b) and Anderson and Magruder (2017); a third party researcher split our data into two samples, gave us half and kept half. Using the first, "training," half of the sample, we conducted the initial analysis and made changes to incorporate feedback. We registered the full set of tables from the estimations on the "training" sample as a final component to our pre-analysis plan on our EGAP registration. The third party researcher then repeated our estimations on the full sample and on the "testing" half as a robustness check. This combination approach allows more flexibility in departures from our initial, non-binding pre-analysis plan, addressing concerns about limiting learning from the data during analysis (Deaton, 2012; Humphreys *et al.*, 2013; Olken, 2015). We present the results from the full sample; the results from the training and testing samples are included in Web Appendix C, and the results from the preliminary PAP specification in Web Appendix D.

Multiple Indicator Cluster Survey

The second dataset is the provincially representative Punjab Multiple Indicator Cluster Survey, collected by the Punjab government and UNICEF. We use MICS cross sections from 2003, 2007, 2011, and 2014. The questionnaire covers BISP and other government transfer programs received in the last year, as well as household assets and income. The MICS also includes a standardized wealth index based on a list of household assets, calculated by the Bureau of Statistics using the methodology in Filmer and Pritchett (2001).[10] We also use the wealth variables in the MICS to estimate an approximation of the official "poverty

[8] There is some evidence that households' caste/clan identification may change over a period of generations (Cassan, 2015). For such changes to affect our estimates, caste identity would have to adjust in response to each electoral cycle, which seems unlikely. To confirm this, we test whether in our home village sample, households are more likely to report the same clan as the candidate from their village if he won (Table 1, first row). We find no evidence of this.

[9] Because our data is retrospective, we cannot identify households that might have left the village during this period. However, for differential out-migration to drive our findings, households who are more likely to receive BISP (for reasons unrelated to favoritism) when the winner selects recipients *and* then be disqualified from BISP after the reform would have to be more likely to leave rivals' villages; this seems unlikely.

[10] The assets used in the index include: number of rooms for sleeping per member; material used for floor, roof and wall

score" formula used to target BISP. Starting in 2007, the MICS also included a question on approval of social safety net programs (whether the respondent perceives transfer programs as "beneficial to the common man").

Administrative data

The third dataset is the Punjab province portion of the administrative dataset that NADRA used for constructing the "poverty score". It provides household-level information on the assets used in the poverty score: dwelling characteristics, livestock, and ownership of land.

4 Empirical Strategy and Results

Section 4.1 presents our empirical strategy and results on how the reform affected favoritism in targeting. Section 4.2 then quantifies the extent to which the reform made targeting more pro-poor overall. We examine mechanisms for the improvements in targeting (4.3), and then for the remaining favoritism (4.4). Section 4.5 investigates how the reform changed public perceptions of social assistance. Section (4.6) tests for evidence of potential unintended consequences of the reform. Finally, we present estimates of the welfare effects of improved targeting quality (4.7).

4.1 Favoritism

We use our home village survey to examine the effect of the reform on favoritism. To quantify favoritism, we need a causal estimate of the effect of a connection to an official on receiving cash transfers. However, elected officials come from socially and economically privileged groups (Table A1), so households connected to them are likely to be wealthier ex ante and less likely to need or qualify for assistance. To identify the effect of the connection with an official on receipt of the transfer, we compare households in an elected official's village with those in the villages of his rivals. Villages connected to the official and his rivals within a constituency are likely to be similar, other than their access to the official who won office. We test whether households in the origin villages of winners and rivals in our sample differ (Table 1, Columns 1-6). The two groups are similar, although households in the winners' villages appear to be more likely to own land. If anything, this suggests that they are slightly *less* likely to be eligible for the cash transfer, which would bias us against finding a favoritism effect.

of dwelling; type of cooking fuel; electricity; gas; radio; television; cable television; mobile and non-mobile phone; computer; internet access; refrigerator; air conditioner; washing machine; cooler; microwave; sewing machine; iron; water filter; motorised pump; watch; bicycle; motorcycle/scooter; animal-drawn cart; car or truck; source of drinking water and type of sanitation facility.

Similarly, we identify households linked to winners through clan. We divide households in the winner's village into two groups: the winner's clan, and all other clans. In each case, we will compare the incidence of BISP with the same group of clans in the rival's village.

Figure 1 shows the incidence of BISP before the reform. Households in the winner's village are approximately three times as likely to receive BISP as those in the rival's village (panel A). Households in the winner's village in the winner's clan are by far the most likely to receive BISP pre reform (panel B), despite the fact that they are typically wealthier than other clans (Table A1). In contrast, none of the households from the winner's clan in the rival's village receive BISP pre reform. The pattern is similar but less pronounced for other clans.

We formally estimate home-village favoritism for each of the two groups (winner's clan and other clans) and test how the reform changes this pattern in a difference-in-differences framework.[11] We estimate the following linear probability model:

$$\begin{aligned} B_{ict} =& \alpha + \beta_1 WinVil_i \times WinClan_i \times PreReform_t + \beta_2 WinVil_i \times OtherClan_i \times PreReform_t \\ &+ \beta_3 WinVil_i \times WinClan_i \times PostReform_t + \beta_4 WinVil_i \times OtherClan_i \times PostReform_t \\ &+ \gamma PostReform_t + \xi WinClan_i + \phi WinClan_i \times PostReform_t + \delta X_i + \eta Z_c + u_i + \epsilon_{ict} \end{aligned} \tag{1}$$

Where $B_{ict} = 1$ if household i in constituency c received BISP transfers during time period t (pre / post reform). $WinVil$ is a dummy variable for the winner's origin village, $WinClan$ is a dummy for the winner's clan, and $OtherClan$ a dummy for all other clans. X and Z are vectors of household- and constituency-level observables, including the 2008 vote margin. $PreReform$ and $PostReform$ are dummies for before and after the BISP reform. u_i is a household-level idiosyncratic term; we estimate a household-level random effects model to increase efficiency (Wooldridge, 2002), and show that the results are robust to OLS and household fixed effects specifications in the appendix. We cluster standard errors at the village level. We also present bootstrapped standard errors clustered at the village level and Cameron *et al.* (2008) wild-cluster bootstrapped standard errors clustered by constituency.

β_1 compares BISP received by households of the winner's clan in the winner's village to households of the winner's clan in the rival's village, pre reform. Thus β_1 is the pre-reform "village favoritism" effect for the winner's clan. The four categories $WinClan \times PreReform$, $OtherClan \times PreReform$, $WinClan \times PostReform$, and $OtherClan \times PostReform$ are mutually exclusive and exhaustive; thus each of the individual terms

[11] 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 on precise recall of years in which the household received the transfer.

$\beta_1 - \beta_4$ can be interpreted as village favoritism for the specified group. That is, β_2 is the village favoritism effect for households in other clans pre-reform; β_3 is the village favoritism effect for the winner's clan post-reform; and β_4 is the village favoritism effect for other clans post reform.

We test for a decrease in favoritism for each group, $H_0 : \beta_1 = \beta_3$ and $H_0 : \beta_2 = \beta_4$. The entire period we study occurred during one electoral term, so the same elected officials were in office throughout. In addition, most aspects of the program besides the reform stayed constant. The only other change over this period besides the elimination of discretion was an increase in the number of recipients of the program. Since the number of recipients increases overall after the reform, our tests for a decrease in the level of favoritism are conservative. For example, if BISP increased from 2% to 5% of households in rivals' villages, and from 4% to 7% in winners' villages, there would be no decrease in the estimated level of favoritism (two percentage points more for winners' villages) despite the fact that BISP coverage would be converging between the two groups.

We include close-elections regression discontinuity estimates with the subsample of 11 constituencies with less than a 5% margin in 2008 (Lee, 2008) alongside our main estimates. These estimates include only the villages of the 2008 winner (treatment) and runner-up (control) in each constituency in estimating 1.[12]

Table 2 shows the results of Equation 1. Columns 1-3 show the main estimates, while columns 4-5 show the close-elections specification. Echoing the pattern shown in Figure 1, the estimates show that at the inception of the program, when elected officials selected recipients, households from the winner's village were dramatically more likely to receive BISP. All clans benefited from having an official from their village elected, but the estimates are largest for the official's clan: they are three percentage points more likely to receive the BISP cash transfer. Pre-reform, other clans also benefit from having an official from their village elected; the point estimates are sometimes smaller than those for the winner's clan, but not significantly different. In contrast, in the rival's villages, none of the households of the winner's clan and 1% of other households receive BISP. Overall, before the reform, households in the winner's village are approximately three times as likely to receive BISP as those in the rival's village. Households in the winner's village in the winner's clan are six times as likely to receive BISP as an average household in the rival's village.[13]

After the reform, favoritism disappeared for the winner's clan in his own village: our estimates of fa-

[12]To assess the validity of this strategy in our sample, we test for potential manipulation of the running variable using the test proposed by McCrary (2008), and find no evidence of any party being more likely to win in close races (Web appendix Figure B1). Recent work on U.S. elections (Sekhon and Caughey, 2011; Grimmer *et al.*, 2011) shows just-winners differ from just-losers, biasing RDD estimates. However, for this to occur, a candidate must closely monitor projected voting outcomes and allocate resources to achieve a vote count just above that of his competitor and no more (cf. Simpser (2013); Gehlbach and Simpser (2015)), which seems unlikely given limited information here (cf. Eggers *et al.* (2013)).

[13]Table A3 shows a version of Equation 1 without the clan interaction, and Web Appendix Table B1 shows a similar estimate incorporating neighboring villages into the sample; the pattern of results is consistent with the results shown in Table 2.

voritism for this group post reform, $\hat{\beta}_3$, are close to zero. We can reject $\beta_1 = \beta_3$ across specifications and alternative approaches to clustering. The overall number of these households receiving BISP actually decreased significantly after the reform, despite the fact that more households received BISP overall. The reform successfully limited favoritism for the households with the closest connection to the elected official. The results are robust to pooled OLS and household fixed effects specifications (Table A2), as well as to out-of-sample testing (Web Appendix C). They are also not driven by powerful or senior officials (Web appendix Table B2).

The post-reform results in columns 1-3 suggest that households in the winner's village but from *other* clans continued to benefit from their connection with him, although these results are not robust to the close-elections sample (columns 4-5).[14] These households are poorer on average and more likely to be eligible than the official's clan (Table A1), so this may suggest that after the reform officials provide their *eligible* connections with assistance in getting access to the program; we explore this further in Section 4.4.

Alternative explanations for home village results

Saturation: The increase in the number of BISP households might have affected the composition of recipients. For example, if the post-reform targeting scheme covered all the elected official's preferred recipients as well as other households in rivals' villages, this could appear to eliminate favoritism, when it only made it unnecessary for the official to intervene to ensure his preferred recipients were targeted. However, this is not the case: 50% of officials' initial nominations were disqualified under the original minimal criteria and 75% of the remainder were disqualified through the reform; thus such a "saturation" effect should not drive our results.

Anticipation: If elected officials anticipated the change in targeting in their initial selection, they might have chosen more individuals they preferred, but anticipated would be disenrolled later, to ensure that these individuals received at least some funds. However, it seems unlikely that officials believed that any such rules would be enforced, given that 50% of initial nominations were rejected even based on the minimal criteria used in the pre-reform period (described in detail in Section 2).

Differential recall bias: Our measure of BISP in the home village survey is recall-based. Recall bias could affect our estimates if involvement of elected officials in BISP distribution is more salient to households in their own villages, and these households remember earlier BISP transfers more readily. Section 3 describes the survey methodology we used to minimize recall bias. As discussed in Section 4, we drop observations of BISP receipt immediately before and after the reform to reduce the effect of recall error on our estimates. In

[14] Table A3 shows the effect of a connection with the winner's village, pooled for households of all clans. Pooling masks the heterogeneity between clans: there is clear evidence of favoritism before the reform, but because the reduction in favoritism is driven by the official's clan, it is not significant in the pooled specification.

addition, we test whether our results hold when using a different construction of the dependent variable which does not rely on reported dates of receipt of the transfer. We estimate a cross-section version of our main specification, using as the dependent variable whether the respondent answered "yes" to the question "Has your household ever received BISP?". Since the transfer is well known by name and is a substantial source of income to low-income recipient households, 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. The patterns mirror our main results (Table A4).

4.2 Targeting the poor

The BISP reform reduced favoritism towards the households best connected to officials in power. But did it shift allocation of the transfer towards poorer households? Under discretionary targeting before the reform, officials could have used local knowledge to select recipients who were poor, but who might not appear to be poor based on the indicators in the formula (Alatas *et al.*, 2012; Alderman, 2002). If this was the case, we might see an improvement in targeting based on the formula indicators but no similar effect in other indicators. Thus, we assess the poverty targeting of BISP based on two sets of indicators. The first set are indicators that overlap with the new targeting formula. This allows us to assess whether the new targeting rules were followed in practice, and whether they were binding. The second set includes indicators which are *not* in the targeting formula, but predict household income in pre-BISP survey data (Web Appendix Table B3). This allows us to assess whether formula-based targeting effectively improved allocation more broadly.

We use the MICS repeated cross-section data, which covers the province of Punjab. We exploit the rollout of the BISP reform district by district. 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 verify that these districts appear similar to the rest of the province, and in particular and have similar pre-trends to other districts on economic and political variables and the targeting of pre-BISP transfers (Figure 2, Table 1).

We estimate:

$$\begin{aligned} B_{idt} = {} & \beta_0 + \beta_1 Poor_{idt} + \beta_2 Y2014_t + \beta_3 Poor_{idt} \times Y2014_t + \beta_4 Pilot_d \times Poor_{idt} \\ & + \beta_5 IMPL_{dt} + \beta_6 IMPL_{idt} Poor_{idt} + \alpha_d + \epsilon_{idt} \end{aligned} \tag{2}$$

Where B is a dummy for household i in district d receiving BISP. $Poor$ is a binary poverty proxy. $Pilot$ is a dummy for the pilot districts in which the reform was rolled out first, and α are district fixed effects. $IMPL$ is a dummy which takes value 1 if the reform has been implemented in district d at time t. We estimate 2 using two rounds (2011 and 2014) of MICS data, i.e. all the data collected after the BISP program had begun. In 2011, the reform had been implemented in pilot districts only, and in 2014, it had been implemented in all districts; thus $IMPL = 1$ for pilot districts in year 2011 and 0 for all other districts in 2011; $IMPL = 1$ for all districts in 2014. Thus 2 is a reverse difference-in-differences estimator capturing the effect of the district-phase-in.

Because the rollout of the reform also increased the number of households targeted, we examine how BISP was distributed among (relatively) wealthy and poor households as it was scaled up. β_1 captures the extent to which the transfer was targeted towards poor households pre-reform, β_2 an overall increase in BISP distribution, and β_3 any overall improvement in targeting over time not driven by the reform. β_4 captures any differences in baseline targeting between pilot and non-pilot districts. β_5 is our estimate of how much the reform increased BISP distribution to wealthy households, and β_5 the additional increase for poor households.[15]

Table 3 shows the results. Each column shows the same specification with a different wealth proxy. Column 1 shows the results using the top two quintiles of the MICS wealth index as the wealth proxy: households in the top two quintiles are categorized as wealthy, while the rest are poor. By this measure, at baseline in the control group, approximately 2% of wealthy households and 4% of poor households received the BISP transfer. The reform led to a 5 percentage point increase in the probability of a poor household receiving BISP, and no detectable increase for a wealthy household. Column 2 examines targeting with a more restrictive wealth criterion: only households in the top quintile of the wealth index are considered wealthy. Here we see that the reform actually reduced the probability of a wealthy household receiving BISP to zero, while increasing the probability of a poor household receiving BISP increased by about 3 percentage points (4.8 - 1.5).

A similar pattern holds for most of the wealth indicators. Thus the reform led to a substantial improvement in targeting - a decrease in the proportion of recipients who are wealthy, or "inclusion error". This pattern holds for key poverty indicators that are *not* included in the official poverty score formula, the physical characteristics of the dwelling (solid roof, walls and floor). This pattern of results suggests that the

[15] Receiving the cash transfer might increase household wealth measures. However, note that (a) the amount of the transfer over the duration in the period we are studying would be unlikely to be sufficient to shift the wealth indicators we use here; (b) if this did take place, this would bias us towards a finding of worse targeting, as households receiving BISP would be more likely to appear wealthy based on the assets they acquire.

reform improved targeting broadly, rather than only ensuring it conformed to the specific formula used.

These estimates could be biased if respondents misreport their assets in response to learning about BISP's formula based targeting, in the hope that the survey will make them eligible for BISP. If so, we could see an apparent improvement in targeting quality due to differential strategic misreporting. However, the pattern persists for dwelling characteristics, which are directly observed by the survey enumerator from outside the home, helping to allay this concern. We also test for evidence that the reform led to differential misreporting in household surveys in Section 4.6, and find no evidence of this.

4.3 Mechanisms for improved targeting

The reform both created new administrative data on household assets and used that data to remove discretion from targeting. The removal of discretion was only possible because of the new administrative database, which provided the information needed for formula-based targeting. However, an alternative policy could have been to gather the data, release it to officials so they would have complete information on household poverty, yet still leave them discretion to select recipients. If a pure information effect explains our results, this would also improve targeting. By combining proxies in the data and local knowledge, this might then be the optimal policy.

To explore this possibility, we examine an environment where the official has extensive information: his own home village. In Table 5, we restrict the sample to households in the winner's village only, and examine how targeting changed within the village. We find that even among households in winners' villages, there is an improvement in targeting: households with observable markers of wealth (solid housing materials, agricultural land, and higher education levels) are much *less* likely to receive BISP after the reform than before. The reform stopped transfers to households in the elected official's own home village, who are better off than their neighbors on characteristics that can be observed directly by any neighbor. Improved information alone cannot explain this result, because the official would have had easy access to this information about his neighbors pre-reform.

In addition, we examine characteristics of households in the elected officials' home villages who started receiving BISP only after the reform (Table A5). These households are poor by measures that would be easily observable to other members of the community: over 30% lived in rudimentary houses; three quarters do not own agricultural land. The majority of these households report that they know the elected official through personal interaction; thus it is not plausible that the officials would be unaware of these households or their poverty.

Taken together, these results show that the improvement in targeting due to the BISP reform was driven by the elimination of discretion that was enabled by administrative data, rather than a pure information effect.

4.4 Mechanisms for remaining favoritism

The suggestive evidence of continued favoritism in Table 2 raises the question: how could elected officials influence targeting after the reform, since they were formally removed from all parts of the process? We consider two main classes of mechanisms: influence on the administrative data used to qualify households for BISP; and assistance with downstream administrative processes such as appeals of the scorecard eligibility process or getting an ID card required to withdraw the cash.

Influence on administrative data

First, the administrative data could be subject to manipulation at several stages. Households could misreport their assets to government enumerators. Officials might help households connected to them anticipate how to respond strategically. To collect administrative asset data, the BISP teams typically interviewed respondents at the doorstep, so households could misreport their assets; there is no penalty other than potential discontinuation of the transfer.

Differential misreporting: We test directly for differential misreporting in home villages in the administrative asset data for Punjab province. We use village information in the administrative data to identify and restrict the sample to home villages of elected officials in 2008 and their rivals (as defined in Section 3). We then regress reported wealth in the administrative data on a connection to the 2008 winner (the elected official in office):

$$X_{ivc} = \beta_0 + \beta_1 WinVil_v + \eta_c + \epsilon_{ic} \tag{3}$$

Where X_{ivc} is a reported wealth indicator for household i in village v in constituency c in the administrative data, $WinVil_v$ is a dummy for the home village of an elected official in office (2008 winners), and η are constituency fixed effects. Standard errors are clustered by constituency. If elected officials encourage or assist in household under-reporting of assets to become eligible for the BISP, we expect $\beta_1 < 0$. In particular, we expect $\beta_1 < 0$ for unobservable assets. Table 6 shows the results. Panel A includes component variables which the government enumerator is more likely to be able to observe directly, even if she does not enter the dwelling: the ownership of cattle, the size of the house, and whether the house has an air conditioner; Panel B includes variables that are harder to observe from outside, including the land owned by the household, the education of the household head, and the number of family members who live in the household. We find no evidence of differential misreporting between winners' and rivals' villages, on either the observable or less observable wealth indicators.

It is possible that this overall finding masks finer manipulation within the distribution, for example shifting households to just below the eligibility cutoff. To investigate this, we estimate:

$$SCOREBIN_{ivc} = \beta_0 + \beta_1 WinVil_v + \eta_c + \epsilon_{ivc} \tag{4}$$

Where $SCOREBIN_{ivc}$ is a dummy for whether household i in village v constituency c has an overall poverty score in a given bin of the "poverty score". Standard errors are clustered at the constituency level. We estimate Equation 4 separately for each of a series of bins. Figure 3, Panel A, shows the results: again, we find no systematic differences in patterns of scores between winners' and rivals' villages.

Differential coverage: A second mechanism for favoritism to continue after the reform could be differential survey *coverage*. Ensuring coverage of all households was a challenge for teams moving through rural areas to complete the data collection. Elected officials might have been able to influence these teams to make sure that eligible households in their preferred areas were included, e.g. through extra followups with households who were not present when the team visited their area. To investigate this, we test whether elected officials' villages have a larger population of households in the administrative data than the villages of rivals. We estimate:

$$log(HHsCoveredBin)_{vc} = \beta_0 + \beta_1 WinVil_v + \eta_c + \epsilon_{vc} \tag{5}$$

Where the unit of observation is the village v; $log(HHsCoveredBin)_{vb}$ is the log number of households in the administrative data in a given poverty score bin. Standard errors are clustered at the constituency level. Figure 3, Panel B shows the results: again, we find no systematic differences in household coverage in each bin between winners' and rivals' villages.

Assistance with administrative processes

Appeals and grandfathering: All the beneficiaries who were removed from the BISP lists during the reform received an official notification, which also mentioned the appeals process, through which households within a higher band of poverty scores could be included in BISP if they launched an appeal. A connection with an elected official might help such households successfully appeal and stay on the list. To examine this, we estimate the cross-sectional version of our main specification on the subsample of households who had *not* received BISP in the pre-reform period (Table A6). The coefficients of interest are significantly different between the estimates on the full sample and subsample, suggesting that "grandfathering" represents one mechanism of the post-reform favoritism. However, the winner village term is still significant in the subsample, and suggests that 80% of post-reform favoritism went to households who did *not* receive BISP before the reform. "Grandfathering" explains some, but not all, the post-reform favoritism. We cannot test directly whether this occurred through appeals, but the disenrollment process for officials' discretionary nominees created a clear opportunity for such a mechanism.

National identity cards: To receive BISP, a female household member had to obtain a National Identity Card. Since elected officials in Punjab often provide assistance with getting an ID issued (Chaudhry and Vyborny, 2013), this is a potential mechanism for officials to assist their connections in getting BISP after the reform. Table A7 presents two suggestive tests for ID cards as a mechanism for continued assistance after the reform. Column 1 shows our basic estimations with receipt of a new ID card as the dependent variable. The results are 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 2-3 show the same specifications with BISP transfer as the dependent variable; in column 3 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: the coefficient on post-reform favoritism drops to zero after controlling for the ID card. These results suggest that elected officials may have assisted individuals in their villages to get ID cards, and that this played a role in any advantage these households had after the reform. However, we treat these results as suggestive, as it is also possible that continued favoritism operated through other channels and these households subsequently had ID cards made in order to complete the enrollment process.

Despite the transition to a rules-based system for selecting recipients, some potential recipients still faced administrative hurdles in getting the transfer; our results suggest that assistance from an elected official helped to clear these hurdles. Formally eliminating discretion may not be enough to put connected and unconnected households on an even footing. Simplifying administrative processes or assisting households to understand and access them may be necessary to ensure that delivery of benefits is fair de facto.

4.5 Public perceptions of social assistance programs

In contexts with weak institutions, voters may see capture or clientelism as legitimate, because ethnic networks help to hold politicians to account to deliver assistance. It is not obvious that the BISP reform would necessarily improve public perceptions. We use the phase-in of the reform to identify effects on public perceptions of the government's social safety net program across the province.

We estimate:

$$APPR_{idt} = \beta_1 + \beta_2 IMPL_{dt} + \gamma PREV_{-i,dt} + \alpha_d + u_t + \epsilon_{idt} \tag{6}$$

Where $APPR$ is a dummy taking value 1 if the respondent agrees with the statement "government schemes benefit the common man," asked in the context of a survey module about targeted transfers. $IMPL_{dt}$ is again a dummy which takes value 1 if the reform has been implemented in district d at time t. α_d are district fixed effects, and u_t round fixed effects. Because the number of beneficiaries increased over this period, we control flexibly for the prevalence of BISP by incorporating $PREV_{-i,dt}$, fixed effects for each decile of BISP prevalence in the data, excluding respondent i. We estimate equation 6 using MICS province-wide data from 2011 and 2014 (the only rounds in which the approval question was asked).

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. The pilot districts are similar to the rest of the province, including in pre-trends of political outcomes (Figure 2).

Table 4 shows the results. Implementation of the reform increases approval of social protection by approximately 13 percentage points, a 50% increase over the control group mean. Including the flexible controls for BISP prevalence $PREV_{-i,dt}$ (column 2) does not change the results.

Columns 3 and 4 show the results for the wealthiest subsample of respondents (the top two quintiles of the wealth index). The reform robustly increased approval of social assistance among the wealthy, who finance the program through their taxes, despite the fact it made them less likely to benefit from it directly.

We then investigate how this effect varies across the province by support for the incumbent party. We present a version of Equation 6 interacted with district level quantiles of the vote share for the incumbent party (PPP), which initiated the BISP transfer program:

$$APPR_{idt} = \alpha + \beta \cdot IMPL_{dt} \cdot PPPVOTE_d + \eta \cdot PPPVOTE_d \cdot Y2014_t + \gamma PREV_{-i,dt} + \psi Y2014_t + \alpha_d + \epsilon_{idt} \tag{7}$$

Figure 4 shows the results. In the overall sample (Panel A), the effects are most pronounced among *opposition* areas. In the wealthy subsample who would be ineligible for the program (Panel B), the effects are similar across the political spectrum. These results suggest that the reduction in favoritism by elected officials and the improvement in targeting helped to broaden support for social protection as a program, beyond households who might expect to benefit from favoritism or patronage based transfers.

4.6 Potential unintended consequences of the reform

It is possible that the BISP reform might induce a substitution effect, in which elected officials try to assist their connections with access to other government programs or services instead of BISP. In our home village survey we gathered information on all targeted assistance programs the household received. 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 find some evidence of a small increase in favoritism in other benefits in the pooled sample (significant at the 10% level in the pooled sample only, not the training and testing subsamples - Table A8). The magnitude of the estimate is 50 PKR, or approximately 50 US cents over the course of the three-year period. Compared to an average of over $7.20 for BISP favoritism of two percentage points (i.e. 2% of the households in the winner's village receiving BISP at $10 per month for three years), this is a small increase. It is also small relative to the size of the control mean (16% of the mean in rivals' villages, whereas pre-reform BISP favoritism was 200%).

Another potential unintended consequence of the BISP reform would be to increase household misreporting of assets in other door-to-door data collection such as national sample household surveys; this could reduce the quality of data used for policy analysis and planning. Respondents may not distinguish between the BISP team and such survey teams, or they might under-report their assets everywhere for consistency (cf. Hurst *et al.* (2014)). 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, might lead a respondent who has misreported assets in the BISP data to repeat the false report. In the MICS representative sample, we again exploit the rollout of the reform, with assets reported in the survey as the dependent variable. If formula-based targeting led to misreporting, we should expect the reform to decrease reports of wealth measures, particularly unobservable ones. The reform has no significant effect on self reports of either observable assets (dwelling characteristics, electricity connection wires) or unobservable assets (ownership of land and cattle) (Table A9).

4.7 Welfare benefits of improved targeting

Creating the BISP administrative database and using it to select BISP recipients improved targeting. How did the benefits compare to the costs?

To quantify the welfare benefits of eliminating discretion in BISP, we use the provincially representative MICS data to simulate the distribution of household expenditure (a) without BISP; (b) with BISP targeted by discretion; and (c) with BISP targeted by the new formula. We then estimate the welfare gains from using the formula as compared to discretion, and assess how much administrative costs would offset this gain. We implement this as follows, using an approach similar to Alatas *et al.* (2019).

1) Predict probability of receiving BISP under formula targeting

First, we simulate which households would receive BISP under the formula-based approach post reform (Proxy Means Test). We model the probability of receiving the transfer post reform as a function of a vector of household characteristics:

$$Pr(BISP_i) = \beta_0 + \gamma X_i + \epsilon_i \tag{8}$$

Where X is a vector including variables both included and not included in the official formula, which are unlikely to be shifted by receiving the transfer. This includes the ratio of rooms in the household to the household size; the number of dependents age 5-16; the number of working adults age 18-65; the ratio of dependents to working adults; ownership of the dwelling; ownership of agricultural land; whether the dwelling is solid; whether the household is female headed; whether the household is rural; and dummies for whether the household head has completed primary, middle, or matriculate education. Equation 8 is estimated using a Probit model with the data from MICS 2014, when the reform was complete in all districts. We use this to generate predicted probabilities of receiving BISP in earlier, pre-reform rounds, under the counterfactual scenario that the formula was used to target BISP, i.e. $Pr(\widehat{BISP}_{i,FORM})$.

2) Compare recipients under discretion to predicted recipients under formula targeting

Second, we compare the predicted recipients based on formula-based targeting to the households that actually received BISP pre reform. We focus on the 2011 data, collected when BISP was still distributed under discretionary targeting by officials in many districts. In these districts, we observe the actual receipt of BISP under discretion, $BISP_{i,DISC}$. We assign a counterfactual transfer of BISP under the formula based targeting to households in the same sample, $\widehat{BISP}_{i,FORM}$, using the estimated probabilities from 8 and holding the total number of recipients constant between the two targeting schemes.

3) Predict household expenditure with and without BISP cash transfer

We use pre-BISP data (from 2003) to predict per-capita household expenditure in the absence of the transfer:

$$ln(exp)_i = \beta_0 + \gamma X_i + \epsilon_i \tag{9}$$

Where X is the same vector of fixed household characteristics listed above. We use these estimates to generate predicted values of household expenditure in the absence of the transfer, $ln\widehat{(exp)}_i$.

We then simulate per-capita household expenditure including the transfer amount under different targeting schemes. We calculate households' expected monthly expenditure under BISP targeted with discretion, including predicted expenditure from 9 plus the monthly value of the transfer for households that receive it. Similarly, we calculate households' expected monthly expenditure under BISP targeted with the formula:

$$\widehat{exp}_{i,DISC} = \widehat{exp}_i + BISP_{i,DISC} \times 1200 \tag{10}$$

$$\widehat{exp}_{i,FORM} = \widehat{exp}_i + \widehat{BISP}_{i,FORM} \times 1200 \tag{11}$$

4) Estimate social welfare

Finally, we calculate welfare under each targeting approach as the sum of log expenditure without BISP ($\sum_i ln(\widehat{exp_i})$), under BISP targeted with discretion ($\sum_i ln(\widehat{exp}_{i,DISC})$) and BISP targeted with the formula ($\sum_i ln(\widehat{exp}_{i,FORM})$).[16] We then repeat this calculation adjusting the "budget" for BISP transfers down from 1200 per month by 2%, the reported administrative cost of the administrative data collection as a percentage of the transfers delivered in the period between its initial collection and when it was first updated (see Section 2), and a larger benchmark of 15% administrative costs.

Figure 5 summarizes the results. Panel 5a shows the distribution of counterfactual per-capita expenditure in the absence of BISP, $ln\widehat{(exp)}_i$, for households targeted under discretion and those predicted to be targeted under the formula. Relative to discretion, formula based targeting selects poorer households. This is reflected in the welfare estimates in Panel 5b. The bars are shown as percentage changes in welfare gain relative to the scenario with no cash transfers. The reform increases the welfare gains from the cash transfer by 15%, even after accounting for administrative costs. Even a hypothetical 15% administrative cost - over seven times the reported administrative costs of the reform - would still improve welfare compared to discretionary

[16] This approach makes several assumptions. First, it ignores the incidence of taxation. This is a reasonable approximation given that the main comparison of interest, the targeting decision, is separable from the incidence of taxation. Second, it ignores potential margins of response such as reduction of labor supply. However, there is limited empirical evidence of this response (Banerjee *et al.*, 2017b), and to the extent it occurs, households still experience a welfare benefit through the increase in leisure.

targeting.

5 Discussion

In this paper, we quantify the effects of the development and use of administrative data for social protection in a low capacity context. We find that a major reform to generate administrative data on household assets and use it for formula-based targeting reduced favoritism and transfers to wealthier households. Our welfare estimates suggest that the value of this targeting improvement was large, exceeding the reported administrative costs of the reform by a factor of seven. These estimates include only the benefit from improved targeting of the cash transfer. They do not account for any benefit from more reliable and continuous payment of the transfer through political transitions; thus they are a lower bound on the welfare benefits of the reform.

We demonstrate that the reform improved public perceptions of the program, including among wealthier households may be net payers for the program through their taxes and opposition areas of the province. Prior to BISP, social safety net policy in Pakistan has been unstable. Successive governments have reduced the budget of previous programs and set up new programs in parallel. This would incur higher administrative costs, but also make the government's social protection program as a whole less effective, as recipients cannot rely on a consistent stream of benefits - the purpose of transfer programs as a safety net was undermined. After a new party took power in Pakistan's 2013 elections, there was widespread speculation that the BISP program would similarly be discontinued or defunded. However, the new government continued the program, placing the photos of their own party's leaders on program materials. Again in 2018, a third party took power at the national and provincial level; and rather than defunding the program the government rebranded it and made it the center piece of an expanded social protection program. In addition, the government of Pakistan has increased the domestically funded proportion of the BISP program budget. We cannot attribute these decisions solely to the improvement in public perceptions of the program after the reform. However, expert interviews suggest that the widespread recognition that the program was fairly distributed may have played a role. In addition, both successor governments moved towards using the new national administrative poverty database created during the reform to identify recipients for other programs. Over thirty government and non-government programs use the database for targeting to date. In April 2020, the government announced it would use the same database to target rapid response emergency cash transfers to assist households with the economic impact of COVID-19; because this administrative data was already established, it enabled the government institutions to provide far more timely emergency response than would have been possible otherwise.

The development of a new administrative database, and its use to enable non-discretionary targeting, helped to build institutional capacity for social protection in a low capacity context. Ultimately, such institutional strengthening may make government programs more politically stable, continuous and effective in the long run.

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. (2010). *The Role of Institutions in Growth and Development*, vol. 1.

—, 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. (2019). Does Elite Capture Matter? Local Elites and Targeted Welfare Programs in Indonesia. *AEA Papers and Proceedings*, **109**, 334–339.

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

—, Banerjee, A. V., Hanna, R., Olken, B. A., Purnamasari, R. and Wai-poi, M. (2016). Self-Targeting: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy*, **124** (2).

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, M. L. and Magruder, J. (2017). Split sample strategies for avoiding false discoveries. *NBER Working Paper 23544*.

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

Anker, A. S. T., Doleac, J. L. and Landerss, R. (2019). The Effects of DNA Databases on the Deterrence and Detection of Offenders.

Banerjee, A., Hanna, R., Kyle, J., Olken, B. A. and Sumarto, S. (2017a). Tangible Information and Citizen Empowerment: Identification Cards and Food Subsidy Programs in Indonesia. *Journal of Political Economy*, **126** (2), 451–491.

— 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.

—, Hanna, R., Kreindler, G. E. and Olken, B. A. (2017b). Debunking the Stereotype of the Lazy Welfare Recipient : Evidence from Cash Transfer Programs. *World Bank Research Observer*, (September), 155–184.

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 Mitra, S. (2014). Political participation, clientelism and targeting of local government programs: analysis of survey results from rural West Bengal, India. *Working Paper, Boston University*, **02215**, 1–42.

—, —, Mookherjee, D. and Nath, A. (2014). Changing Voting Patterns in Rural West Bengal: Role of Clientelism and Local Public Goods. *Economic and Political Weekly*, **49** (11), 54–62.

—, —, — and Sarkar, A. (2009). Local Democracy and Clientelism: Implications for Political Stability in Rural West Bengal. pp. 46–58.

— 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.

Barnwal, P. (2019). Curbing leakage in public programs: Evidence from India's direct benefit transfer policy. (August), 1–64.

Basurto, P., Dupas, P. and Robinson, J. (2019). Decentralization and Efficiency of Subsidy Targeting: Evidence from Chiefs in Rural Malawi. *Working paper, Stanford University.*

Beg, S. (2020). Digitization and Development: Formalizing Property Rights and its impact on Land and Labor Markets. *Mimeo, University of Delaware.*

Bertrand, M., Duflo, E. and Mullainathan, S. (2004). How much should we trust difference-in-differences estimates? *Quarterly Journal of Economics*, **119** (1).

Besley, T. (1990). Means Testing versus Universal Provision in Poverty Alleviation Programmes. *Economica*, **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. (2012). Just Rewards? Local Politics and Public Resource Allocation in South India. *The World Bank Economic Review*, **26** (2), 191–216.

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.

Burgess, R., Jedwab, R., Miguel, E., Morjaria, A. 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.

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. *The 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, 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).

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. (2002). Community Based Targeting Mechanisms for Social Safety Nets: A Critical Review. *World Development*, **30** (3).

Currie, J. (2006). The Take Up of Social Benefits. In A. J. Auerbach, D. Card and J. M. Quigley (eds.), *Public Policy and the Income Distribution*, New York: Russell Sage Foundation Publications.

De La O, A. (2013). Do Conditional Cash Transfers Affect Electoral in Mexico. *American Journal of Political Science*, **57** (1), 1–14.

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.

Diaz-Cayeros, A., Esteves, F. and Magaloni, B. (2016). *Political Logic of Poverty Relief: Electoral Strategies and Social Policy in Mexico*. Cambridge University Press.

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

Doleac, J. L. (2017). The Effects of DNA Databases on Crime. *American Economic Journal: Applied Economics*, **9** (1), 165–201.

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., Folke, O., Fowler, A., Hainmueller, J., Hall, A. B. and Snyder, J. M. (2013). 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. (2017a). Do Politicians' Relatives Get Better Jobs? Evidence from Municipal Elections. *Journal of Law, Economics, and Organization*.

— and — (2017b). Using Split Samples to Improve Inference about Causal Effects. *Political Analysis*, **25**, 465–482.

Fair, C. C., Kuhn, P., Malhotra, N. A. and Shapiro, J. (2017). *Natural Disasters and Political Engagement: Evidence from the 201011 Pakistani Floods*. September 2016.

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.

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? *Economic & Political Weekly*, **xlvi** (28), 59–66.

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

Giné, X., Goldberg, J. and Yang, D. (2012). Credit Market Consequences of Improved Personal Identification: Field Experimental Evidence from Malawi. *American Economic Review*, **102** (6), 2923–2954.

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

Hanmer, L. and Elefante, M. (2016). The Role of Identification in Ending Child Marriage. *World Bank*.

Hanna, R. and Olken, B. A. (2018). Universal Basic Incomes versus Targeted Transfers: Anti-Poverty Programs in Developing Countries. *Journal of Economic Perspectives*, **32** (4), 201–226.

Hickey, S. (2008). Conceptualising the Politics of Social Protection in Africa. In A. Barrientos and D. Hulme (eds.), *Social Protection for the Poor and Poorest: Concepts, Policies and Politics*, October.

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

Hou, X. (2011). Poverty Scorecard for Pakistan: An Update Using the PSLM 2007-2008 Data and the Choice of the Cut-off Score. *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*.

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.

Hussam, R., Rigol, N. and Roth, B. (2018). Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design In The Field. *Working paper, Harvard Business School*, pp. 1–45.

Imai, K., King, G. and Rivera, C. V. (2020). Do Nonpartisan Programmatic Policies Have Partisan Electoral Effects? Evidence from Two Large-Scale Experiments. *Journal of Politics*, **82** (2).

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.

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*.

Kilic, T., Whitney, E. and Winters, P. (2015). Decentralised beneficiary targeting in large-scale development programmes: Insights from the Malawi farm input subsidy programme. *Journal of African Economies*, **24** (1), 26–56.

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

Kleven, H. J. and Kopczuk, W. (2011). Transfer Program Complexity and the Take-Up of Social Benefits. *American Economic Journal: Economic Policy*, **3** (February), 54–90.

Kosec, K. and Mo, C. H. (2019). Does Relative Deprivation Condition the Effects of Social Protection Programs on Political Attitudes? Experimental Evidence from Pakistan. *IFPRI Discussion Paper 01842*, (May).

Labonne, J. (2013). The local electoral impacts of conditional cash transfers. Evidence from a field experiment. *Journal of Development Economics*, **104**, 73–88.

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.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, **142**, 698–714.

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 Indi. *American Economic Review*, **106** (10), 2895–2929.

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

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

Nguyen, K.-T., Do, Q.-A. and Tran, A. (2017). One Mandarin Benefits the Whole Clan: Hometown Infrastructure and Nepotism in an Autocracy. *American Economic Journal: Applied Economics*, **9** (4), 1–29.

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.

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

Pomeranz, D. (2015). No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax. *American Economic Review*, **105** (8), 2539–2569.

Pritchett, L. (2005). A Lecture On The Political Economy of Targeted Safety Nets. (0501).

Sarwar, M. B. (2018). The political economy of cash transfer programmes in Brazil , Pakistan and the Philippines. *ODI Working Paper 543*.

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.

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

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*, vol. 58. MIT Press.

World Bank (2007). Social Protection in Pakistan: Managing Household Risks and Vulnerability. *Report No. 35472-PK*, (35472).

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

World Bank (2018a). The Role of Digital Identification for Healthcare.

World Bank (2018b). *The State of Social Safety Nets 2018.*

Table 1: Summary statistics and balance

	Home village survey - Full sample			Home village survey - RDD subset			MICS province-wide survey		
	Rival's village	Winner's village	SE difference	Rival's village	Winner's village	SE difference	Pilot districts	Other districts	
							Mean		SE difference
Same clan as origin-village official / rival	0.169	0.123	(0.074)	0.181	0.062	(0.087)	-	-	
Related to elected or local official	0.039	0.036	(0.019)	0.045	0.010	(0.019)*	-	-	
Female head	0.031	0.028	(0.007)	0.031	0.026	(0.010)	0.04	0.05	(0.00799)**
Any daughters aged 18-25	0.391	0.413	(0.030)	0.389	0.445	(0.036)	-	-	
Rudimentary house	0.310	0.318	(0.087)	0.325	0.375	(0.142)	0.52	0.44	(0.0706)
No agricultural land	0.745	0.821	(0.045)	0.741	0.885	(0.051)**	0.68	0.65	(0.0727)
No cattle	0.699	0.728	(0.053)	0.707	0.808	(0.069)	0.60	0.59	(0.0874)
No residential land	0.151	0.186	(0.085)	0.118	0.296	(0.138)	-	-	
HH head 5th grade or higher	0.459	0.442	(0.050)	0.445	0.461	(0.075)	0.47	0.49	(0.0317)
HH head 8th grade or higher	0.321	0.324	(0.045)	0.306	0.315	(0.061)	0.47	0.49	(0.0317)
HH head 10th grade or higher	0.215	0.227	(0.033)	0.212	0.212	(0.048)	0.31	0.33	(0.0298)
Years HH has lived in village	78.951	77.529	(5.045)	76.883	73.367	(8.061)	-	-	
Years squared	6740.765	6571.038	(641.872)	6484.715	5972.873	(1057.884)	-	-	
HH member received ID (2002-07)	0.495	0.545	(0.057)	0.514	0.478	(0.064)	-	-	
N	3970	3733		3246	1478		7980	83095	

Notes: Unit of observation is the household. Home village survey: Recall data from 2007 are used for data for house materials and cattle ownership. Standard errors are estimated by regressing the household characteristic on a dummy for winner's village, clustered by village. MICS province-wide survey: Standard errors are estimated by regressing the household characteristic on a dummy for pilot district, clustered by district. * p < 0.1, ** p < .05, *** p < .01.

Table 2: Impact of BISP reform on favoritism for officials' village and clan

	(1)	(2)	(3)	(4)	(5)
			HH received BISP cash transfer		
β_1: Winner's village x winner's clan x pre	0.030	0.032	0.010	0.029	0.025
	(0.010)***†‡	(0.007)***†‡	(0.016)	(0.012)**†	(0.011)**†
β_2: Winner's village x other clan x pre	0.017	0.021	0.016	0.021	0.027
	(0.006)***†‡	(0.006)***†‡	(0.007)**	(0.006)***†‡	(0.007)***†‡
β_3: Winner's village x winner's clan x post	-0.015	-0.010	-0.032	0.005	0.001
	(0.012)	(0.009)	(0.013)**	(0.012)	(0.015)
β_4: Winner's village x other clan x post	0.024	0.027	0.023	0.007	0.017
	(0.016)‡	(0.012)**‡	(0.009)**	(0.025)	(0.015)
Vote margin		X	X		X
HH controls		X	X		X
HH controls x post		X	X		X
Constituency FE			X		
RD subset				X	X
Observations	15390	15390	15390	9440	9440
Mean Y:					
Rival village, other clan, pre			0.01		
Rival village, other clan, post			0.04		
P-values, village favoritism effect equal for:					
Winner's clan, pre = post ($\beta_1 = \beta_3$)	0.00***†‡	0.00***†‡	0.00***	0.00***†‡	0.10
Other clan, pre = post ($\beta_2 = \beta_4$)	0.64	0.60	0.60	0.58	0.58
Pre, winner's clan = other clan ($\beta_1 = \beta_2$)	0.22	0.16	0.67	0.44	0.81
Post, winner's clan = other clan ($\beta_3 = \beta_4$)	0.06*	0.02**†‡	0.00***	0.94	0.53

Notes: Home village pooled sample. Unit of observation is the household-round (pre / post reform). Equation 1; each coefficient shows the estimated favoritism for households in the winner's village, compared to households in the rivals' village. The four groups (winner's clan pre / post reform; other clans pre / post reform) are mutually exclusive and exhaustive (as shown in Figure 1). All specifications include control for post reform, winner's clan, and winner's clan x post. Household controls in columns 2, 3, and 5 include female-headed household; no agricultural land; rudimentary house; any daughters aged 18-25; no cattle; no residential land; household head 5th, 8th, 10th grade or higher; years household has lived in the village; years squared; any household member had an ID card issued from 2002-2007; winner's clan; and rival's clan. * $p < 0.1$; ** $p < .05$; *** $p < .01$ (robust standard errors clustered at the village level); † $p < 0.1$ for bootstrapped standard errors clustered at the village level; ‡ $p < 0.1$ for cluster wild bootstrap by constituency. Preferred specification uses random effects for efficiency; results are robust to pooled OLS and household fixed effects estimation (Table A2) and to out-of-sample testing (Web Appendix C).

Table 3: BISP reform and targeting the poor: provincial sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Poor	0.020***	0.023***	0.021***	0.026***	0.012***	0.016***	-0.001
	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.005)
Reform Implemented	-0.007	-0.015*	0.021	0.016	0.013	0.020	0.002
	(0.005)	(0.008)	(0.014)	(0.018)	(0.014)	(0.021)	(0.022)
Reform implemented × poor	0.053**	0.048*	0.010	0.014	0.029*	0.024*	0.104***
	(0.025)	(0.025)	(0.011)	(0.014)	(0.016)	(0.014)	(0.028)
Observations	133562	133562	133562	133562	133562	133562	133562
Baseline mean - C group - poor HHs	0.040	0.036	0.041	0.037	0.041	0.043	0.038
Baseline mean - C group - wealthy HHs	0.017	0.008	0.020	0.015	0.024	0.027	0.028
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Round FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Poverty proxy	Wealth Q1-3	Wealth Q1-4	Head ed <8	No ag land	Solid roof	Solid walls	Solid floors
Proportion of HHs categorized as poor	0.580	0.800	0.507	0.691	0.360	0.200	0.185
Overlap with official formula?	Yes	Yes	Yes	Yes	No	No	No

Notes: Provincially representative Multiple Indicator Cluster Survey, 2011 and 2014 rounds. Equation 2. "Reform implemented" is an indicator variable defined at the district-year level. "Wealthy" and "Poor" are indicators for wealth proxy=1 and wealth proxy=0, based on the wealth proxy named in the table footer for each specification. Terms for "poor," "year 2014," "poor x 2014," "pilot x poor" are included in the estimate but omitted from the table. Standard errors in parentheses, clustered at the district level. * p < .1, ** p < .05, *** p < .01.

Table 4: BISP reform and public approval of social assistance

	Government schemes are beneficial			
	(1)	(2)	(3)	(4)
Reform Implemented	0.127***	0.128**	0.0923**	0.0714*
	(0.0446)	(0.0487)	(0.0428)	(0.0381)
Observations	130176	119773	54667	53189
Sample mean	0.264	0.264	0.272	0.272
District FE	Yes	Yes	Yes	Yes
Round FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Flexible Controls BISP prevalence	No	Yes	No	Yes
Sample	All	All	Wealthy	Wealthy

Notes: Provincially representative Multiple Indicator Cluster Survey, 2011 and 2014 rounds. "Reform implemented" is an indicator variable defined at the district-year level as in Equation 2. Standard errors in parentheses, clustered at the district level. * p < .1, ** p < .05, *** p < .01. Wealthy Sample restriction exclude all households that do not fall under 4th and 5th wealth quintiles.

Table 5: Impact of reform on targeting: winner villages only

	(1)
	BISP
Post reform	0.005 (0.066) {0.068}
Male head	-0.059 (0.033)* {0.034}*
Male head x post reform	0.044 (0.039) {0.043} 0.33
Solid house	-0.010 (0.007) {0.007}
Solid house x post reform	-0.043 (0.018)** {0.020}** 0.12
Any agricultural land	-0.011 (0.004)*** {0.004}***
Any ag land x post reform	-0.029 (0.010)*** {0.008}*** 0.02
Any residential land	-0.022 (0.019) {0.018}
Any res land x post reform	0.027 (0.035) {0.034} 0.52
HH head ed 8 yrs+	-0.004 (0.006) {0.007}
Ed 8 yrs+ x post reform	-0.014 (0.007)* {0.008}* 0.08
Observations	7928

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. Estimates are followed by robust standard errors clustered at the village level (parentheses); bootstrapped standard errors clustered at the village level {braces}, and p-value for cluster wild bootstrap test of $H_0 : \beta = 0$. Results are robust to out-of-sample testing; tables available upon request. We also examine targeting patterns in the full sample of home villages in Web Appendix Tables B7-B8.

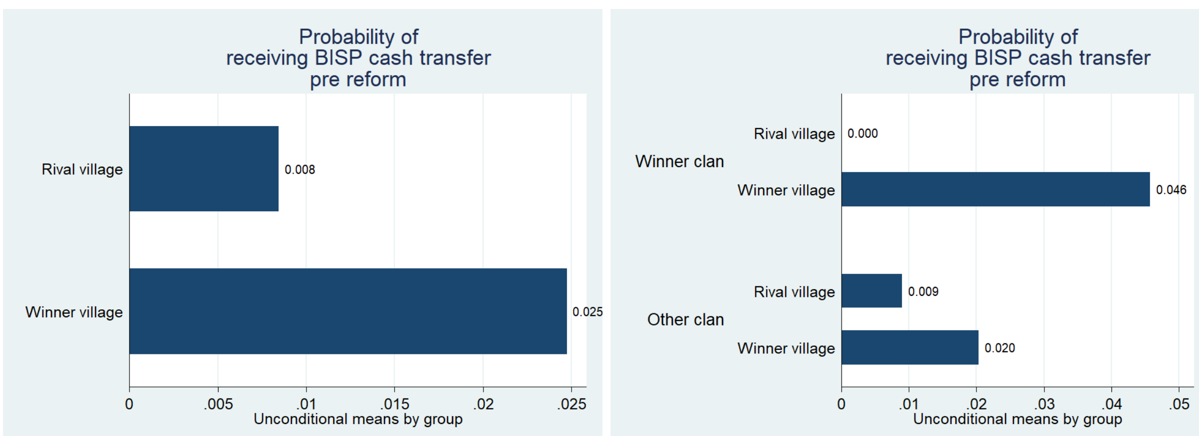
Table 6: Mechanism test: differential misreporting in administrative data on observable and unobservable assets

Panel A: Components of poverty score observable to government enumerator				
	(1)	(2)	(3)	(4)
	Any cows	Any buffalos	Air conditioner	N rooms
Winner's Village	0.036*	0.002	0.000	0.037
	(0.019)	(0.014)	(0.000)	(0.041)
Observations	107315	107315	107315	107315
Sample mean	0.069	0.056	0.001	1.308
Constituency FE	Yes	Yes	Yes	Yes

Panel B: Components of poverty score unobservable to government enumerator			
	(1)	(2)	(3)
	Acres land	Years education	N HH members
Winner's Village	0.017	-0.048	0.087
	(0.043)	(0.059)	(0.082)
Observations	107314	94120	107315
Sample mean	0.111	0.841	2.863
Constituency FE	Yes	Yes	Yes

Notes: Administrative poverty score data. Robust standard errors clustered at the constituency level. * $p < 0.1$; ** $p < .05$; *** $p < .01$.

Figure 1: Distribution of BISP pre reform

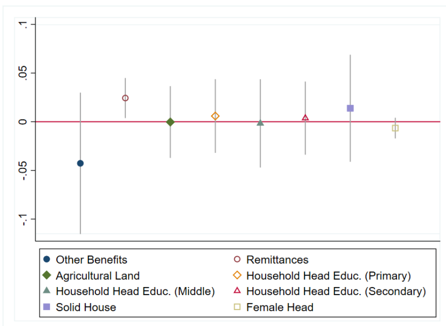


(a) By village

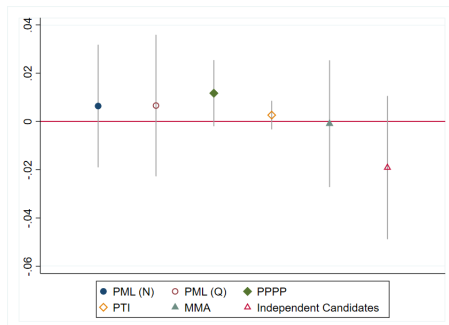
(b) By village and clan

Notes: Home village training sample, pre-reform observations only. Unit of observation is the household.

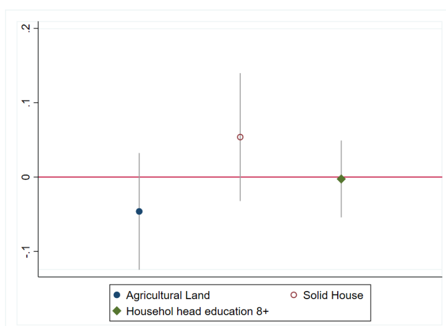
Figure 2: Pre-Trends across Pilot and non-Pilot Districts



(a) Economic Pre-trends: Household characteristics (indicator variables)



(b) Political Pre-trends: Vote share by party

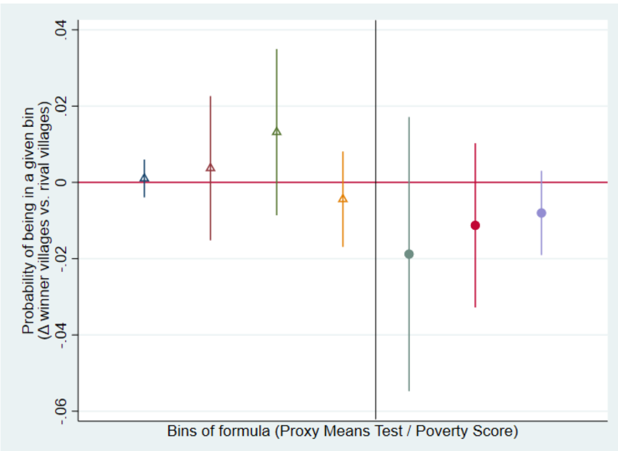


(c) Targeting Pre-trends: Triple Interaction

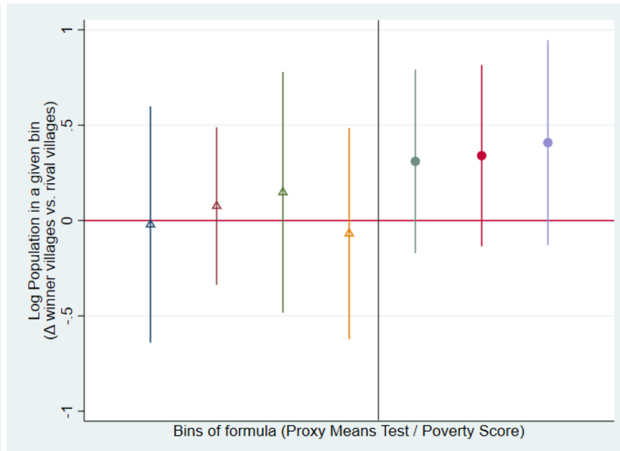
Notes: Panels A-B: Estimated $\hat{\beta}_1$ from $Y_{idt} = \beta_0 + \beta_1 PILOT_d \times YEAR2003_t + \alpha_d + \mu_t + \epsilon_{idt}$. where Y are the characteristics noted in the legend (government benefits, household characteristics, and vote share). Panel C: Triple interaction across pilot district, year 2003 and a wealth proxy: $\hat{\beta}_5$ from $BENEFITS_{idt} = \beta_0 + \beta_1 WEALTH_i + \beta_2 WEALTH_i \times PILOT_d + \beta_3 WEALTH_i \times YEAR2003_t + \beta_4 PILOT_d \times YEAR2003_t + \beta_5 WEALTH_i \times PILOT_d \times YEAR2003_t + \alpha_d + \mu_t + \epsilon_{idt}$, where $BENEFITS$ is a dummy for receiving any pre-BISP transfer program and $WEALTH$ are binary wealth proxies shown in the legend. Standard errors are clustered at the district level in all specifications.

Figure 3: Mechanism tests for favoritism: Influence on administrative Poverty Score data

(a) Test of differential misreporting: probability of a household's official poverty score falling in each bin



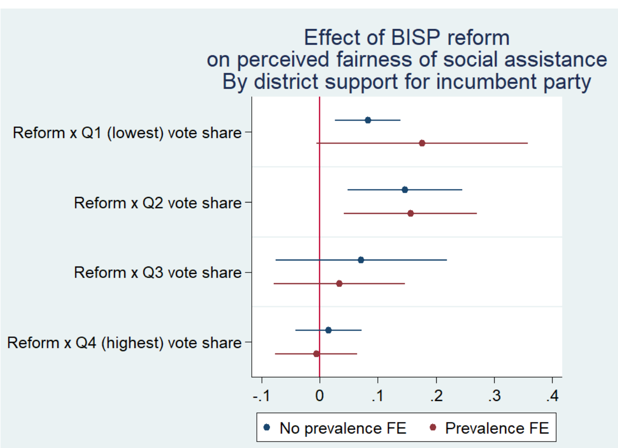
(b) Test of differential coverage: Number of households recorded in village in each bin



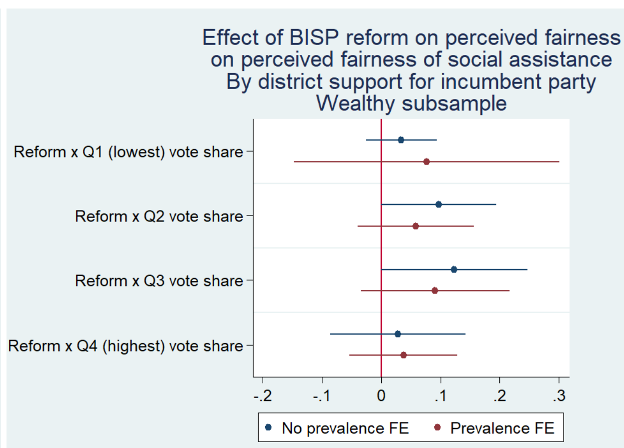
Notes: Administrative poverty score data. Panel A: Each coefficient is the estimate of β_1 (comparing the winner's village to rival's village) in Equation 4 for one bin of the poverty score; the dependent variable in each case is a dummy for whether the household's administrative poverty score falls in that bin. Panel B: Each coefficient is the estimate of β_1 (comparing the winner's village to rival's village) in Equation 5; the dependent variable in each case is the log population in village i in administrative poverty score bin j. Both panels: All estimates include constituency fixed effects and are clustered at the constituency level. The vertical line shows the cutoff for BISP eligibility.

Figure 4: Effects of BISP reform on public approval of social assistance by political alignment

(a) Full sample



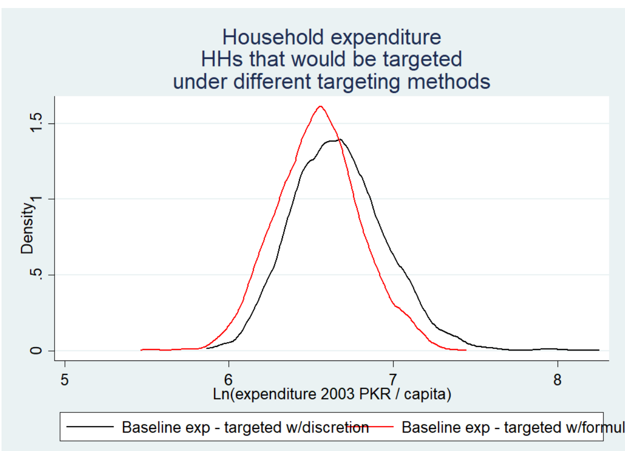
(b) Wealthy subsample



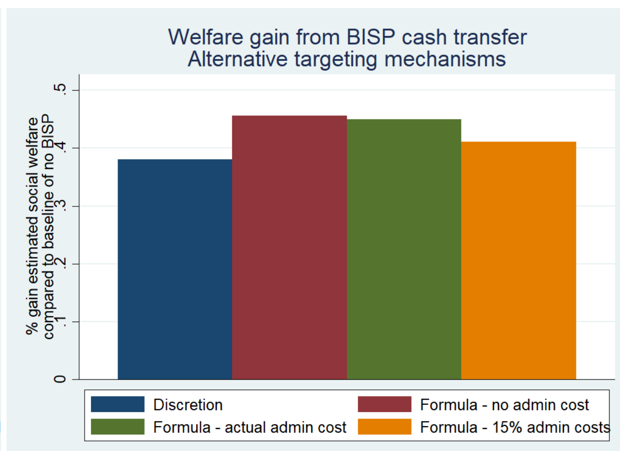
Notes: MICS provincial survey data. Coefficients from Equation7. All estimates include controls for differential time trends for each quantiles of vote share ×. Standard errors clustered by district. 95% confidence intervals shown.

Figure 5: Welfare estimates

(a) Baseline expenditure of households that would be targeted under discretionary vs. formula-based targeting



(b) Welfare gain from BISP reform under alternative assumptions of administrative costs



Notes: Authors' calculations using MICS 2003-2014 household survey data, following the approach described in Section 4.7.

A Appendix: Supplemental tables

Table A1: Summary statistics by clan

	(1)	(2)	(3)	(4)
	Candidate's clan	Others	Difference	SE difference
Female head	0.03806	0.0283	0.00976	(0.00646)
Any daughters currently aged 18-25	0.3822	0.405	-0.0228	(0.0276)
Rudimentary house (lag - 07)	0.2486	0.325	-0.0764	(0.0833)
No ag land	0.56	0.820	-0.260	(0.0657)***
No cattle (lag - 07)	0.608	0.731	-0.123	(0.0401)***
No residential land	0.1045	0.179	-0.0745	(0.0592)
HH head 5th grade or higher	0.589	0.427	0.162	(0.0704)**
HH head 8th grade or higher	0.47	0.297	0.173	(0.0685)**
HH head 10th grade or higher	0.336	0.202	0.134	(0.0533)**
Years in village	78.2738	78.26	0.0138	(3.492)
Years squared	6635.84	6662.4	-26.56	(448.3)
HH member received ID (02-07)	0.52579	0.518	0.00779	(0.0588)
Related to elected or local official	0.1222	0.0230	0.0992	(0.0362)***

Notes: Home village sample. Candidate clan is defined as 1 for any household i that is the same clan as the winner (elected official) or rival from i's village. Standard errors are clustered at the village level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Impact of reform on favoritism: robustness to pooled OLS and HH FE

	(1)	(2)	(3)	(4)	(5)
	HH received BISP cash transfer				
	Panel A: Pooled OLS				
Winner's village x winner's clan x pre	0.046	0.047	0.023	0.045	0.041
	(0.022)**	(0.017)***	(0.026)	(0.026)*	(0.022)*
	0.020**	0.022**		0.029	0.024*
Winner's village x other clan x pre	0.011	0.017	0.009	0.012	0.018
	(0.006)*	(0.006)***	(0.008)	(0.006)*	(0.008)**
	0.006**	0.008**		0.006**	0.012
Winner's village x winner's clan x post	-0.011	-0.010	-0.034	0.017	0.000
	(0.017)	(0.015)	(0.018)*	(0.013)	(0.020)
	0.015	0.021		0.014	0.029
Winner's village x other clan x post	0.023	0.028	0.020	0.005	0.014
	(0.019)	(0.013)**	(0.010)*	(0.028)	(0.016)
	0.016	0.016*		0.024	0.017
Observations	7356	7356	7356	4520	4520
	Panel B: HH FE				
Winner's village x winner's clan x post	-0.057	-0.058	-0.058	-0.028	-0.041
	(0.017)***	(0.024)**	(0.024)**	(0.013)*	(0.023)*
Winner's village x other clan x post	0.012	0.011	0.011	-0.007	-0.004
	(0.017)	(0.014)	(0.014)	(0.027)	(0.018)
Observations	7356	7356	7356	4520	4520
Vote margin		X	X		X
Controls		X	X		X
Controls x post		X	X		X
Constituency FE			X		
RD subset				X	X

Notes: Home village "training" sample. Estimates of key parameters are followed by robust standard errors clustered at the village level (parentheses); bootstrapped standard errors clustered at the village level {braces}. All specifications include controls for winner's clan and post reform. Household controls include female-headed household; no agricultural land; rudimentary house; any daughters aged 18-25; no cattle; no residential land; household head 5th, 8th, 10th grade or higher; years household has lived in the village; years squared; any household member had an ID card issued from 2002-2007; winner's clan; and rival's clan. * $p < 0.1$; ** $p < .05$; *** $p < .01$.

Table A3: Impact of BISP reform on favoritism for officials' villages

	(1)	(2)	(3)	(4)	(5)
	HH received BISP cash transfer				
Winner's village x pre reform	0.018	0.022	0.017	0.022	0.027
	(0.006)***	(0.006)***	(0.007)**	(0.006)***	(0.007)***
	{0.006}***	{0.006}***		{0.006}***	{0.010}***
	0.00	0.00		0.01	0.01
Winner's village x post reform	0.017	0.024	0.019	0.002	0.015
	(0.014)	(0.010)**	(0.008)**	(0.023)	(0.013)
	{0.016}	{0.011}**		{0.026}	{0.018}
	0.15	0.03		0.92	0.22
Vote margin		X	X		X
HH controls		X	X		X
HH controls x post		X	X		X
Constituency FE			X		
RD subset				X	X
Observations	15390	15390	15390	9440	9440
Mean Y:					
Rival village, pre reform				0.01	
Rival village, post reform				0.04	
P-values:					
Winner's village, pre = post					
Robust SE clustered by village	0.92	0.86	0.86	0.39	0.50
BS SE clustered by village	0.92	0.85		0.45	0.58
Wild BS clustered by constituency	0.89	0.84		0.30	0.57

Notes: Home village sample. Estimates of key parameters are followed by robust standard errors clustered at the village level (parentheses); bootstrapped standard errors clustered at the village level {braces}; and P-values for cluster wild bootstrap test of null that coefficient equals zero. All specifications include control for post reform. Household controls include female-headed household; no agricultural land; rudimentary house; any daughters aged 18-25; no cattle; no residential land; household head 5th, 8th, 10th grade or higher; years household has lived in the village; years squared; any household member had an ID card issued from 2002-2007; winner's clan; and rival's clan. * $p < 0.1$; ** $p < .05$; *** $p < .01$. Results are robust to out-of-sample testing (Web Appendix C)

Table A4: Robustness check for recall bias

	(1)	(2)	(3)	(4)
	HH received BISP cash transfer:			
	Ever	Last year	Ever	Last year
Winner's village	0.052	0.024		
	(0.011)***	(0.010)**		
	{0.015}***	{0.012}**		
	0.00	0.03		
Winner's village x winner's clan			0.037	-0.010
			(0.014)**	(0.009)
			{0.017}**	{0.013}
			0.05	0.32
Winner's village x other clan			0.053	0.028
			(0.013)***	(0.012)**
			{0.015}***	{0.013}**
			0.00	0.03
P-values, cross-equation test of equal coefficients:				
Winner's village	0.00			
Winner's village x other clan			0.00	
Winner's village x winner's clan			0.00	
Observations	7695	7695	7695	7695

Notes: Home village sample. All specifications include controls for winner's clan, party fixed effects, and household controls including female-headed household; no agricultural land; rudimentary house; any daughters aged 18-25; no cattle; no residential land; household head 5th, 8th, 10th grade or higher; years household has lived in the village; years squared; any household member had an ID card issued from 2002-2007; winner's clan; and rival's clan. Estimates of key parameters are followed by robust standard errors clustered at the village level (parentheses); bootstrapped standard errors clustered at the village level {braces}; and P-values for cluster wild bootstrap test of null that coefficient equals zero. * p < 0.1; ** p < .05; *** p < .01. Results are robust to out-of-sample testing; tables available on request.

Table A5: Characteristics of households in officials' villages who did not receive BISP pre-reform

	(1)	(2)
	All	Knows elected official
Female headed HH	0.030	0.031
Rudimentary house	0.306	0.316
No ag land	0.741	0.722
PML-N won	0.504	0.508
PPP won	0.203	0.237
Years in village	79.136	78.112
HH head ed 5 yrs+	0.462	0.510
HH head ed 8 yrs+	0.323	0.359
HH head ed 10 yrs+	0.217	0.244
No cattle	0.698	0.694
No res land	0.146	0.099
No assets (land or cattle)	0.109	0.058
Knows elected official	0.637	1
Observations	3861	2459

Notes: Sample of households in winning official's village who did not receive BISP. 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 elected official through personal interaction. Results are similar in testing sample; tables available on request.

Table A6: Mechanism test: 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	0.022	0.018		
	(0.011)*	(0.010)*		
Winner's village x winner's clan			-0.009	-0.008
			(0.011)	(0.010)
Winner's village x other clan			0.025	0.021
			(0.012)*	(0.011)*
HH control variables	X	X	X	X
N	7695	7571	7695	7571
P-values, cross-equation tests of equal coefficients:				
Winner's village		0.04		
Winner's village x other clan				0.03

Notes: Home village sample, post reform data only. Household controls include female-headed household; no agricultural land; rudimentary house; any daughters aged 18-25; no cattle; no residential land; household head 5th, 8th, 10th grade or higher; years household has lived in the village; years squared; any household member had an ID card issued from 2002-2007; winner's clan; and rival's clan. Robust standard errors clustered at the village level. * p < 0.1; ** p < .05; *** p < .01. Results are robust to out-of-sample testing; tables available on request.

Table A7: Mechanism test: ID cards

	(1)	(2)	(3)
	Dependent variable:		
	HH received new ID card	HH received BISP cash transfer	
Winner's village, winner's clan, pre-reform	0.032	0.032	0.005
	(0.007)***†‡	(0.007)***†‡	(0.001)***†‡
Winner's village, winner's clan, post-reform	0.004	-0.010	-0.014
	(0.010)	(0.009)	(0.004)***†
Winner's village, other clan, pre-reform	0.021	0.021	0.003
	(0.007)***†‡	(0.006)***†‡	(0.001)***†‡
Winner's village, other clan, post-reform	0.035	0.027	-0.002
	(0.012)***†‡	(0.012)**†‡	(0.004)
HH received new ID card			0.834
			(0.022)***
Observations	15406	15390	15390
Mean Y:			
Rival village, pre reform		0.01	
Rival village, post reform		0.05	
P-values:			
Winner's village, winner's clan, pre = post	0.03		
Winner's village, other clan, pre = post	0.25		
Winner's village, winner's clan, pre reform equal model 2 vs. 3			0.00
Winner's village, winner's clan, post reform equal model 2 vs. 3			0.66
Winner's village, other clan, pre reform equal model 2 vs. 3			0.00
Winner's village, other clan, post reform equal model 2 vs. 3			0.00

Notes: Home village sample. All specifications include controls for post reform, vote margin, female-headed household; no agricultural land; rudimentary house; any daughters aged 18-25; no cattle; no residential land; household head 5th, 8th, 10th grade or higher; years household has lived in the village; years squared; any household member had an ID card issued from 2002-2007; winner's clan; and rival's clan. * p < 0.1; ** p < .05; *** p < .01 (robust standard errors clustered at the village level); † p < 0.1 for bootstrapped standard errors clustered at the village level; ‡ p< 0.1 for cluster wild bootstrap by constituency. Results are robust to out-of-sample testing (Web Appendix C).

Table A8: Substitution with other programs

	(1)	(2)
	Total value of benefits - excluding BISP	
Winner's village x pre reform	-5.428	
	(30.543)	
Winner's village x post reform	51.083	
	(38.213)	
Winner's village x winner's clan x pre		-16.267
		(166.897)
Winner's village x other clan x pre		-4.367
		(27.103)
Winner's village x winner's clan x post		102.791
		(97.099)
Winner's village x other clan x post		45.520
		(38.001)
Mean Y:		
Rival village, pre reform	292.37	
Rival village, other clan, pre reform		310.23
Rival village, winner clan, pre reform		291.16
Observations	15406	15406
P-values:		
Winner's village, pre = post	0.09	
Winner's village, winner's clan, pre = post		0.18
Winner's village, other clan, pre = post		0.18
Winner's village, winner's clan = other clan, pre		0.94
Winner's village, winner's clan = other clan, post		0.54

Notes: Home 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 in PKR (100 PKR = 1 USD) based on administrative data on each program as described in the Pre-Analysis Plan. All specifications include controls for post reform, vote margin, female-headed household; no agricultural land; rudimentary house; any daughters aged 18-25; no cattle; no residential land; household head 5th, 8th, 10th grade or higher; years household has lived in the village; years squared; any household member had an ID card issued from 2002-2007; winner's clan; and rival's clan. (Parentheses: robust standard errors clustered at the village level.) * $p < 0.1$; ** $p < .05$; *** $p < .01$. Results are qualitatively similar but pre-post test is insignificant in training and testing subsamples; tables available on request.

Table A9: Impact of reform on differential misreporting on survey data: 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
Reform implemented	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 in Equation 2. Standard errors in parentheses, clustered at the district level. * p < .1, ** p < .05, *** p < .01.