Building State Capacity: Evidence from Biometric Smartcards in India*

Karthik Muralidharan† UC San Diego

Paul Niehaus‡ UC San Diego

Sandip Sukhtankar§ Dartmouth College

July 2, 2015

Abstract

Anti-poverty programs in developing countries are often difficult to implement; in particular, many governments lack the capacity to deliver payments securely to targeted beneficiaries. We evaluate the impact of biometrically-authenticated payments infrastructure ("Smartcards") on beneficiaries of employment (NREGS) and pension (SSP) programs in the Indian state of Andhra Pradesh, using a large-scale experiment that randomized the rollout of Smartcards over 157 subdistricts and 19 million people. We find that, while incompletely implemented, the new system delivered a faster, more predictable, and less corrupt NREGS payments process without adversely affecting program access. For each of these outcomes, treatment group distributions first-order stochastically dominated those of the control group. The investment was cost-effective, as time savings to NREGS beneficiaries alone were equal to the cost of the intervention, and there was also a significant reduction in the "leakage" of funds between the government and beneficiaries in both NREGS and SSP programs. Beneficiaries overwhelmingly preferred the new system for both programs. Overall, our results suggest that investing in secure payments infrastructure can significantly enhance "state capacity" to implement welfare programs in developing countries.

JEL codes: D73, H53, O30, O31

Keywords: state capacity, corruption, service delivery, biometric authentication, secure payments, electronic benefit transfers, public programs, NREGS, pensions, India

*We thank Santosh Anagol, Abhijit Banerjee, Julie Cullen, Gordon Dahl, Roger Gordon, Rema Hanna, Gordon Hanson, Erzo Luttmer, Santhosh Mathew, Simone Schaner, Monica Singhal, Anh Tran, and several seminar participants for comments and suggestions. We are grateful to officials of the Government of Andhra Pradesh, including Reddy Subrahmanyam, Koppula Raju, Shamsher Singh Rawat, Raghunandan Rao, G Vijaya Laxmi, AVV Prasad, Kuberan Selvaraj, Sanju, Kalyan Rao, and Madhavi Rani; as well as Gulzar Natarajan for their continuous support of the Andhra Pradesh Smartcard Study. We also thank officials of the Unique Identification Authority of India (UIDAI), including Nandan Nilekani, Ram Sevak Sharma, and R Srikar for their support, and Tata Consultancy Services (TCS) and Ravi Marri, Ramanna, and Shubra Dixit for their help in providing us with administrative data. This paper would not have been possible without the outstanding efforts and inputs of the J-PAL/IPA project team, including Vipin Awatramani, Kshitij Batra, Prathap Kasina, Piali Mukhopadhyay, Michael Kaiser, Raghu Kishore Nekanti, Matt Pecenco, Surili Sheth, and Pratibha Shrestha. We are deeply grateful to the Omidyar Network – especially Jayant Sinha, CV Madhukar, Surya Mantha, Ashu Sikri, and Dhawal Kothari – for the financial support and long-term commitment that made this study possible. We also thank IPA, Yale University, and the Bill and Melinda Gates Foundation for additional financial support through the Global Financial Inclusion Initiative.

†UC San Diego, JPAL, NBER, and BREAD. kamurali@ucsd.edu.

‡UC San Diego, JPAL, NBER, and BREAD. pniehaus@ucsd.edu.

§Dartmouth College, JPAL, and BREAD. sandip.sukhtankar@dartmouth.edu.

1 Introduction

Developing countries spend billions of dollars annually on anti-poverty programs, but the delivery of these programs is often poor and plagued by high levels of corruption (World Bank, 2003; Pritchett, 2010). It is therefore likely that investing in state capacity for better delivery of anti-poverty programs may have high returns in such settings. However, while a recent theoretical literature has highlighted the importance of investing in state capacity for economic development (Besley and Persson, 2009, 2010), there is limited empirical evidence on the returns to such investments.

One key constraint in the effective implementation of anti-poverty programs is the lack of a secure payments infrastructure to make transfers to intended beneficiaries. Often, money meant for the poor is simply stolen by officials along the way, with case studies estimating "leakage" of funds as high as 70 to 85 percent (Reinikka and Svensson, 2004; PEO, 2005). Thus, building a secure payments infrastructure can be seen as an investment in state capacity that could improve the implementation of existing anti-poverty programs, and also expand the state's long-term policy choice set.[1]

Recent technological advances have made it feasible to provide people with a biometrically-authenticated unique ID linked to bank accounts, which can be used to directly transfer benefits to intended recipients. Biometric technology is especially promising in developing country settings where high illiteracy rates constrain financial inclusion by precluding the universal deployment of traditional forms of authentication, such as passwords or PIN numbers.[2] The potential for such payment systems to improve the performance of public welfare programs (and also provide financial inclusion for the poor) has generated enormous interest around the world, with a recent survey documenting the existence of 230 programs in over 80 countries that are deploying biometric identification and payment systems (Gelb and Clark, 2013). This enthusiasm is exemplified by India's ambitious *Aadhaar* initiative to provide biometric-linked unique IDs (UIDs) to nearly a billion residents, and then transition social program payments to Direct Benefit Transfers via UID-linked bank accounts. Over 850 million UIDs have been issued as of June 2015, with the former Finance Minister of India claiming that the project would be "a game changer for governance" (Harris, 2013).

At the same time, there are a number of reasons to be skeptical about the hype around these new payment systems. First, their implementation entails solving a complex mix of technical and logistical challenges, raising the concern that the undertaking might fail unless all components are well-implemented (Kremer, 1993). Second, vested interests whose rents

[1] For instance, the ability to securely transfer income to poor households may make it more feasible for governments to replace distortionary commodity subsidies with equivalent income transfers.

[2] Fujiwara (2015) provides analogous evidence from Brazil on the effectiveness of electronic voting technology in circumventing literacy constraints, and on increasing enfranchisement of less educated voters.

are threatened may subvert the intervention and limit its effectiveness (Krusell and Rios-Rull, 1996; Prescott and Parente, 2000). Third, the new system could generate exclusion errors if genuine beneficiaries are denied payments due to technical problems. This would be particularly troubling if it disproportionately hurt the most vulnerable beneficiaries (Khera, 2011). Fourth, reducing corruption on some margins could displace it onto others (e.g. Yang (2008a)) or could paradoxically hurt the poor if it dampened incentives for officials to implement anti-poverty programs in the first place (Leff, 1964). Finally, even assuming positive impacts, cost-effectiveness is unclear as the best available estimates depend on a number of untested assumptions (see e.g. NIPFP (2012)). Overall, there is very limited evidence to support either the enthusiasts or the skeptics of biometric payment systems.

In this paper, we contribute toward filling this gap, by presenting evidence from a large-scale experimental evaluation of the impact of rolling out biometric payments infrastructure to make social welfare payments in India. Working with the Government of the Indian state of Andhra Pradesh (AP),[3] we randomized the order in which 157 sub-districts introduced a new "Smartcard" program for making payments in two large welfare programs: the National Rural Employment Guarantee Scheme (NREGS), and Social Security Pensions (SSP). NREGS is the largest workfare program in the world (targeting 800 million rural residents in India), but has well-known implementation issues including problems with the payment process and leakage (Dutta et al., 2012; Niehaus and Sukhtankar, 2013a,b). SSP programs complement NREGS by providing income support to the rural poor who are not able to work (Dutta et al., 2010). The new Smartcard-based payment system used a network of locally-hired, bank-employed staff to biometrically authenticate beneficiaries and make cash payments in villages. It thus provided beneficiaries of NREGS and SSP programs with the same effective functionality as intended by UID-linked Direct Benefit Transfers.

The experiment randomized the rollout of Smartcards over a universe of about 19 million people, with randomization conducted over entire sub-districts, making it one of the largest randomized controlled trials ever conducted. Evaluating an "as is" deployment of a complex program that was implemented at scale by a government addresses one common concern about randomized trials in developing countries: that studying NGO-led pilots may not provide accurate forecasts of performance at scales relevant for policy-making (see for example Banerjee et al. (2008); Acemoglu (2010); Bold et al. (2013)). The experiment thus provides an opportunity to learn about the likely impacts of India's massive UID initiative, as well as scaled-up deployments of biometric payments infrastructure more generally.

After two years of program rollout, the share of Smartcard-enabled payments across both programs in treated sub-districts had reached around 50%. This conversion rate over two

[3]The original state of AP (with a population of 85 million) was divided into two states on June 2, 2014. Since this division took place after our study, we use the term AP to refer to the original undivided state.

years compares favorably to the pace of electronic benefit transfer rollout in other contexts. For example, the United States took over 15 years to convert all Social Security payments to electronic transfers, while the Philippines took 5 years to reach about 40% coverage in a cash transfer program. On the other hand, the inability to reach a 100% conversion rate (despite the stated goal of senior policymakers to do so) reflects the non-trivial logistical, administrative, and political challenges of rolling out a complex new payment system (see section 3.3 and Mukhopadhyay et al. (2013) for details).

We therefore focus throughout the paper on intent-to-treat analysis, which correctly estimates the average return to as-is implementation following the "intent" to implement the new system. These estimates yield the relevant policy parameter of interest, because they reflect the impacts that followed a decision by senior government officials to invest in the new payments system and are net of all the logistical and political economy challenges that accompany such a project in practice.

We organize our analysis around three main dimensions of program performance: payments logistics, (prevention of) leakage, and program access. Beginning with payment logistics, we find that Smartcards delivered a faster and more predictable payment process for beneficiaries, especially under the NREGS program. NREGS workers spent 22 fewer minutes collecting each payment (20% less than the control group), and collected their payments 10 days sooner after finishing their work (29% faster than the control mean). The absolute deviation of payment delays also fell by 39% relative to the control group, suggesting that payments became more predictable. Payment collection times for SSP beneficiaries also fell, but the reduction was small and statistically insignificant.

Turning to leakage, we find that household NREGS earnings in treated areas increased by 24% while government outlays on NREGS did not change. The net result is a significant reduction in leakage of funds between the government and target beneficiaries. With a few further assumptions (see Section 4.2), we estimate a 12.7 percentage point reduction in NREGS leakage in treated areas (a 41% reduction relative to the control mean). Similarly, SSP benefit amounts increased by 5%, with no corresponding change in government outlays, resulting in a significant reduction in SSP leakage of 2.8 percentage points (a 47% reduction relative to the control mean).

These gains for participants on the intensive margin of program performance were not offset by reduced access to programs on the extensive margin. We find that the proportion of households reporting having worked on NREGS *increased* by 7.1 percentage points (a 17% increase over the control mean of 42%). We show that this result is explained by a significant reduction in the fraction of "quasi-ghost" beneficiaries - defined as cases where officials reported work against a beneficiary's name and claimed payments for this work, but where the beneficiary received neither work nor payments. These results suggest that the introduction of biometric authentication made it more difficult for officials to over-report the amount of work done (and siphon off the extra wages unknown to the beneficiary), and that the optimal response for officials was to ensure that more actual work was done against the claimed wages, with a corresponding increase in payments made to workers. We find no impact on access to pensions, with the rate of SSP enrollment unchanged.

We also examine the distribution of impacts on each margin of performance. We find no evidence that poor or vulnerable segments of the population were made worse off by the new system. For each dimension of performance with significant positive average impacts, treatment distributions first-order stochastically dominate control distributions. Thus, no treatment household was worse off relative to a control household at the same percentile of the outcome distribution. Treatment effects also did not vary significantly as a function of village-level baseline characteristics, suggesting broad-based gains across villages from access to the new payments system.

The Smartcards intervention introduced two main sets of changes to the payments process. First, it changed the organizations responsible for making payments and moved the point of payment closer to the village. Second, it introduced biometric authentication. In a non-experimental decomposition of the treatment effects, we find that improvements in the timeliness of payments are concentrated entirely in villages that switched to the new payment system, but do not vary within these villages across recipients who had or had not received biometric Smartcards. In contrast, increases in payments to beneficiaries and reductions in leakage are concentrated entirely among recipients who actually received biometric Smartcards. This suggests that organizational changes associated with the new payment system drove improvements in the payments process, while biometric authentication was key to reducing fraud.

Overall, the data suggest that Smartcards improved beneficiary experiences in collecting payments, increased payments received by intended beneficiaries, reduced corruption, broadened access to program benefits, and achieved these without substantially altering fiscal burdens on the state. Consistent with these findings, 90% of NREGS beneficiaries and 93% of SSP recipients who experienced Smartcard-based payments reported that they prefer the new system to the old.

Finally, Smartcards appear to be cost-effective. In the case of NREGS, our best estimate of the value of beneficiary time savings ($4.5 million) *alone* exceeds the government's cost of program implementation and operation ($4 million). Further, our estimated NREGS leakage reduction of $38.5 million/year is over nine times greater than the cost of implementing the new Smartcard-based payment system. While gains in the SSP program are more modest, the estimated leakage reduction of $3.2 million/year is still higher than the costs of the program ($2.3 million). The reductions in leakage represent redistribution from corrupt officials to beneficiaries, and are hence not Pareto improvements. However, if a social planner places a greater weight on the gains to program beneficiaries (likely to be poorer) than on the loss of illegitimate rents to corrupt officials, the welfare effects of reduced leakage will be positive.

The first contribution of our paper is as an empirical complement to the recent theoretical literature highlighting the importance of state capacity for economic development (Besley and Persson, 2009, 2010).[4] However, despite the high potential social returns to investing in public goods such as general-purpose implementation capacity, both theory and evidence suggest that politicians may underinvest in these relative to specific programs that provide patronage to targeted voter and interest groups (Lizzeri and Persico, 2001; Mathew and Moore, 2011). Further, politicians may perceive the returns to such investments as accruing in the long-run, while their own electoral time horizon may be shorter. Our results suggest that in settings of weak governance, the returns to investing in implementation capacity can be positive and large over as short a period as two years.[5]

We also contribute to the literature on identifying effective ways to reduce corruption in developing countries (Reinikka and Svensson, 2005; Olken, 2007). Our results highlight the potential for technology-enabled top-down improvements in governance to reduce corruption. They may also help to clarify the literature on technology and service delivery in developing countries, where an emerging theme is that technology may or may not live up to its hype. Duflo et al. (2012) find, for example, that time-stamped photos and monetary incentives increased teacher attendance and test scores in Indian schools (when implemented in schools run by an NGO). Banerjee et al. (2008) find, on the other hand, that a similar initiative to monitor nurses was subverted by vested interests when a successful NGO-initiated pilot program was transitioned to being implemented by the local government. Our results, which describe the effects of an intervention driven from the start by the government's own initiative, suggest that technological solutions *can* significantly improve service delivery when implemented as part of an institutionalized policy decision to do so at scale.

Finally, our results complement a growing literature on the impact of payments and authentication infrastructure in developing countries. Jack and Suri (2014) find that the MPESA mobile money transfer system in Kenya improved risk-sharing; Aker et al. (2013) find that using mobile money to deliver transfers in Niger cut costs and increased women's

[4]Note that political scientists also use the term "state capacity" to represent the set of formal institutions that adjudicate conflicting claims in societies (including legislatures, and judiciaries). Besley and Persson (2010) focus on fiscal and legal state capacity, but do not distinguish the legislative and executive aspects of such capacity. In practice, the poor implementation of existing laws, regulations, and policies in developing countries (including widespread tax evasion and leakage in spending), suggest that the executive side of state capacity is an important constraint in these settings. This is what our study focuses on.

[5]While set in a different sector, the magnitude of our estimated reduction in leakage is consistent with recent evidence from India showing that investing in better monitoring of teachers may yield a tenfold reduction in the cost of teacher absence (Muralidharan et al., 2014). Dal Bó et al. (2013) present complementary evidence on the impact of raising public sector salaries on the quality of public sector workers hired.

intra-household bargaining power; and Gine et al. (2012) show how biometric authentication helped a bank in Malawi reduce default and adverse selection.

From a policy perspective, our results contribute to the ongoing debates in India and other developing countries regarding the costs and benefits of using biometric payments technology for service delivery. We discuss the policy implications of our results and caveats to external validity in the conclusion.

The rest of the paper is organized as follows. Section 2 describes the context, social programs, and the Smartcard intervention. Section 3 describes the research design, data, and implementation details. Section 4 presents our main results. Section 5 discusses cost-effectiveness. Section 6 concludes. We also include an extensive online Appendix with supplemental program details and analysis.

2 Context and Intervention

As the world's largest democracy, India has sought to reduce poverty through ambitious welfare schemes. Yet these schemes are often poorly implemented (Pritchett, 2010) and prone to high levels of corruption or "leakage" as a result (PEO, 2005; Niehaus and Sukhtankar, 2013a,b). Benefits that do reach the poor arrive with long and variable lags and are time-consuming for recipients to collect. The AP Smartcard Project aimed to address these problems by integrating new payments infrastructure into two major social welfare programs managed by the Department of Rural Development, which serve as a comprehensive safety net for both those able (NREGS) and unable (SSP) to work. This section provides a concise description of these programs and how the introduction of Smartcards altered their implementation (further details are provided in Appendix A).

2.1 The National Rural Employment Guarantee Scheme

The NREGS is one of the two main welfare schemes in India and the largest workfare program in the world, covering 11% of the world's population. The Government of India's allocation to the program for fiscal year April 2013-March 2014 was Rs. 330 billion (US $5.5 billion), or 7.9% of its budget.[6] The program guarantees every rural household 100 days of paid employment each year. There are no eligibility requirements, as the manual nature of the work is expected to induce self-targeting.

Participating households obtain jobcards, which list household members and have empty spaces for recording employment and payment. Jobcards are issued by the local Gram

[6]NREGS figures: `http://indiabudget.nic.in/ub2013-14/bag/bag5.pdf`; total outlays: `http://indiabudget.nic.in/ub2013-14/bag/bag4.pdf`, both accessed June 23, 2015.

Panchayat (GP, or village) or mandal (sub-district) government offices. Workers with job-cards can apply for work at will, and officials are legally obligated to provide either work on nearby projects or unemployment benefits (though, in practice, the latter are rarely provided). NREGS projects vary somewhat but typically involve minor irrigation work or improvement of marginal lands. Project worksites are managed by officials called Field Assistants, who record attendance and output on "muster rolls" and send these to the sub-district for digitization, from where the work records are sent up to the state level, which triggers the release of funds to pay workers.

Figure A.1a depicts the payment process in AP prior to the introduction of Smartcards. The state government transfers money to district offices, which pass the funds to mandal offices, which transfer it to beneficiary post office savings accounts. Workers withdraw funds by traveling to branch post offices, where they establish identity using jobcards and passbooks. In practice it is common for workers (especially illiterate ones) to give their documents to Field Assistants who then control and operate their accounts – taking sets of passbooks to the post office, withdrawing cash in bulk, and returning to distribute it in villages.

Issues of payments logistics, leakage, and access have all dogged NREGS implementation. Both prior research (Dutta et al., 2012) and data from our control group suggest that even conditional on doing NREGS work, the payment process is slow and unreliable, limiting the extent to which the NREGS can effectively insure the rural poor.[7] In extreme cases, delayed payments have reportedly led to worker suicides (Pai, 2013).

The payments process is also vulnerable to leakage of two forms: over-reporting and under-payment. Consider a worker who has earned Rs. 100, for example: the Field Assistant might report that he is owed Rs. 150 but pay the worker only Rs. 90, pocketing Rs. 50 through over-reporting and Rs. 10 through under-payment. Two extreme forms of over-reporting are "ghost" workers who do not exist, but against whose names work is reported and payments are made; and "quasi-ghost" workers who do exist, but who have not received any work or payments though work is reported against their names and payments are made. In both cases, the payments are typically siphoned off by officials. Prior work in the same context suggests that over-reporting is the most prevalent form of leakage - perhaps because it involves stealing from a "distant" taxpayer, and can be done without the knowledge of workers (Niehaus and Sukhtankar, 2013a).[8]

Finally, program access is imperfect, although by design NREGS work and payments should be constrained only by worker demand. In practice, supply appears to be the binding

[7] Imperfect implementation of social insurance programs may even be a deliberate choice by local elites to preserve their power over the rural poor, as these elites are often the default providers of credit and insurance. See Anderson et al. (2015) for discussion, and also Jayachandran (2006) who shows how uninsured rainfall shocks benefit landlords and hurt workers (especially those who lack access to credit).

[8] A growing literature has examined over-invoicing as a form of corruption and the effects of government policies on it. See Fisman and Wei (2004); Olken (2007); Yang (2008b); Mishra et al. (2008), among others.

constraint, with NREGS availability being constrained by the level of budgetary allocations and by limited local administrative capacity and willingness to implement projects (Dutta et al., 2012; Witsoe, 2014). We confirm this in our data, where less than 4% of workers in our control group report that they can access NREGS work whenever they want it.

2.2 Social Security Pensions

Social Security Pensions are unconditional monthly payments targeted to vulnerable populations. The program covers over 6 million beneficiaries and costs the state of AP roughly Rs. 18 billion ($360 million) annually. Eligibility is restricted to members of families classified as Below the Poverty Line (BPL) who are residents of the district in which they receive their pension and not covered by any other pension scheme. In addition, recipients must qualify in one of four categories: old age (> 65), widow, disabled, or certain displaced traditional occupations. Pension lists are proposed by village assemblies (Gram Sabhas) and sanctioned by the mandal administration. Pensions pay Rs. 200 (˜$3) per month except for disability pensions, which pay Rs. 500 (˜$8). Unlike the NREGS, pension payments are typically disbursed in the first week of each month in the village itself by a designated village development officer.

The SSP program appears to be better implemented than NREGS. Dutta et al. (2010) find that it is well targeted with relatively low levels of leakage (about 17% in Karnataka, less than half the rate found in comparable programs). We also did not find documented evidence on beneficiary complaints regarding the SSP payment process. This is likely to be because it is a straightforward process, with a mostly fixed list of beneficiaries who receive a fixed amount of payment at a fixed time every month. Our pilots corroborated this view of the SSP payments process, and we therefore did not collect data on payment delays.

2.3 Smartcard-enabled Payments

The Smartcard project was India's first large-scale attempt to implement a biometric payments system.[9] It was a composite intervention, modifying NREGS and SSP payment systems in multiple ways, which we think of as comprising two complementary but conceptually distinct bundles: one set of technological changes, and one set of organizational ones.

Technologically, the intervention changed the way in which beneficiaries were expected to establish their identity when collecting payments. Under the status quo, beneficiaries proved identity by exhibiting identifying documents to the agent issuing payments, who was

[9]The central (federal) government had similar goals for the Aadhaar (UID) platform. However, the initial rollout of Aadhaar was as an enabling infrastructure, and it had not yet been integrated into any of the major welfare schemes as of June 2014. The Smartcard intervention can therefore be seen as a functional precursor to the integration of Aadhaar into the NREGS and SSP.

responsible for verifying these. Under the Smartcards scheme, biometric data (typically all ten fingerprints) and digital photographs were collected during enrollment campaigns and linked to newly created bank accounts. Beneficiaries were then issued a physical "Smartcard" that included their photograph and (typically) an embedded electronic chip storing biographic, biometric, and bank account details. Beneficiaries use these cards to collect payments as follows: (a) they insert them into a Point-of-Service device operated by a Customer Service Provider (CSP), which reads the card and retrieves account details; (b) the device prompts for one of ten fingers, chosen at random, to be scanned; (c) the device compares this scan with the records on the card, and authorizes a transaction if they match; (d) the amount of cash requested is disbursed;[10] and (e) the device prints out a receipt (and in some cases announces transaction details in the local language, Telugu). Figure A.2 shows a sample Smartcard and a fingerprint scan in progress.[11]

Organizationally, the intervention changed the vendors and staff responsible for delivering payments. The government contracted with banks to manage payments for both schemes, and these banks in turn contracted with Technology Service Providers (TSPs) to manage the last-mile logistics of delivery; the TSPs then hired and trained CSPs. Figure A.1b illustrates the flow of funds from the government through banks, TSPs and CSPs to beneficiaries under this scheme. The government assigned each district to a single bank-TSP pairing, and compensated them with a 2% commission on all payments delivered in GPs that were migrated to the new Smartcard-based payment system (banks and TSPs negotiated their own terms on splitting the commission). The government required a minimum of 40% beneficiaries in a GP to be enrolled and issued Smartcards prior to converting the GP to the new payment system; this threshold applied to each program separately. Once a GP was "converted", all payments - for each program in which the threshold was reached - in that GP were routed through the Bank-TSP-CSP system (even for beneficiaries who had not enrolled in or obtained Smartcards).

The government also stipulated norms for CSP selection, and required that CSPs be women resident in the villages they served, have completed secondary school, not be related to village officials, preferably be members of historically disadvantaged castes, and be members of a self-help group.[12] While meeting all these requirements was often difficult and sometimes impossible, the selected CSPs were typically closer socially to beneficiaries than the post-office officials or village development officers (both government employees) who previously

[10] While beneficiaries could in principle leave balances on their Smartcards and thus use them as savings accounts, NREGS guidelines required beneficiaries to be paid in full for each spell of work. As a result the default expectation was that workers would withdraw their wages in full.

[11] Note that a truly "smart" card was not required or always issued: one Bank chose to issue paper cards with digital photographs and bar codes while storing biometric data in the Point-of-Service device (as opposed to on the card). Authentication in this system was otherwise the same.

[12] Self-help groups are groups of women organized by the government to facilitate micro-lending.

disbursed payments (for NREGS and SSP respectively). Moreover, because CSPs were stationed within villages they were also geographically closer to beneficiaries.

2.4 Potential Impacts of Smartcards

Taken as a whole, the Smartcards intervention constituted a significant change to the authentication and payments process in NREGS and SSP programs, and could have affected program performance on multiple dimensions. To help structure the analysis that follows we organize it around three main dimensions of impact: payments logistics, leakage, and program access.

First, payments logistics could improve or deteriorate. Smartcards could speed up payments, for example, by moving transactions from the (typically distant) post office to a point within the village. They could just as easily slow down the process, however, if CSPs were less reliably present or if the checkout process were slower due to technical problems.[13] Similarly, on-time cash availability could either improve or deteriorate depending on how well TSPs managed cash logistics relative to the post office. In a worst-case scenario the intervention could cut off payments to beneficiaries who were unable to obtain cards, lost their cards, or faced malfunctioning authentication devices.[14]

Second, leakage might or might not decrease. In principle, Smartcards should reduce payments to "ghost" beneficiaries as ghosts do not have fingerprints, and also make it harder for officials to collect payments in the name of real beneficiaries as they must be present, provide biometric input, and receive a receipt which they can compare to the amount disbursed. These arguments assume, however, that the field technology works as designed and that CSPs are not more likely to be corrupt than local GP officials and post office workers. Moreover, achieving significant leakage reductions might require near complete implementation and yet the intervention was complex enough that complete implementation was unlikely.[15]

Finally, program access could also improve or suffer. In the case of NREGS, reducing rents may reduce local officials' incentives to create and implement projects, which could reduce access. On the other hand, a reduction in officials' incentives to over-report work done (because the money now goes directly to beneficiaries) might induce them to increase the actual amount of work done (to better correspond to the inflated muster rolls), which could increase access to NREGS. In other words, if Smartcards make it more difficult for

[13] For example, case-study based evidence suggests that manual payments were faster than e-payments in Uganda's cash transfer program (CGAP, 2013).

[14] The tension here between reducing fraud and excluding genuine beneficiaries is an illustration of the general trade-off between making Type I (exclusion) and Type II (inclusion) errors in public welfare programs (see Dahl et al. (2014) for a discussion in the context of adjudicating claims of disability insurance).

[15] Specifically, leakage reduction may be convex in the extent of coverage if those who enroll for Smartcards are genuine workers, and if the non-enrollees are the ghosts. In such a setting, there may be limited impact on leakage reduction unless Smartcard coverage is near complete and uncarded payments are stopped.

officials to siphon off funds, more of these funds could be available for actual work and may lead to NREGS implementation becoming closer to what the program framers intended (with more work, more payments to workers, and more rural assets created). In the case of SSP, reducing leakage could drive up the illicit price of getting on the SSP beneficiary list.

The Smartcards intervention included both *technological* and *organizational* innovations: we present a non-experimental decomposition of the relative contribution of these two components in section 4.6. Finally, we present results for NREGS and SSP programs in parallel to the extent possible, but there is no deep economic reason to treat them similarly or expect similar impacts because the nature of the programs and pre-existing quality of implementation were quite different.[16]

3 Research Design

3.1 Randomization

The AP Smartcard project began in 2006, but took time to overcome initial implementation challenges including contracting, integration with existing systems, planning the logistics of enrollment and cash management, and developing processes for financial reporting and reconciliation. Because the government contracted with a unique bank to implement the project within each district, and because multiple banks participated, considerable heterogeneity in performance across districts emerged over time. In eight of twenty-three districts the responsible banks had made very little progress as of late 2009; in early 2010 the government decided to restart the program in these districts, and re-allocated their contracts to banks that had implemented Smartcards in other districts. This "fresh start" created an attractive setting for an experimental evaluation of Smartcards for two reasons. First, the roll-out of the intervention could be randomized in these eight districts. Second, the main implementation challenges had already been solved in other districts, yielding a "stable" implementation model prior to the evaluation.

Our evaluation was conducted in these eight districts (see Figure C.1), which have a combined rural population of around 19 million. While not randomly selected, they look similar to AP's remaining 13 non-urban districts on major socioeconomic indicators, including proportion rural, scheduled caste, literate, and agricultural laborers (see Appendix D.1). They also span the state geographically, with representation in all three historically distinct socio-cultural regions: 2 in Coastal Andhra and 3 each in Rayalseema and Telangana.

The study was conducted under a formal agreement between J-PAL South Asia and the Government of Andhra Pradesh (GoAP) to randomize the order in which mandals (sub-

[16]The NREGS and SSP programs are both part of the experiment only because they are both run by the AP Department of Rural Development, which led the AP Smartcard initiative.

districts) were converted to the Smartcard system. We assigned a total of 296 mandals to treatment and control status by lottery as follows: 112 mandals were assigned to the treatment group, 139 to a "buffer" group, and 45 to a control group (Figure C.1).[17] We collected survey data only in the treatment and control groups; we created the buffer group to ensure we would have time to conduct endline surveys after Smartcards had been deployed in the treatment mandals but before they were deployed in the control mandals (during which period, enrollment could take place in the buffer group without affecting the control group). The realized lag between program rollout in treatment and control mandals was over two years. Randomization was stratified by district and by a principal component of socio-economic characteristics. Table C.1 presents tests of equality between treatment and control mandals along characteristics used for stratification, none of which (unsurprisingly) differ significantly. Table C.2 reports balance along all of our main outcomes as well as key socio-economic household characteristics from the baseline survey; three of 28 differences for NREGS and two of seventeen for SSP are significant at the 10% level. In the empirical analysis we include specifications that control for the village-level baseline mean value of our outcomes to test for sensitivity to any chance imbalances.

3.2 Data Collection

Our data collection was designed to capture impacts broadly, including both anticipated positive and negative effects; full details are provided in Appendix B. We first collected official records on beneficiary lists and benefits paid, and then conducted detailed baseline and endline household surveys of samples of enrolled participants. Household surveys included questions on receipts from and participation in the NREGS and SSP as well as questions about general income, employment, consumption, and assets. We conducted surveys in August through early October of 2010 (baseline) and 2012 (endline) in order to obtain information about NREGS participation between late May and early July of those years, as this is the peak period of participation in most districts (see Figure 1).[18] The intervention was rolled out in treatment mandals shortly after baseline surveys. We also conducted unannounced audits of NREGS worksites during our endline surveys to independently verify the number of workers who were present.

[17]Note that there were a total of 405 mandals in the eight study districts, but we excluded 109 mandals from the universe of our study (mainly because Smartcard enrollment had started in these mandals before the agreement with GoAP was signed). The remaining 296 mandals comprised the universe of our study and randomization. See Appendix C.1 for full details on the randomization, and D.3 for comparisons between the 109 non-study mandals and the 296 study mandals.

[18]There is a tradeoff between surveying too soon after the NREGS work was done (since payments would not have been received yet), and too long after (since recall problems might arise). We surveyed on average 10 weeks after work was done, and also facilitated recall by referring to physical copies of jobcards (on which work dates and payments are meant to be recorded) during interviews.

Full details and discussion of the sampling procedure used are in Appendix C.2. In brief, we sampled 880 GPs in which to conduct surveys. Within each GP we sampled 10 households, 6 from the frame of NREGS jobcard holders and 4 from the frame of SSP beneficiaries. Our NREGS sample included 5 households in which at least one member had worked during May-June according to official records and one household in which no member had worked. This sampling design trades off power in estimating leakage (for which households reported as working matter) against power in estimating rates of access to work (for which all households matter). For our endline survey we sampled 8,774 households, of which we were unable to survey or confirm existence of 295, while 365 households were confirmed as ghost households, leaving us with survey data on 8,114 households (corresponding numbers for baseline are 8,572, 1,000, 102, and 7,425 respectively).

The resulting dataset is a panel at the village level and a repeated cross-section at the household level. This is by design, as the endline sample should be representative of potential participants at that time. We verify that the treatment did not affect either the size or composition of the sampling frame (Appendix C.3), suggesting that our estimated treatment effects are not confounded by changes in the composition of potential program beneficiaries.

3.3 Implementation, First-Stage, and Compliance

We present a brief description of the implementation of the Smartcard project and the extent of actual roll-out to help interpret our results better. As may be expected, the implementation of such a complex project faced a number of technical, logistical, and political challenges. Even with the best of intentions and administrative attention, the enrollment of tens of millions of beneficiaries, physical delivery of Smartcards and Point-of-Service devices, identification and training of CSPs, and putting in place cash management protocols would have been a non-trivial task. In addition, local officials (both appointed and elected) who benefited from the status quo system had little incentive to cooperate with the project, and it is not surprising that there were attempts to subvert an initiative to reduce leakage and corruption (as also described in Banerjee et al. (2008)). In many cases, local officials tried to either capture the new system (for instance, by attempting to influence CSP selection), or delay its implementation (for instance, by citing difficulties to beneficiaries in accessing their payments under the new system).

On the other hand, senior officials of GoAP were strongly committed to the project, and devoted considerable administrative resources and attention to successful implementation. More generally, GoAP was strongly committed to NREGS and AP was a leader in utilization of federal funds earmarked for the program. Overall, implementation of the Smartcard Program was a priority for GoAP, but it faced an inevitable set of challenges. Our estimates therefore reflect the impacts of a policy-level decision to implement the Smartcard project at scale, and is net of all the practical complexities of doing so.

Figure 2 plots program rollout in treatment mandals from 2010 to 2012 using administrative data. Clearly, implementation was incomplete. By July 2012, 82% (89%) of treatment group mandals were "converted" (defined as having converted at least one GP) for NREGS (SSP) payments. Conditional on being in a converted mandal, 83% (93%) of GPs had converted for NREGS (SSP) payments, where being "converted" meant that payments were made through the new Bank-TSP-CSP system. These payments could include authenticated payments, unauthenticated payments to workers with Smartcards, and payments to workers without Smartcards.[19] Payments made to beneficiaries with Smartcards ("carded payments," both authenticated and unauthenticated) made up about two-thirds of payments within converted GPs by the endline. All told, about 50% of payments in treatment mandals across both programs were "carded" by May 2012.[20]

Turning to compliance with the experimental design, we see that sampled GPs in treated mandals were much more likely to have migrated to the new payment system, with 67% (79%) being "carded" for NREGS (SSP) payments, compared to 0.5% (0%) of sampled control GPs (Table 1). The overall rate of transactions done with carded beneficiaries was 45% (59%) in treatment areas, with no carded transactions reported in control areas. We can also assess compliance using data from our survey, which asked beneficiaries about their Smartcard use. About 38% (45%) of NREGS (SSP) beneficiaries in treated mandals said that they used their Smartcards both generally or recently, while 1% (4%) claimed to do so in control mandals. This latter figure likely reflects some beneficiary confusion between enrollment (the process of capturing biometrics and issuing cards) and the onset of carded transactions themselves, as the government did not allow the latter to begin in control areas until after the endline survey. Note that official and survey figures are not directly comparable since the former describe *transactions* while the latter describe *beneficiaries*.

Overall, both official and survey records indicate that Smartcards were operational albeit incompletely in treatment areas, with minimal contamination in control areas. We therefore focus on intent-to-treat (ITT) estimates which can be interpreted as the average treatment effects corresponding to an approximately half-complete implementation.[21] It is important

[19]Transactions may not be authenticated for a number of reasons, including failure of the authentication device and non-matching of fingerprints.

[20]There was considerable heterogeneity in the extent of Smartcard coverage across the eight study districts, with coverage rates ranging from 31% in Adilabad to nearly 100% in Nalgonda district. Thus, we focus our analysis on ITT effects, and all our estimates include district fixed effects. We present correlates of implementation heterogeneity in Appendix D, and provide a qualitative discussion of implementation heterogeneity in a companion study (Mukhopadhyay et al., 2013).

[21]Note that given implementation heterogeneity across districts and the possibility of non-linear treatment effects in the extent of Smartcard coverage, our results should be interpreted as the average treatment effect across districts with different levels of implementation (averaging to around 50% coverage) and not as the impact of a half-complete implementation in all districts.

to note, however, that the 50% rate of Smartcard coverage achieved in two years compares favorably with the performance of changes in payments processes elsewhere. For example, a conditional cash transfer program in the Phillippines (4Ps) took 5 years to reach 40% coverage (2008-13) (Bohling and Zimmerman, 2013).

3.4 Estimation

We report ITT estimates, which compare average outcomes in treatment and control areas. All outcomes are estimated at the individual beneficiary level for SSP, and at the level which they were asked - individual, individual by week, or household - for NREGS, unless aggregation is necessary in order to compare with official data. All regressions are weighted by inverse sampling probabilities to obtain average partial effects for the populations of NREGS jobcard holders or SSP beneficiaries. We include district fixed effects and the first principal component of a vector of mandal characteristics used to stratify randomization (PC_{md}) in all regressions, and cluster standard errors at the mandal level. We thus estimate

$$Y_{imd} = \alpha + \beta Treated_{md} + \delta District_d + \lambda PC_{md} + \epsilon_{imd} \tag{3.1}$$

where Y_{imd} is an outcome for household or individual i in mandal m and district d, and $Treated_{md}$ is an indicator for a mandal in wave 1. When possible, we also report specifications that include the baseline GP-level mean of the dependent variable, $\overline{Y}^0_{pmd}$, to increase precision and assess sensitivity to any randomization imbalances. We then estimate

$$Y_{ipmd} = \alpha + \beta Treated_{md} + \gamma \overline{Y}^0_{pmd} + \delta District_d + \lambda PC_{md} + \epsilon_{ipmd} \tag{3.2}$$

where p indexes panchayats or GPs. Note that we easily reject $\gamma = 1$ in all cases and therefore do not report difference-in-differences estimates.

4 Effects of Smartcard-enabled Payments

4.1 Effects on Payment Logistics

Data from our control group confirm that NREGS payments are typically delayed. Recipients in control mandals waited an average of 34 days after finishing a given spell of work to collect payment, more than double the 14 days prescribed by law (Table 2). The collection process is also time-consuming, with the average recipient in the control group spending almost two hours traveling and waiting in line to collect a payment.

Smartcards substantially improved this situation. The total time required to collect a NREGS payment fell by 22 minutes in mandals assigned to treatment (20% of the control mean). Time to collect payments also fell for SSP recipients, but the reduction is not statistically significant (Table 2; columns 1-2 for NREGS, columns 3-4 for SSP). We also find that over 80% of both NREGS and SSP beneficiaries who had received or enrolled for Smartcards reported that Smartcards had sped up payments (Table 6).

NREGS recipients also faced shorter delays in receiving payments after working, and these lags became more predictable. Columns 5 and 6 of Table 2 report that assignment to treatment lowered the mean number of days between working and collecting NREGS payments by 10 days, or 29% of the control mean (and 50% of the amount by which this exceeds the statutory limit of 14 days). There is also suggestive evidence that uncertainty about the timing of payments fell. While we do not directly measure beliefs, columns 7 and 8 show that the *variability* of payment lags – measured as the absolute deviation from the median mandal level lag, thus corresponding to a robust version of a Levene's test – fell by 39% of the control mean. This reduced variability is potentially valuable for credit-constrained households that need to match the timing of income and expenditure.[22]

4.2 Effects on Payment Amounts and Leakage

Recipients in treatment mandals also received more money. For NREGS recipients, columns 3 and 4 of Table 3a show that earnings per week during our endline study period increased by Rs. 35, or 24% of the control group mean. For SSP beneficiaries, earnings per beneficiary during the three months preceding our endline survey (May-July) increased by Rs. 12, or 5% of the control mean. In contrast, we see no impacts on fiscal outlays. For the workers sampled into our endline survey, we find no significant difference in official NREGS disbursements between treatment and control mandals. Similarly, SSP disbursements were also unaltered (columns 1 and 2 of Tables 3a and 3b respectively).

The fact that recipients report receiving more while government outlays are unchanged implies a reduction in leakage on both programs. Columns 5 and 6 of Table 3a confirm that the difference between official and survey measures of earnings per week on NREGS fell significantly by Rs. 25.[23] Results on the SSP program mirror the NREGS results: we find a reduction in leakage of Rs. 7 per pension per month. This represents a 2.8 percentage point reduction in leakage relative to fiscal outlays, which is a 47% reduction relative to the control mean (Table 3b).

While we find evidence of a significant reduction in NREGS leakage, estimating the magnitude of this reduction as a proportion of average leakage requires additional data. We cannot

[22]We did not collect analogous data on date of payment from SSP beneficiaries as payment lags had not surfaced as a major concern for them during initial fieldwork.

[23]Note that because we estimate results in a representative sample of jobcards, they are not affected by changes in the extensive margin of participation in or payment for the program.

simply compare what a given *household* reports receiving to what the government reported paying out on the *jobcard* based on which that household was sampled because, in practice, many households have more than one jobcard listed in their name.[24] Using official records to count the total number of jobcards in our study districts, and data from the 68th round of the National Sample Survey (July 2011-June 2012) to estimate the number of households in those districts with at least one jobcard, we calculate that the number of jobcards exceeds the number of households with jobcards by an average factor of 1.9. This implies that we will substantially under-estimate leakage if we do not account for multiple jobcards. Indeed, Table 3a shows that the naive estimate for the control group is a *negative* leakage rate of Rs. 20 per week.

To obtain a consistent estimate of average leakage we need to correct for multiple jobcards. We do so as follows: we scale up official records of payments issued in each district by the estimated number of jobcards per jobcard-holding household in that district, calculated as above. We then compare average amount disbursement per household (as opposed to per jobcard) to the average received per household. Using this method we estimate an endline leakage rate of 30.7% in control areas and 18% in treatment areas (Table E.1), implying that Smartcards reduced leakage by roughly 41%.[25]

4.2.1 Margins of Leakage Reduction

We examine leakage reduction along the three margins discussed earlier (ghosts, over-reporting, and under-payment), and find that reduced over-reporting appears to be the main driver of lower NREGS leakage. Reductions in NREGS ghost beneficiaries are insignificant, though the incidence of ghosts is a non-trivial 11% (Table 4a, columns 1-2). This is not surprising given the incomplete coverage of Smartcards, and the government's political decision to not ban unauthenticated payments. Thus, beneficiary lists were not purged of ghosts, and payments to these jobcards are likely to have continued. We also find limited impact on under-payment, measured as whether a bribe had to be paid to collect payments (Table 4a, columns 5 and 6). As we find little evidence of under-payment to begin with (control group incidence rate of 2.6%), Smartcards may have limited incremental value on this margin.

However, over-reporting in the NREGS drops substantially, with the proportion of jobcards that had positive official payments reported but zero survey amounts (excluding ghosts) dropping significantly by 8.4 percentage points, or 32% (Table 4a, columns 3-4). This result is mirrored in Figure 3, which presents quantile treatment effect plots on official and survey

[24]This issue is not solved by only including survey reports of individuals listed on the sampled jobcard - which we indeed do - since payments made to those individuals may be listed on other jobcards.

[25]However, this procedure leads to a loss of precision, as scaling up by a constant increases variance by the square of the constant (p-value 0.11). Appendix E.1 provides more detail on this procedure as well as an example to illustrate how the multiple-jobcard issue affects our calculations.

payments; here we see (a) no change in official payments at any part of the distribution, (b) a significant reduction in the incidence of beneficiaries reporting receiving zero payments, and (c) no significant change in amounts received relative to control households who were reporting positive payments.

These results suggest that leakage reduction was mainly driven by a reduction in the incidence of "quasi-ghosts": real beneficiaries who did not previously get any NREGS work or payments, though officials were reporting work and claiming payments on their behalf. If some of these households were to have enrolled for a Smartcard, it would no longer be possible for officials to siphon off payments without their knowledge, following which officials' optimal response appears to have been to provide actual work and payments to these households (see results on access below). A similar decomposition of the reduction in SSP leakage (Table 4b), reveals a reduction in all three forms of leakage, suggesting that Smartcard may have improved SSP performance on all dimensions (though none of the individual margins are significant).

The reduction in NREGS over-reporting raises an additional question: If Smartcards reduced officials' rents on NREGS, why did they not increase the total amounts claimed (perhaps by increasing the number of ghosts) to make up for lost rents? Conversations with officials suggest that the main constraint in doing so was the use of budget caps within the NREGS in AP that exogenously fixed the maximum spending on the NREGS for budgeting purposes (also reported by Dutta et al. (2012)). If enforced at the local level, these caps would limit local officials' ability to increase claims in response to Smartcards.

While we cannot directly test this, our result finding no significant increase in official payments in treated areas (Table 3a) holds even when we look beyond our study period and sampled GPs. Figure 1 shows the evolution of official disbursements in *all* GPs in treatment and control mandals, and for *every* week in 2010 and 2012 (baseline and endline years). The two series track each other closely, with no discernible differences at baseline, endline, or other times in those years. Because of randomization, it is not surprising that the series overlap each other up to and through our baseline study period. What is striking, however, is how closely they continue to track each other after Smartcards began to roll out in the summer of 2010, with no discernible gap emerging. This strongly suggests the existence of constraints that limited local officials' ability to increase the claims of work done.[26]

[26]Note that budgetary allocations are likely to be the binding constraint for NREGS volumes in AP because the state implemented NREGS well and prioritized using all federal fiscal allocations. In contrast, states like Bihar had large amounts of unspent NREGS funds, and ethnographic evidence suggests that the binding constraint in this setting was the lack of local project implementation capacity (Witsoe, 2014).

4.3 Effects on Program Access

Although Smartcards may have benefitted participants by reducing leakage, they could make it harder for others to participate in the first place. Access could fall for both mechanical and incentive reasons. Mechanically, beneficiaries might be unable to participate if they cannot obtain Smartcards or successfully authenticate. Further, by reducing leakage, Smartcards could reduce officials' primary motive for running programs in the first place. This is particular true for the NREGS which – despite providing a de jure entitlement to employment on demand – is de facto rationed (Dutta et al., 2012). Indeed, in our control group 20% (42%) of households reported that someone in their household was unable to obtain NREGS work in May (January) when private sector demand is slack (tight); and only 3.5% of households said that anyone in their village could get work on NREGS anytime (Table 5). Thus, the question of whether Smartcards hurt program access is a first order concern.

We find no evidence that this was the case. If anything, households with jobcards in treated mandals were 7.1 percentage points *more* likely to have done work on the NREGS during our study period, a 17% increase relative to control (Table 5, columns 1 and 2). Combined with the results in the previous section showing a significant reduction in the incidence of quasi-ghost NREGS workers, these results suggest that the optimal response of officials to their reduced ability to report work without providing any work or payments to the corresponding worker, was to provide more actual work (this section) and payments (previous section) to these workers. Beyond the increase in actual work during our survey period, columns 3 through 6 show that self-reported access to work also improved at other times of the year. The effects are insignificant in all but one case, but inconsistent with the view that officials "stop trying" once Smartcards are introduced. Bribes paid to access NREGS work were also (statistically insignificantly) lower (columns 7 and 8).

Given the theoretical concerns about potential negative effects of reducing leakage on program access, how should we interpret the lack of adverse effects in the data? One hypothesis is that officials simply had not had time to adapt their behavior (and reduce their effort on NREGS) by the time we conducted our endline surveys. However, the average converted GP in our data had been converted for 14.5 months at the time of our survey, implying that it had experienced two full peak seasons of NREGS under the new system. More generally, we find no evidence of treatment effects emerging over time in any of the official outcomes which we can observe weekly (e.g. Figure 1). On balance it thus appears more likely that we are observing a steady-state outcome.

A more plausible explanation for our results is that the main NREGS functionary (the Field Assistant) does not manage any other government program, which may limit the opportunities to divert rent-seeking effort. Further, despite the reduction in rent-seeking opportunities, implementing NREGS projects may have still been the most lucrative activity for the Field Assistant (note that we still estimate leakage rates of 20% in the treatment mandals). This may have mitigated potential negative extensive margin effects.[27]

We similarly find no evidence of reduced access to the SSP program. Since pensions are valuable and in fixed supply, the main concern here would be that reducing leakage in monthly payments simply displaces this corruption to the registration phase, increasing the likelihood that beneficiaries must pay bribes to begin receiving a pension in the first place. We find no evidence that reduced SSP leakage increased the incidence of bribes at the enrollment stage. Columns 9 and 10 of Table 5 show that the incidence of these bribes among SSP beneficiaries who enrolled after Smartcards implementation began is in fact 5.5 percentage points lower in treated mandals (73% of the control mean), although this result is not statistically significant.

4.4 Heterogeneity of Impacts

Even if Smartcards benefited the average program participant, it is possible that it harmed some. For instance, vulnerable households might have a harder time obtaining a Smartcard and end up worse off as a result. While individual-level treatment effects are by definition not identifiable, we can test the vulnerability hypothesis in two ways.

First, we examine quantile treatment effects for official payments, and survey outcomes that show a significant mean impact (time to collect payment, payment delays, and payments received). We find that the treatment distribution first-order stochastically dominates the control distribution for each of these outcomes (Figure 3). Thus, no treatment household is worse off relative to a control household at the same percentile in the outcome distribution.

Second, we examine whether treatment effects vary as a function of baseline characteristics at the village level. We begin with heterogeneity as a function of the baseline value of the outcome variable. The first row of Table F.1 suggests broad-based program impacts at all initial values of these outcomes. Overall, the data do not identify any particular group that appears to have suffered on these margins. We discuss the remainder of Table F.1 in Appendix F.

4.5 Beneficiary Perceptions of the Intervention

The estimated treatment effects thus far suggest that Smartcards unambiguously improved service delivery. It is possible, however, that our outcome measures miss impacts on some dimension of program performance that deteriorated. We therefore complement our impact

[27]The limited jurisdiction of the NREGS Field Assistant also suggests that there may have been limited opportunities for displacement of corruption to other programs (Yang (2008a)). While we cannot measure corruption in other sectors, we find no evidence of strategic displacement of NREGS corruption to non-treated mandals (see Appendix E.3).

estimates with beneficiaries' stated preferences regarding the Smartcard-based payment system as a whole. We asked recipients in converted GPs within treatment mandals who had been exposed to the Smartcard-based payment system to describe the pros and cons of the new process relative to the old one and state which they preferred.

Responses (Table 6) reflect many of our own ex ante concerns, but overall are overwhelmingly positive. Many recipients report concerns about losing their Smartcards (63% NREGS, 71% SSP) or having problems with the payment reader (60% NREGS, 67% SSP). Most beneficiaries do not yet trust the Smartcards system enough to deposit money in their accounts. Yet strong majorities (over 80% in both programs) also agree that Smartcards make payment collection easier, faster, and less manipulable. Overall, 90% of NREGS beneficiaries and 93% of SSP beneficiaries prefer Smartcards to the status quo, with only 3% in either program disagreeing, and the rest neutral.[28]

While stated preferences have well-known limitations, it is worth highlighting their value from a policy point of view. Senior officials in government were much more likely to hear field reports about problems with Smartcards than about positive results. This bias was so severe that GoAP nearly scrapped the entire Smartcards system in 2013, and their decision to not do so was partly in response to reviewing these stated preference data. The episode thus provides an excellent example of the political economy of concentrated costs (to low-level officials who lost rents due to Smartcards, and were vocal with negative feedback) versus diffuse benefits (to millions of beneficiaries, who were less likely to communicate positive feedback) (Olson, 1965).[29]

4.6 Mechanisms of Impact

As discussed earlier, the Smartcards intervention involved both technological changes (biometric authentication) and organizational changes (payments delivered locally by CSPs). The composite nature of the intervention does not allow us to decompose their relative contributions experimentally. We can, however, compare outcomes within the treatment group to get a sense of the relative importance of these two components of the Smartcards in-

[28]These questions were asked when beneficiaries had received a Smartcard and used it to pick up wages or had enrolled for, but not received, a physical Smartcard. We are thus missing data for those beneficiaries who received but did not use Smartcards (10.4% of NREGS beneficiaries and 3.4% of SSP beneficiaries who enrolled). Even if all of these beneficiaries for whom data is missing preferred the old system over Smartcards, approval ratings would be 80% for NREGS and 90% for SSP.

[29]Note also that vested interests trying to subvert the program would typically not do so by admitting that their rents were being threatened, but by making plausible arguments for why the new system would make poor beneficiaries worse off. Our data suggest that some of these concerns are very real (over 60% of beneficiaries report concerns about losing their Smartcards or encountering a non-functioning card reader), and highlight both the ease with which vested interests can hide behind plausibly genuine concerns, and the value of data from large, representative samples of beneficiaries.

tervention.[30] We have variation in our data both in whether CSPs were used for payment (because not all GPs converted) and in whether biometric IDs were used for authentication (because not all beneficiaries in converted GPs received or used biometric IDs).

Table 7 presents a non-experimental decomposition of the total treatment effects along these dimensions. For each of the main outcomes that are significant in the overall ITT estimates (payment process, leakage, and access), we find significant effects only in the carded GPs, suggesting that the new Smartcard-based payment system was indeed the mechanism for the ITT impacts we find.

In addition, we find that in converted GPs, uncarded beneficiaries benefit just as much as carded beneficiaries for payment process outcomes such as time to collect payments and reduction in payment lags (columns 1-4). These non-experimental decompositions provide suggestive evidence that converting a village to carded payments may have been the key mechanism by which there were improvements in the process of collecting payments, and also suggest that the implementation protocol followed by GoAP did not inconvenience uncarded beneficiaries in GPs that were converted to the new system. The lack of negative impacts for uncarded beneficiaries may be due to GoAP's decision to not insist on carded payments for all beneficiaries (due to the political cost of denying payments to genuine beneficiaries). While permitting uncarded payments may have allowed some amount of leakage to continue even under the new system, it was probably politically prudent to do so in the early stages of Smartcard implementation.

However, reductions in leakage appear to be concentrated among households with Smartcards, and we see no evidence of reduced leakage for uncarded beneficiaries (column 10), suggesting that biometric authentication was important for leakage reduction. Note that the lower official and survey payments to uncarded beneficiaries in converted GPs could simply reflect less active workers (who will be paid less) being less likely to have enrolled for the Smartcards, and so our main outcome of interest is leakage. The decomposition of program access is less informative for the same reason (since more active workers are more likely to have enrolled in the Smartcard), but we again see that all the increases in access are concentrated among households who had received a Smartcard. This is consistent with the pattern observed in Figure 3 suggesting that possession of a Smartcard made it more difficult for officials to report work on the corresponding jobcards without providing actual work and payments to households.

In short, the data suggest that the organizational shift to routing payments through banks

[30] While only suggestive, this is a policy-relevant question because these are aspects of the intervention that could in principle have been deployed individually. For instance, the government could have transitioned responsibility for payments delivery to banks and TSPs without requiring biometric authentication. Alternately, the government could have retained the status quo payment providers and required biometric authentication.

and ultimately through village-based CSPs is what drove improvements in the payments process, while the biometric authentication technology is what drove leakage reductions.

4.7 Robustness

In this section we address two main threats to the validity of the leakage results: differential mis-reporting on our survey and spillovers. Mis-reporting may be deliberate, because respondents collude with officials and report higher payments than they are entitled to, or inadvertent due to recall problems. If treatment affects collusion or recall, our results may be biased. We present several pieces of evidence that differential mis-reporting is not driving the results, and provide further details and additional checks in Appendix E.

First, note that Figure 3 shows a significant increase mainly in payments received by those who would have otherwise received no payments (relative to the control group). Since there is no reason to expect collusion only with this sub-group (if anything, it would arguably be easier for officials to collude with workers with whom they were already transacting), this pattern is difficult to reconcile with a collusion-based explanation. Since recalling whether one worked or not is easier than recalling the precise payment amount, this pattern also suggests our leakage results are not driven by differential recall.

Second, we conducted independent audits of NREGS worksites in treatment and control mandals during our endline surveys, and counted the number of workers who were present during unannounced visits to worksites. While imprecise, we find an insignificant 39.3% increase in the number of workers found on worksites in treatment areas during our audits (Table E.2), and cannot reject that this is equal to the 24% increase in survey payments reported in Table 3a. Thus, the audits find that the increase in survey payments reported are proportional to the measured increase in workers at worksites, suggesting that misreporting either because of collusion or recall bias is unlikely.

In addition, we directly test for differential rates of false survey responses by asking survey respondents to indicate whether they had ever been asked to lie about NREGS participation - using the "list method"[31] to elicit mean rates of being asked to lie without forcing any individual to reveal their answer - and find no significant difference between the treatment and control groups on this measure (Table E.3). Next, we saw that beneficiaries overwhelmingly prefer the new payment system to the old, which would be unlikely if officials were capturing most of the gains. We also find evidence that Smartcards increased wages in the *private* sector, consistent with the interpretation that it made NREGS employment a more remunerative alternative, and a more credible outside option for workers (see section 5).

[31] The list method is a standard device for eliciting sensitive information and allows the researcher to estimate population average incidence rates for the sensitive question, though the answers cannot be attributed at the respondent level (Raghavarao and Federer, 1979; Coffman et al., 2013).

Finally, we use the fact that our survey was spread over two months to check whether there was indeed differential recall. Holding constant the week in which work was actually done, survey lag does not affect the estimated treatment effect on leakage (Table E.4). While each of these pieces of evidence is only suggestive, taken together, they strongly suggest that our results do not reflect differential rates of collusion or recall bias in treatment mandals.

So far we have assumed that the Stable Unit Treatment Value Assumption (SUTVA) is satisfied; however, it is possible that one mandal's treatment status affects outcomes in other mandals. Such spillovers could occur if, for example, higher level officials reallocate funding to control mandals as it is easier to steal from them. We address this issue in two ways.

First, we note that there is no reallocation of funds to control mandals from treatment mandals; Figure 1 shows that average official spending is virtually identical in the two in both baseline and endline years. This is inconsistent with "strategic" spillover effects in which senior officials route funds to the places where they are easiest to steal. Second, we test for spatial spillovers by estimating the effect of a measure of exposure to treatment in the neighborhood of each GP (controlling for own treatment status). We find no evidence of spatial spillovers across any of our main outcomes (Table E.7).

Appendix E explores two additional robustness checks. Since we asked directly about when completed payments were made, we can check that our survey reports do not simply reflect the fact that treatment reduced payment delays so more respondents in treatment areas would have been paid by the time they were surveyed (Table E.2). Next, we designed our data collection activities to allow us to test whether the activities themselves affected measurement, and find no indication that they did (Hawthorne effects, Table E.9).

5 Cost-Effectiveness

We next estimate the cost-effectiveness of Smartcards as operating at the time of our endline survey. Some of the effects we measure are inherently redistributive, so that any valuation of them depends on the welfare weights we attach to various stakeholders. We therefore quantify costs and efficiency gains before discussing redistribution.

We assume that the cost of the Smartcard system was equal to the 2% commission that the government paid to banks on payments in converted GPs. This commission was calibrated to cover *all* implementation costs of banks and TSPs (including the one-time costs of enrollment and issuing of Smartcards), and is a conservative estimate of the incremental social cost of the Smartcard system because it does not consider the savings accruing to the government from decommissioning the status-quo payment system (e.g. the time of local officials who previously issued payments).[32] Using administrative data on all NREGS payments in 2012,

[32] Note that we do not include the time cost of senior officials in overseeing the Smartcard program because

and scaling down this figure by one-third (since costs were only paid in carded GPs, and only two-thirds of GPs were carded), we calculate the costs of the new payment system at \$4 million in our study districts. The corresponding figure for SSP is \$2.3 million.[33]

The efficiency gains we find include reductions in time taken to collect payment, and reductions in the variability of the lag between doing work and getting paid for it. We cannot easily price the latter, though we note that unpredictability is generally thought to be very costly for NREGS workers. To price the former, we estimate the value of time saved conservatively using reported agricultural wages during June, when they are relatively low. Using June wages of Rs. 130/day and assuming a 6.5 hour work-day (estimates of the length of the agricultural work day range from 5 to 8 hours/day), we estimate the value of time at Rs. 20/hour. We assume that recipients collect payments once per spell of work (as they do not keep balances on their Smartcards). Time to collect fell 22 minutes per payment (Table 2), so we estimate the value of time saved at Rs 7.3 per payment. While modest, this figure applies to a large number of transactions; scaling up by the size of the program in our study districts, we estimate a total saving of \$4.5 million for NREGS, suggesting that the value of time savings to beneficiaries *alone* may have exceed the government's implementation costs (for NREGS).

Redistributive effects include reduced payment lags (which transfer the value of interest "float" from banks to beneficiaries) and reduced leakage (which transfers funds from corrupt officials to beneficiaries). To quantify the former, we assume conservatively that the value of the float is 5% per year, the mean interest rate on savings accounts. Multiplied by our estimated 10-day reduction in payment lag and scaled up by the volume of NREGS payments in our study districts, this implies an annual transfer from banks to workers of \$0.4 million.[34] To quantify the latter, we multiply the estimated reduction in leakage of 12.7% by the annual NREGS wage outlay in our study districts and obtain an estimated annual reduction in leakage of \$38.5 million. Similarly, the estimated reduction in SSP leakage of 2.8% implies an annual savings of \$3.2 million.[35]

While valuing these redistributive effects requires subjective judgments about welfare weights, the fact that they both transferred income from the rich to the poor suggests

they would have had to exercise oversight of the older system as well.

[33]Note that our estimated impacts are ITT effects and are based on converting only two-thirds of GPs. An alternative approach would be to use the randomization as an instrument to generate IV estimates of the impact of being a carded GP. However, this will simply scale up both the benefit and cost estimates linearly by a factor of 1.5. We prefer the ITT approach because it does not require satisfying an additional exclusion restriction.

[34]Note that given the costs of credit-market intermediation, workers may value the use of capital well above the 5% deposit rate, as is suggested by the 26% benchmark interest rate for micro-loans, which are the most common form of credit in rural AP. In this case, the value of the reduced payment lag to beneficiaries may exceed the cost to the banks, implying an efficiency gain.

[35]Total NREGS wage outlays for the eight study districts in 2012 were \$303.5 million; SSP disbursements in these districts totalled \$112.7 million.

that they should contribute positively to a utilitarian social planner (assuming, for example, a symmetric utilitarian social welfare function with concave individual utility functions). Moreover, if taxpayers or the social planner place a low weight on losses to corrupt officials (as these are "illegitimate" earnings), then the welfare gains from reduced leakage are large.

The estimates above are based on measuring the direct impact of the Smartcards project on the main targeted outcomes of improving the payment process and reducing leakage. In preliminary work we have also found evidence that the intervention led to significant increases in rural private-sector wages, a general equilibrium effect which most likely represents the spillover effects to private labor markets of a better implemented NREGS (Imbert and Papp, 2015; Zimmermann, 2015). Since improving the outside options of rural workers in the lean season was a stated objective of the NREGS (Dreze, 2011), these results further suggest that Smartcards improved the capacity of the government to implement NREGS as intended.[36]

6 Conclusion

While a theoretical literature has emphasized the importance of investing in state capacity for economic development (Besley and Persson, 2009, 2010), the political viability of these investments depends on the magnitude and immediacy of their returns. Advocates argue that improved payments infrastructure may be a high-return investment in state capacity with the potential to significantly improve the implementation of public welfare programs in developing countries. The arguments are appealing, but there are many reasons to be skeptical. Implementations of new payments technology must overcome both logistical complexity and the resistance of vested interests. Those that do could potentially backfire by benefiting some while hurting the most vulnerable, or by eroding the incentives of bureaucrats to implement programs they previously viewed as sources of rents. Finally, technologies like biometric authentication could simply cost more than they are worth.

This paper has examined these issues empirically in the context of one of the largest randomized experiments yet conducted: an as-is evaluation of a new payment system built on biometric authentication and electronic benefit transfers introduced into two major social programs in the Indian state of Andhra Pradesh. We find that concerns about barriers to implementation are well-founded, as conversion was limited to 50% of transactions by the end of the study. Yet the poor gained significantly from the reform: beneficiaries received payments faster and more reliably, spent less time collecting payments, received a higher proportion of benefits, and paid less in bribes. These average gains did not come at the ex-

[36]Note that a better implemented NREGS could in principle also have efficiency costs, distorting the allocation of labor to the private sector. A full examination of such effects is beyond the scope of the current paper, which focuses on the impact of Smartcards on the quality of program implementation. We expect to study the GE effects of a better-implemented NREGS on rural labor markets in future work.

pense of vulnerable beneficiaries, as treatment distributions stochastically dominated those in control. Nor did they come at the expense of program access, which if anything appears to improved due to a reduction in over-reporting. Non-experimental decompositions suggest that organizational changes drove improvements in quality of service to beneficiaries, while biometric authentication drove reductions in fraud. Finally, beneficiaries themselves overwhelmingly reported preferring the new payment system to the old, and conservative cost-benefit calculations suggest that Smartcards more than justified their costs.

The fact that the theoretically-posited perverse side-effects did not materialize raises the question of what the Smartcards initiative did to minimize them. While we cannot provide definitive answers without further experimental variation, our extensive field experience evaluating the project leads us to conjecture that the government's decision to encourage but *not* mandate Smartcard-based payments may have played an important role. While this left open a major loophole for graft – likely explaining, for example, the lack of impact on ghost beneficiaries – it also ensured that beneficiaries could continue to access their NREGS and SSP benefits even if they were unable to obtain Smartcards or to authenticate. This tradeoff is particularly salient given the Indian Supreme Court's decision prohibiting the government from making possession of a UID mandatory for participation in federal welfare schemes. It also aptly illustrates the more general tradeoff between Type I and Type II errors in the administration of social programs, and suggests that it may be prudent to proceed with UID-linked benefit transfers by making it more attractive to beneficiaries, rather than making it mandatory.

A further conjecture supported by the AP Smartcard experience is that reducing leakage incrementally, as opposed to trying to eliminate it rapidly, may mitigate potential negative effects. For instance, the fact that NREGS Field Assistants still found it lucrative to implement projects (albeit with lower rents than before) may explain the lack of adverse effects on the extensive margin of program access. The gradual reduction of leakage may have also reduced the risk of political vested interests subverting the entire program.[37]

As usual, extrapolating this result to other settings requires care. While the overall level of development in AP almost precisely matches all-India averages, the state is generally perceived as well-administered, and devoted significant resources and senior management time to implementing the Smartcard program well. This raises the possibility that implementation would be more difficult in other settings. On the other hand, the problems that Smartcards were designed to address – slow, unpredictable, and leaky payments – are probably more

[37]The Government of India's pilot project (in 2013) on migrating in-kind subsidies for cooking gas to UID-linked cash transfers of the equivalent subsidy provides a cautionary tale. The pilot stopped benefits to those without UID-linked accounts, which sharply reduced official disbursements of subsidies since many beneficiaries were fake, but triggered strong political opposition following which it was shelved (see Barnwal (2014)).

severe elsewhere, implying greater potential upside. On net it is unclear whether the social returns would be higher or lower elsewhere. Similarly, forecasting the future evolution of the program requires care. Benefits could deteriorate if interest groups gradually find ways to subvert or capture the Smartcards infrastructure. On the other hand, benefits could increase if the government is able to increase coverage and plug remaining loopholes. Overall, our results are best interpreted as pointing to the potential for large returns in a relatively short time horizon should other governments choose to implement similar biometric payment systems to improve the delivery of public welfare programs.

More broadly, secure payments infrastructure may also facilitate future increases in the scale and scope of private economic transactions. In the absence of such infrastructure, payments often move through informal networks (Greif, 1993) or not at all. Thus, in addition to improving the delivery of public programs, investments in secure payments systems can be seen as building public infrastructure – akin to roads, railways, or the internet, which while initially set up by governments for their own use (e.g. moving soldiers to the border quickly or improving intra-government communication) eventually generated substantial benefits for the private sector as well. The gains reported in this paper do not reflect potential future benefits to other public programs or to private sector actors, and are thus likely to be a lower bound on the total long-term returns of investing in secure payments infrastructure.

References

Acemoglu, Daron, "Theory, General Equilibrium, and Political Economy in Development Economics," *Journal of Economic Perspectives*, 2010, *24* (3), 17–32.

Aker, Jenny, Rachid Boumnijel, Amanda McClelland, and Niall Tierney, "How do Electronic Transfers Compare? Evidence from a Mobile Money Cash Transfer Experiment in Niger," Technical Report, Tufts University 2013.

Anderson, Siwan, Patrick Francois, and Ashok Kotwal, "Clientilism in Indian Villages," *American Economic Review*, 2015, *105* (6), 1780–1816.

Banerjee, Abhijit, Rachel Glennerster, and Esther Duflo, "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System," *Journal of the European Economic Association*, 2008, *6* (2-3), 487–500.

Barnwal, Prabhat, "Curbing Leakage in Public Programs with Biometric Identification Systems: Evidence from India's Fuel Subsidies," Technical Report, Columbia University 2014.

Besley, Timothy and Torsten Persson, "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, September 2009, *99* (4), 1218–44.

_ **and** _ , "State Capacity, Conflict, and Development," *Econometrica*, 01 2010, *78* (1), 1–34.

Bohling, Kristy and Jamie Zimmerman, "Striving for E-payments at Scale: The Evolution of the Pantawid Pamilyang Pilipino Program in the Philippines," Technical Report, Consultative Group to Assist the Poor (CGAP), World Bank 2013.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur, "Interventions and Institutions: Experimental Evidence on Scaling up Education Reforms in Kenya," Technical Report, Center for Global Development 2013.

CGAP, "Electronic Payments with Limited Infrastructure: Uganda's Search for a Viable E-payments Solution for the Social Assistance Grants for Empowerment," Technical Report, World Bank 2013.

Coffman, Katherine, Lucas Coffman, and Keith Marzilli Ericson, "Privacy is Not Enough: The Size of the LGBT Population and the Magnitude of Anti-Gay Sentiment are Substantially Underestimated," Technical Report, Ohio State University 2013.

Dahl, Gordon B., Andreas Ravndal Kostøl, and Magne Mogstad, "Family Welfare Cultures," *The Quarterly Journal of Economics*, 2014, *129* (4), 1711–1752.

Dal Bó, Ernesto, Frederico Finan, and Martín Rossi, "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service," *The Quarterly Journal of Economics*, 2013, *128* (3), 1169–1218.

Dreze, Jean, "Employment Guarantee and the Right to Work," in Reetika Khera, ed., *The Battle for Employment Guarantee*, Oxford University Press, 2011.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan, "Incentives Work: Getting Teachers to Come to School," *American Economic Review*, 2012, *102* (4), 1241–78.

Dutta, P., S. Howes, and R. Murgai, "Small but effective: India's targeted unconditional cash transfers," *Economic and Political Weekly*, 2010, *45* (52), 63–70.

Dutta, Puja, Rinku Murgai, Martin Ravallion, and Dominique van de Walle, "Does India's Employment Guarantee Scheme Guarantee Employment?," Policy Research Working Paper Series 6003, World Bank 2012.

Fisman, Raymond and Shang-Jin Wei, "Tax Rates and Tax Evasion: Evidence from "Missing Imports" in China," *Journal of Political Economy*, April 2004, *112* (2), 471–500.

Fujiwara, Thomas, "Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil," *Econometrica*, 3 2015, *83* (2), 423–464.

Gelb, Alan and Julia Clark, "Identification for Development: The Biometrics Revolution," Working Paper 315, Center for Global Development 2013.

Gine, Xavier, Jessica Goldberg, and Dean Yang, "Credit Market Consequences of Improved Personal Identification: Field Experimental Evidence from Malawi," *American Economic Review*, October 2012, *102* (6), 2923–54.

Greif, Avner, "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition," *American Economic Review*, 1993, *83* (3), pp. 525–548.

Harris, Gardiner, "India Aims to Keep Money for Poor Out of Others' Pockets," *New York Times*, January 5 2013.

Imbert, Clement and John Papp, "Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee," *American Economic Journal: Applied Economics*, 2015, *7* (2), 233–263.

Jack, William and Tavneet Suri, "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution," *American Economic Review*, 2014, *1*, 183–223.

Jayachandran, Seema, "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries," *Journal of Political Economy*, 2006, *114* (3), pp. 538–575.

Khera, Reetika, "The UID Project and Welfare Schemes," *Economic and Political Weekly*, 2011, *46* (9).

Kremer, Michael, "The O-Ring Theory of Economic Development," *The Quarterly Journal of Economics*, 1993, *108* (3), 551–575.

Krusell, Per and Jose-Victor Rios-Rull, "Vested Interests in a Positive Theory of Stagnation and Growth," *The Review of Economic Studies*, 1996, *63* (2), 301–329.

Leff, Nathaniel, "Economic Development through Bureaucratic Corruption," *American Behavioural Scientist*, 1964, *8*, 8–14.

Lizzeri, Alessandro and Niccola Persico, "The Provision of Public Goods under Alternative Electoral Incentives," *American Economic Review*, 2001, *91* (1), pp. 223–239.

Mathew, Santhosh and Mick Moore, "State incapacity by design: Understanding the Bihar story," *IDS Working Papers*, 2011, *2011* (366), 1–31.

Mishra, Prachi, Arvind Subramanian, and Petia Topalova, "Tariffs, enforcement, and customs evasion: Evidence from India," *Journal of Public Economics*, October 2008, *92* (10-11), 1907–1925.

Mukhopadhyay, Piali, Karthik Muralidharan, Paul Niehaus, and Sandip Sukhtankar, "Implementing a Biometric Payment System: The Andhra Pradesh Experience," Technical Report, University of California, San Diego 2013.

Muralidharan, Karthik, Jishnu Das, Alaka Holla, and Aakash Mohpal, "The Fiscal Cost of Weak Governance: Evidence from Teacher Absence in India," Working Paper 20299, National Bureau of Economic Research 2014.

Niehaus, Paul and Sandip Sukhtankar, "Corruption Dynamics: The Golden Goose Effect," *American Economic Journal: Economic Policy*, 2013, *5.*

_ and _ , "The Marginal Rate of Corruption in Public Programs: Evidence from India," *Journal of Public Economics*, 2013, *104*, 52 – 64.

NIPFP, "A Cost-Benefit Analysis of Aadhaar," Technical Report, National Institute for Public Finance and Policy 2012.

Olken, Benjamin A., "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy*, April 2007, *115* (2), 200–249.

Olson, Mancur, *The Logic of Collective Action: Public Goods and the Theory of Groups*, Harvard University Press, 1965.

Pai, Sandeep, "Delayed NREGA payments drive workers to suicide," *Hindustan Times*, December 29 2013.

PEO, "Performance Evaluation of Targeted Public Distribution System," Technical Report, Planning Commission, Government of India March 2005.

Prescott, Edward and Stephen Parente, *Barriers to Riches*, Cambridge: MIT Press, 2000.

Pritchett, Lant, "Is India a Flailing State? Detours on the Four Lane Highway to Modernization," Working Paper RWP09-013, Harvard Kennedy School 2010.

Raghavarao, Damaraju and Walter T. Federer, "Block total response as an alternative to the randomized response method in surveys," *Journal of the Royal Statistical Society*, 1979, *41* (1), 40–45.

Ravi, Shamika and Monika Engler, "Workfare as an Effective Way to Fight Poverty: The Case of India's NREGS," Technical Report, Indian School of Business 2013.

Reinikka, Ritva and Jakob Svensson, "Local Capture: Evidence From a Central Government Transfer Program in Uganda," *The Quarterly Journal of Economics*, May 2004, *119* (2), 678–704.

_ **and** _ , "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda," *Journal of the European Economic Association*, 04/05 2005, *3* (2-3), 259–267.

Witsoe, Jeffrey, "The Practice of Development: An Ethnographic Examination of the National Rural Employment Guarantee Act in Bihar," Mimeo, Union College 2014.

World Bank, "World Development Report 2004: Making Services Work for Poor People," Technical Report, World Bank 2003.

Yang, Dean, "Can Enforcement Backfire? Crime Displacement in the Context of Customs Reform in the Philippines," *The Review of Economics and Statistics*, November 2008, *90* (1), 1–14.

_ , "Integrity for Hire: An Analysis of a Widespread Customs Reform," *Journal of Law and Economics*, 02 2008, *51* (1), 25–57.

Zimmermann, Laura, "Why Guarantee Employment? Evidence from a Large Indian Public-Works Program," Working Paper, University of Georgia April 2015.

Table 1: Official and self-reported use of Smartcards

(a) NREGS

	Official data		Survey data	
	(1)	(2)	(3)	(4)
	Carded GP	Mean fraction carded payments	Payments generally carded (village mean)	Most recent payment carded (village mean)
Treatment	.67***	.45***	.38***	.38***
	(.045)	(.041)	(.043)	(.042)
District FE	Yes	Yes	Yes	Yes
Adj R-squared	.45	.48	.36	.36
Control Mean	.0046	.0017	.039	.013
N. of cases	880	880	818	818
Level	GP	GP	GP	GP

(b) SSP

	Official data		Survey data	
	(1)	(2)	(3)	(4)
	Carded GP	Mean fraction carded payments	Payments generally carded (village mean)	Most recent payment carded (village mean)
Treatment	.79***	.59***	.45***	.45***
	(.042)	(.038)	(.052)	(.049)
District FE	Yes	Yes	Yes	Yes
Adj R-squared	.57	.57	.38	.38
Control Mean	0	0	.069	.044
N. of cases	880	880	878	878
Level	GP	GP	GP	GP

This table analyzes usage of Smartcards for NREGS and SSP payments as of July 2012. Each observation is a gram panchayat ("GP": administrative village). "Carded GP" is a gram panchayat that has moved to Smartcard-based payment, which usually happens once 40% of beneficiaries have been issued a card. "Mean fraction carded payments" is the proportion of transactions done with carded beneficiaries in treatment mandals. Both these outcomes are from official data. Columns 3 and 4 report survey-based measures of average beneficiary use of Smartcards or a biometric-based payment system in the GP. The difference in number of observations between official and survey measures for NREGS is due to missing data for (mainly control) GPs where enrollment had not even started; assuming that there were no carded payments in these GPs increases the magnitude of the treatment effect on implementation. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 2: Access to payments

	Time to Collect (Min)				Payment Lag (Days)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					Ave Payment Delay		Deviation	
Treatment	-22**	-22**	-6.1	-3.5	-5.8*	-10***	-2.5**	-4.7***
	(9.2)	(8.7)	(5.2)	(5.4)	(3.5)	(3.5)	(.99)	(1.6)
BL GP Mean		.079*		.23***		.013		.042
		(.041)		(.07)		(.08)		(.053)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week Fe	No	No	No	No	Yes	Yes	Yes	Yes
Adj R-squared	.06	.08	.07	.11	.17	.33	.08	.17
Control Mean	112	112	77	77	34	34	12	12
N. of cases	10191	10120	3789	3574	14213	7201	14213	7201
Level	Indiv.	Indiv.	Indiv.	Indiv.	Indiv-Week	Indiv-Week	Indiv-Week	Indiv-Week
Survey	NREGS	NREGS	SSP	SSP	NREGS	NREGS	NREGS	NREGS

The dependent variable in columns 1-4 is the average time taken to collect a payment (in minutes), including the time spent on unsuccessful trips to payment sites, with observations at the beneficiary level. The dependent variable in columns 5-6 is the average lag (in days) between work done and payment received on NREGS. The outcome in columns 7-8 is the absolute deviation from the week-specific median mandal-level lag. Since the data for columns 5-8 are at the individual-week level, we include week fixed effects to absorb variation over the study period. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 3: Official and survey reports of program benefits

(a) NREGS

	Official		Survey		Leakage	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	11	9.6	35**	35**	-24*	-25*
	(12)	(12)	(16)	(16)	(13)	(13)
BL GP Mean		.13***		.11***		.096**
		(.027)		(.037)		(.038)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.03	.05	.05	.06	.04	.04
Control Mean	127	127	146	146	-20	-20
N. of cases	5143	5107	5143	5107	5143	5107

(b) SSP

	Official		Survey		Leakage	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	4.3	5.1	12**	12*	-7.5*	-7*
	(5.3)	(5.4)	(5.9)	(6.1)	(3.9)	(3.9)
BL GP Mean		.16*		.0074		-.022
		(.092)		(.022)		(.026)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.00	.01	.01	.01	.01	.01
Control Mean	251	251	236	236	15	15
N. of cases	3330	3135	3330	3135	3330	3135

This table reports regressions of program benefits (in Rupees) as reported in official or survey records. Both panels include all sampled households (NREGS)/beneficiaries (SSP) who were a) found by survey team to match official records or b) listed in official records but confirmed as "ghosts". "Ghosts" refer to households or beneficiaries within households that were confirmed not to exist, or who had permanently migrated before the study period started on May 28, 2012. In panel (a), each outcome observation refers to household-level average weekly amounts for NREGS work done during the study period (May 28 to July 15 2012). "Official" refers to amounts paid as listed in official muster records. "Survey" refers to payments received as reported by beneficiaries; we only include beneficiaries listed on the officially sampled jobcard. "Leakage" is the difference between these two amounts. "BL GP Mean" is the GP average of household-level weekly amounts for NREGS work done during the baseline study period (May 31 to July 4 2010). In panel (b), each outcome observation refers to the average SSP monthly amount for the period May, June, and July 2012. "Official" refers to amounts paid as listed in official disbursement records. "Survey" refers to payments received as reported by beneficiaries. "Leakage" is the difference between these two amounts. "BL GP Mean" is the GP average SSP monthly amounts for the baseline period of May, June, and July 2010. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 4: Illustrating channels of leakage reduction

(a) NREGS

	Ghost households (%)		Other overreporting (%)		Bribe to collect (%)	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-.012	-.011	-.082**	-.084**	-.0083	-.0087
	(.021)	(.022)	(.033)	(.036)	(.013)	(.013)
BL GP Mean		-.013		.016		.0087
		(.069)		(.044)		(.023)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.02	.02	.05	.04	.02	.02
Control Mean	.11	.11	.26	.26	.026	.026
N. of cases	5143	5107	3953	3672	7119	7071
Level	Hhd	Hhd	Hhd	Hhd	Indiv.	Indiv.

(b) SSP

	Ghost payments (Rs)		Other overreporting (Rs)		Underpayment (Rs)	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-2.9	-2.4	-2.7	-3.1	-2.3	-2.4
	(2.7)	(2.7)	(2.9)	(3)	(1.9)	(2)
BL GP Mean		.19		.024**		-.02
		(.16)		(.01)		(.045)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.01	.01	.01	.01	.01	.01
Control Mean	11	11	1.7	1.7	2.5	2.5
N. of cases	3330	3135	3165	2986	3165	2986
Level	Indiv.	Indiv.	Indiv.	Indiv.	Indiv.	Indiv.

This table analyzes channels of reduction in leakage. Panel (a) reports the incidence of the three channels - ghosts, overreporting, and underpayment - for NREGS, while panel (b) decomposes actual amounts (in Rupees) into these channels in the case of SSP. In both tables, "Ghost households" refer to households (or all beneficiaries within households) that were confirmed not to exist, or who had permanently migrated before the study period started on May 28, 2012 (May 31, 2010 for baseline). "Other overreporting" for NREGS is the incidence of jobcards that had positive official payments reported but zero survey amounts (not including ghosts); note that the drop in observations as compared to Table 3a is because here we drop jobcards with 0 official payments. "Bribe to collect" refers to bribes paid in order to receive payments on NREGS. "Other overreporting" for SSP is the difference between what officials report beneficiaries as receiving and what beneficiaries believe they are entitled to (not including ghosts). "Underpayment" for SSP is the monthly amount paid in order to receive their pensions in May-July 2012. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 5: Access to programs

	Proportion of Hhds doing NREGS work		Was any Hhd member unable to get NREGS work in...		Is NREGS work available when anyone wants it		Did you have to pay anything to get this NREGS work?		Did you have to pay anything to start receiving this pension?	
	(1) Study Period	(2) Study Period	(3) May	(4) January	(5) All Months	(6) All Months	(7) NREGS	(8) NREGS	(9) SSP	(10) SSP
Treatment	.072**	.071**	-.023	-.027	.027*	.024	-.0003	-.00054	-.046	-.055
	(.033)	(.033)	(.027)	(.033)	(.015)	(.015)	(.0015)	(.0015)	(.031)	(.039)
BL GP Mean		.14***				-.023		-.0064**		.025
		(.038)				(.027)		(.0031)		(.046)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.05	.06	.10	.11	.02	.02	.00	.00	.05	.05
Control Mean	.42	.42	.2	.42	.035	.035	.0022	.0022	.075	.075
N. of cases	4943	4909	4748	4496	4755	4715	7185	6861	581	352

This table analyzes household level access to NREGS and SSP. Columns 1-2 report the proportion of households doing work in the 2012 endline study period (May 28-July 15). If any member of the household did work on NREGS during that period, the household is considered "working". In columns 3-4, the outcome is an indicator for whether any member of household was unable to obtain work despite wanting to work during May (slack labor demand) or January (peak labor demand). In columns 5-6, the outcome is an indicator for whether the respondent believes anyone in the village who wants NREGS work can get it at any time. In columns 7-8, the outcome is an indicator for whether the respondent had to pay a bribe in order to obtain NREGS work during the endline study period. Note that only NREGS beneficiaries who worked during the endline study period are considered in columns 7-8. In columns 9-10, the outcome is an indicator for whether the respondent had to pay a bribe to get on the SSP beneficiary list in the years 2011 and 2012 (hence only new enrollees are included). All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 6: Beneficiary opinions of Smartcards

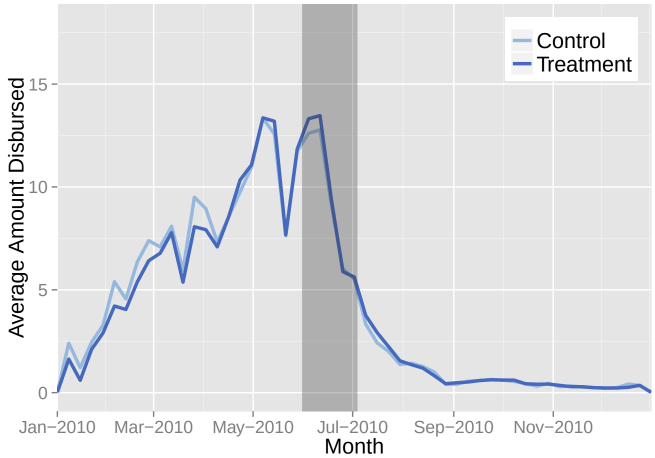
	NREGS				SSP			
	Agree	Disagree	Neutral/ Don't know	N	Agree	Disagree	Neutral/ Don't know	N
Positives:								
Smartcards increase speed of payments (less wait times)	.83	.04	.13	3336	.87	.07	.06	1451
With a Smartcard, I make fewer trips to receive my payments	.78	.04	.18	3334	.83	.04	.12	1450
I have a better chance of getting the money I am owed by using a Smartcard	.83	.01	.16	3333	.86	.03	.11	1450
Because I use a Smartcard, no one can collect a payment on my behalf	.82	.02	.16	3331	.86	.03	.11	1446
Negatives:								
It was difficult to enroll to obtain a Smartcard	.19	.66	.15	3338	.29	.60	.11	1451
I'm afraid of losing my Smartcard and being denied payment	.63	.15	.21	3235	.71	.15	.14	1403
When I go to collect a payment, I am afraid that the payment reader will not work	.60	.18	.22	3237	.67	.18	.14	1403
I would trust the Smartcard system enough to deposit money in my Smartcard account	.29	.41	.30	3334	.31	.46	.24	1448
Overall:								
Do you prefer the smartcards over the old system of payments?	.90	.03	.07	3346	.93	.03	.04	1454

This table analyzes beneficiaries' perceptions of the Smartcard program in GPs that had switched over to the new payment system (carded GPs). These questions were asked when NREGS and SSP beneficiaries had received a Smartcard and used it to pick up wages; and also if they had enrolled for, but not received, a physical Smartcard. We are thus missing data for those beneficiaries who received but did not use Smartcards (10.5% of NREGS beneficiaries and 3.5% of SSP beneficiaries who enrolled).

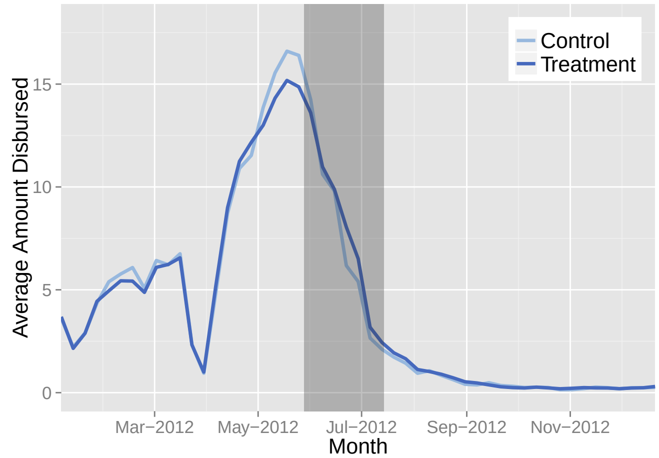
Table 7: Non-experimental decomposition of treatment effects by carded status

	Time to collect		Payment lag		Official		Survey		Leakage		Proportion of Hhds doing NREGS work	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Carded GP	-33***		-4.9*		10		40**		-29**		.078**	
	(8.1)		(2.8)		(13)		(15)		(13)		(.035)	
Have SCard, Carded GP		-33***		-4.7		93***		169***		-75***		.31***
		(8.4)		(2.9)		(17)		(23)		(22)		(.041)
No SCard, Carded GP		-33***		-5.4*		-14		-11		-4.6		-.042
		(8.5)		(2.9)		(14)		(17)		(13)		(.043)
No Info SCard, Carded GP		.33		-5.9		-109***		-128***		18		-.38***
		(20)		(3.7)		(13)		(15)		(13)		(.037)
Not Carded GP	4.9	5	-7.4	-7.4	8.4	6.6	23	20	-14	-13	.056	.046
	(13)	(13)	(5)	(5)	(16)	(16)	(22)	(22)	(19)	(19)	(.04)	(.042)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	No	No	Yes	Yes	No	No	No	No	No	No	No	No
BL GP Mean	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
p-value: carded GP = not carded GP	<.001***		.46		.9		.35		.39		.5	
p-value: Have SC = No SC		.88		.67		<.001***		<.001***		.0019***		<.001***
Adj R-squared	.1	.1	.17	.17	.046	.095	.057	.13	.04	.051	.058	.16
Control Mean	112	112	34	34	127	127	146	146	-20	-20	.42	.42
N. of cases	10120	10120	14213	14213	5107	5107	5107	5107	5107	5107	4909	4909
Level	Indiv.	Indiv.	Indiv-Week	Indiv-Week	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd

This table shows the main ITT effects decomposed by levels of program implementation. “Carded GP” is a gram panchayat that has moved to Smartcard based payments (5056 individuals, 2541 households). “Have SCard, Carded GP” (2619 individuals, 1398 households) and “No SCard, Carded GP” (2419 individuals, 957 households) are based on whether the beneficiary or household lives in a carded GP and self-reported receiving a Smartcard (at least one Smartcard in the household for household-level variables). “No Info SCard, Carded GP” (18 individuals, 186 households) are those whose Smartcard ownership status is unknown, either because they did not participate in the program and hence were not asked questions about Smartcards, or because they are ghost households (since we cannot verify whether those households did have a fake Smartcard in their name). “Not Carded GP” is a gram panchayat in a treatment mandal that has not yet moved to Smartcard-based payments (2261 individuals, 1131 households). For each outcome, we report the p-values from a test of equality of the coefficients on “Carded GP” and “Not Carded GP” (odd columns), and “Have SCard” and “No Scard” (even columns). A specification with the baseline mean is not reported for the payment lag outcome due to a large number of missing baseline observations, which makes decomposition difficult. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$



(a) 2010



(b) 2012

Figure 1: Official disbursement trends in NREGS

This figure shows official NREGS payments for all workers averaged at the GP-week level for treatment and control areas. The grey shaded bands denote the study periods on which our survey questions focus (baseline in 2010 - May 31 to July 4; endline in 2012 - May 28 to July 15).

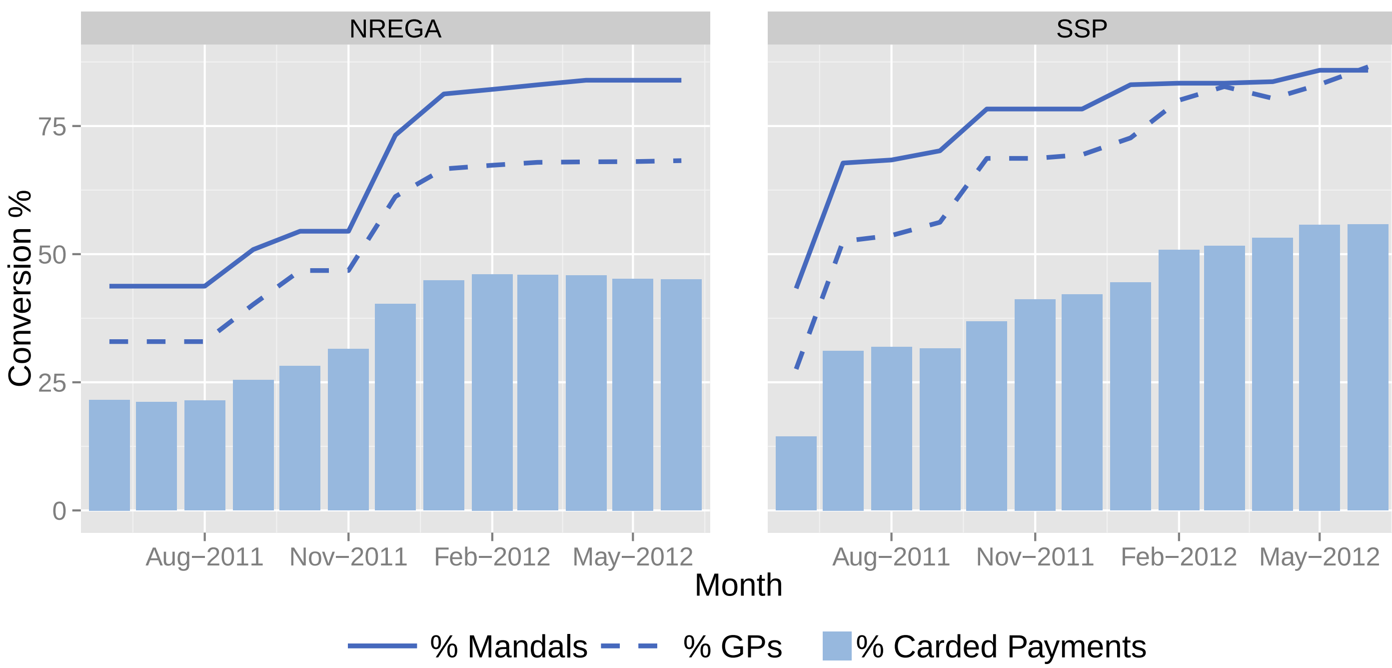
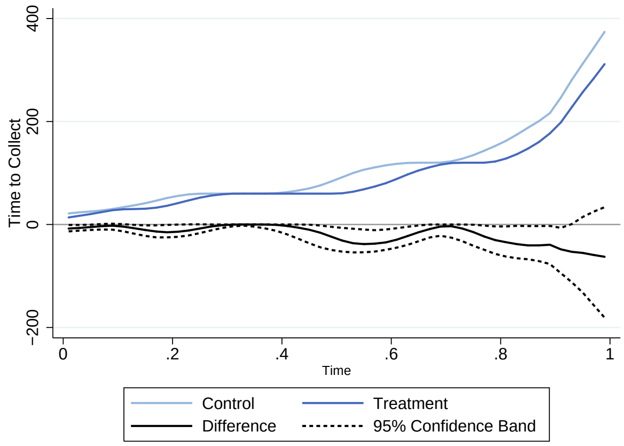
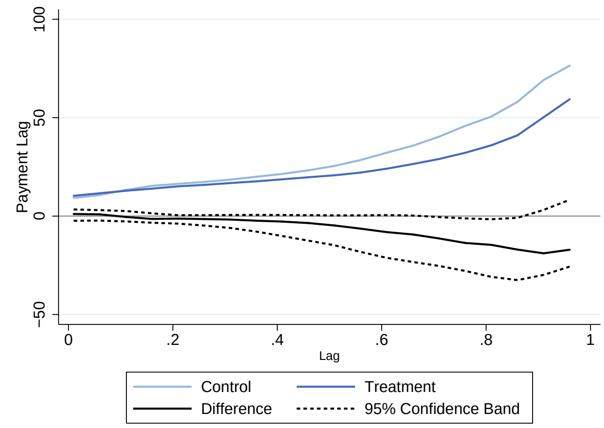


Figure 2: Rollout of Smartcard integration with welfare programs

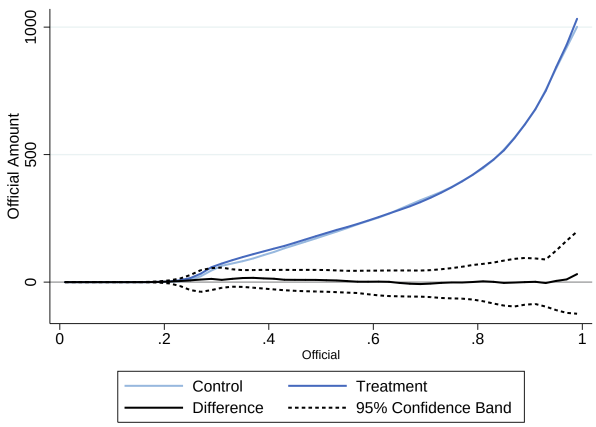
This figure shows program rollout in aggregate and at different conversion levels. Each unit converts to the Smartcard-enabled system based on beneficiary enrollment in the program. "% Mandals" is the percentage of mandals converted in a district. A mandal converts when at least one GP in the mandal converts. "% GPs" is the percentage of converted GPs across all districts. "% Carded Payments" is obtained by multiplying % Mandals by % converted GPs in converted mandals and % payments to carded beneficiaries in converted GPs.



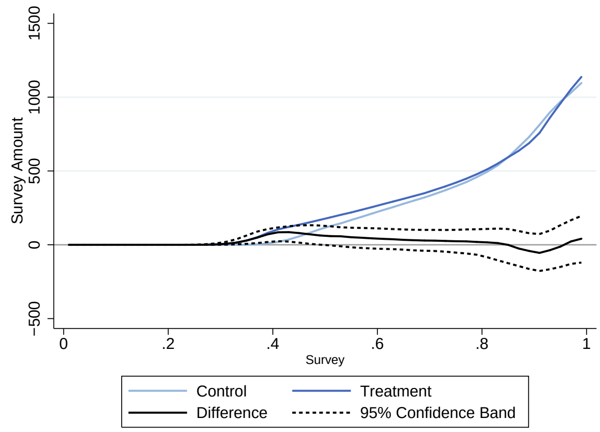
(a) Time to collect: NREGS



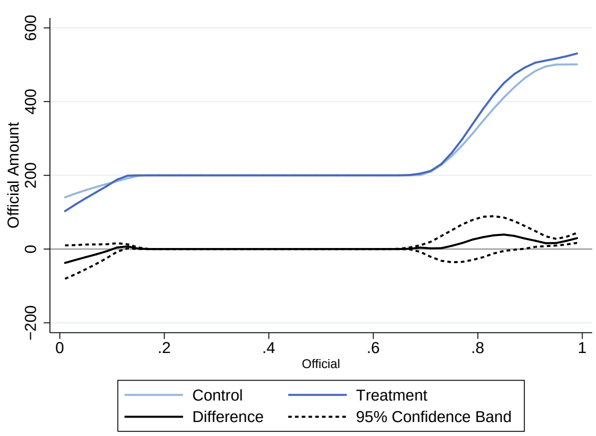
(b) Payment Lag: NREGS



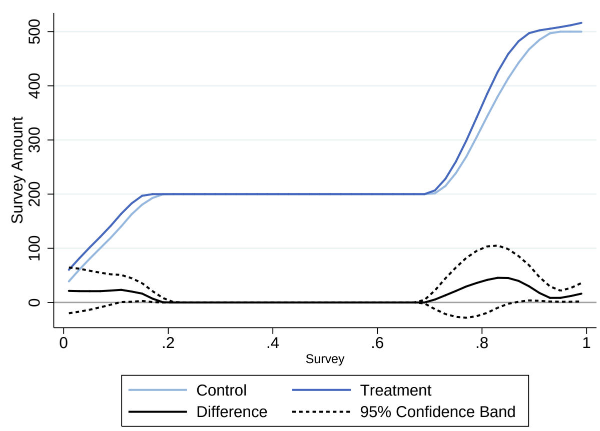
(c) Official: NREGS



(d) Survey: NREGS



(e) Official: SSP



(f) Survey: SSP

Figure 3: Quantile Treatment Effects on Key Outcomes

Panels (a)-(f) show nonparametric treatment effects. "Time to collect: NREGS" is the average time taken to collect a payment, including the time spent on unsuccessful trips to payment sites. "Payment Lag: NREGS" is the average lag (in days) between work done and payment received under NREGS. The official payment amounts, "Official: NREGS" and "Official: SSP", refer to payment amounts paid as listed in official muster/disbursement records. The survey payment amounts, "Survey: NREGS" and "Survey: SSP" refer to payments received as reported by beneficiaries. The NREGS data is taken from the study period (endline was 2012 - May 28 to July 15), while SSP official data is an average of June, July and August disbursements. All lines are fit by a kernel-weighted local polynomial smoothing function with Epanechnikov kernel and probability weights, with bootstrapped standard errors. The dependent variable is the vector of residuals from a linear regression of the respective outcome with the first principal component of a vector of mandal characteristics used to stratify randomization and district fixed effects as regressors.

FOR ONLINE PUBLICATION ONLY

A Further Background on Programs and Smartcard Intervention

This Section provides further information on the two welfare programs - NREGS and SSP - as well as the Smartcards intervention that changed the payment system for the two programs, focusing on supplemental information that was not provided in the main text in order to conserve space.

A.1 NREGS

The National Rural Employment Guarantee Act (NREGA) of 2005 - ex-post renamed the Mahatma Gandhi National Rural Employment Guarantee Act (MNREGA) - mandated federal and state governments to set up employment programs which would guarantee one hundred days of paid employment to any rural household in India. The employment programs, or "schemes", which are collectively referred to as NREGS, are meant to be a self-targeting safety net, with those in need of wage labor accessing work during slack labor seasons. There is no eligibility requirement in order to get work through the program.

The first step in obtaining NREGS employment is to obtain a jobcard. This is a household level document that lists all adult members of the household, and also has assigned pages for recording details of work done and payment owed, including dates of employment and payment. Obtaining a jobcard is generally a simple process, and 65.7% of rural households in Andhra Pradesh have jobcards according to National Sample Survey data; this likely comprises the universe of households who might consider working on NREGS.

Program beneficiaries do (mainly) physical labor at minimum wages. These wages are set at the state level, and can be daily wages or piece rates. Most of the work done in Andhra Pradesh is paid on the basis of piece rates. These rates vary by difficulty of task, and are supposed to enable workers to attain the daily minimum wage with roughly a day's worth of effort. Available tasks depend on the project undertaken, which generally include road construction, field clearing, and irrigation earthworks.

Local village officials are responsible for the implementation of NREGS projects, which are meant to be chosen in advance at a village-wide meeting (the "Gram Sabha"). Project worksites are managed by officials called Field Assistants, who record attendance and output on "muster rolls" and send these to the sub-district for digitization, from where the work records are sent up to the state level, which triggers the release of funds to pay workers. In the status quo, payment was made often by the same Field Assistants in workers' villages, or through the local post office, with no formal authentication procedure required.

Although the program is meant to be demand driven, in practice work is available when there is a project active in the village, and not otherwise. As Figure 1 suggests, there is very high seasonality in when the program is active, with the main periods of activity being the dry season months of April, May and June. Thus the 100 day limit rarely binds per se for particular households, particularly since it may be possible to get around the limit by creating multiple jobcards per household. For example, Imbert and Papp (2015) note that in 2009-10 the median household worked for only 30 days out of the year (mean was 38 days). Moreover, participation varies at high frequency as participants move in and out of the program; Ravi and Engler (2013) find that only about 30% of households in a panel survey of ultra poor households (very likely NREGS participants) in Andhra Pradesh worked in both 2007 and 2009 even though the survey was conducted at the same time of year.

In addition to rationing, other implementation issues are also rife. NREGS workers have to wait over a month to receive payments after working, spend about 2 hours per payment to collect payments, and face much uncertainty over when exactly they will be paid. Of these issues, the long wait to be paid has created some outcry in the media, who have reported on beneficiaries committing suicide because of the inordinate delay (Pai, 2013).

Workers must also worry about whether they will receive the full payment due to them, as corrupt officials may pocket earnings along the way (Niehaus and Sukhtankar, 2013a,b). Leakage from the labor budget may take two forms: underpayment, in which an official simply pays the worker less than she is owed, and over-reporting, in which the official invoices the government for more than what the worker is owed, and pockets the difference. Over-reporting includes invoicing for "ghost" workers, i.e. workers who do not exist, or "quasi-ghost" workers, who exist in the database but have actually not participated on the program at all. Leakage from other parts of the budget is also possible, for example by overinvoicing for materials, but as can be seen in Table E.8, spending on wages is over 91% of the overall budget.

A.2 SSP

The Social Security Pension (SSP) program is a welfare scheme that contrasts with the NREGS on multiple dimensions. First, there are clear eligibility criteria, with pensions restricted to those who are below the poverty line and have restricted earnings ability in some form, due to old age, disability, or being member of a traditional and now outdated profession. Second, if the eligibility criteria are satisfied, the program provides an unconditional cash transfer: these are no work or other requirements. Finally, in contrast to the NREGS in which participation varies at high frequency, SSP beneficiaries are more or less permanent participants after enrollment. The only churn is as a result of death or migration, although these rates are higher than those of the general population given that SSP beneficiaries are targeted for being elderly and disabled.

While there is far less academic research on this program as compared to the NREGS, the little that is available suggests that the program is far better implemented. Dutta et al. (2010) examine the program functioning in Karnataka and Rajasthan, and find that it is well targeted, with poorer households far more likely to obtain benefits than richer households. Moreover, levels of leakage are low: about 17% in Karnataka, less than half comparable rates on an in-kind transfer program (the Public Distribution System) in the same sample.

We did not find any documented evidence on the functioning of the actual payment process for SSP, likely because it is a straightforward process and does not suffer from the types of problems observed in the NREGS programs. The SSP program has a more or less fixed list of beneficiaries, who receive a fixed amount of payment at a fixed time every month (usually in the first week of the month). Our pilots on this issue corroborated this view of the payments process on SSP, and we therefore did not collect data on this aspect of the program.

Overall, we can think of SSP beneficiaries as salaried permanent employees, and NREGS beneficiaries as spot workers on the casual labor market who may or may not show up to obtain work on a given day. The pensioners are paid a fixed wage (entitlement) each month of the year at a specific time of the month (like receiving a monthly paycheck or direct deposit at the end of the month). Meanwhile, NREGS workers are paid based on how much work they did, and this participation varies at high frequency.

A.3 Smartcards intervention

The Smartcards project began in Andhra Pradesh in 2006 in order to improve the payments system for two main welfare schemes in the state. By 2010, Smartcards had been rolled out in 13 out of 21 non-urban districts in the state. The Smartcards system was implemented by private and public sector banks who worked with Technology Service Providers (TSPs) to manage the technological details last-mile delivery and authentication. Each district was assigned to a single bank via a system of competitive bidding. In Nalgonda district, the winning entity was actually the post office. Banks were paid 2% of every transaction in villages in which they handled the payment system. The bank was responsible for sharing this commission with the TSP as per their contract.

In some cases TSPs subcontracted the actual last-mile delivery to another entity, called a "banking correspondent," (BC) who handled the village level Customer Service Providers (CSPs) who actually made the payments. The TSP or BC was responsible for hiring CSPs as per the criteria laid down by the government, and making sure actual cash was delivered to these local agents. Typically a mandal-level coordinator handled the delivery of cash to CSPs, and assisted in training and providing other support to the CSPs.

Figure A.1 demonstrates the payment flow system. Banks opened "no-frills" accounts for NREGS and SSP beneficiaries, and payments were deposited into these accounts. Individual beneficiaries could not however access these payments directly; they could only be paid through CSPs, who were supposed to verify beneficiary identity via fingerprint authentication.

Authentication was performed via the use of small Point-of-Service (PoS) devices, pictured in Figure A.2. The devices did not require internet connectivity in order to authenticate, as they simply matched the fingerprint placed on the device with the biometric information stored on the Smartcard that was inserted into the device at the same time. One TSP did not store biometric information on the Smartcard, but rather on a data storage device in the PoS device itself, which restricted the use of the machine to only those beneficiaries whose biometrics were stored on the device (typically at the GP level). All machines were battery powered, so did not need to be plugged in to an external source of electricity. At the end of the day, after cash was dispensed, the machines could be charged back up and connected via GPRS to the banks' network for reconciliation of accounts.

The Smartcards system was a precursor to the nationwide Aadhaar/ biometric Unique ID system. While functionally equivalent for making NREGS and SSP payments, there are some differences between Aadhaar and Smartcards. Most importantly, Aadhaar requires connectivity to a central server for authentication, while Smartcards authentication is offline. Aadhaar can thus be used across various platforms across states, while the use of Smartcards was restricted to making payments for NREGS and SSP beneficiaries within Andhra Pradesh.

B Data

This section describes various data we use in the paper, as well as the collection process involved in obtaining the data.

B.1 Official data

B.1.1 NREGS

We received two types of data from Tata Consultancy Services, which manages the Monitoring and Information System for the Department of Rural Development of the Government of Andhra Pradesh. The first dataset is the full jobcard database, i.e. every single jobcard in the system at the moment of data transfer in each of our study districts. Each jobcard entry in this database contains a listing of family members, including name, sex, age, as well as caste status of the household and address details. The second dataset is the muster roll or disbursement data, which contains details of participation on NREGS for the study period. These details include the jobcard number, dates worked, project worked on, and amount disbursed by the government.

We received both sets of data at two separate points in time: in mid-July 2010 prior to the baseline survey, and mid-July 2012 prior to the endline survey. Note that treatment did not affect the collection or reporting of data in any way, which was managed by the same officials at the village level and the same agency at the state level in all areas at all times over the course of this study. We explain the sampling procedure, which uses both these sets of data, in section C.2 below.

B.1.2 SSP

The official SSP data mirrored those from the NREGS, with one dataset corresponding to the full list of SSP beneficiaries and the second dataset pertaining to recent disbursements. The Department of Rural Development of the Government of Andhra Pradesh directly gave us both datasets in mid-July 2010 and 2012. The SSP beneficiary list contains data on the individual beneficiary, including name, sex, age, caste states, address, and type of pension. The disbursement list contains beneficiary names and disbursement amounts for May, June, and July. Since benefit amounts do not change over the course of our study and we already have the list of beneficiaries, the only advantage of the disbursement data is that it may reflect slightly more current information on payments, and basically serve as confirmation that money was indeed disbursed by the government. Like the NREGS program, the Smartcards intervention did not affect collection or reporting of official data.

B.2 Survey data

We conducted two rounds of household surveys, a baseline survey in August-September 2010 and an endline survey in August-September 2012. We also conducted a midline survey in

September 2011, but that survey collected process data on the progress of the Smartcards intervention rather than data on outcomes. Accordingly, there were only 996 households surveyed in that round as compared to the 7425 at baseline and 8114 at endline. In addition to the household survey, we also had a village-level survey answered by a village elder, schoolteacher, or local official; we do not use these data in this paper. Finally, we also attempted to survey the mandal coordinators and CSPs, but had limited success in reaching them in the time frame that the survey team was in the area, with less than a 50% response rate for these surveys.

The household survey was comprised of seven modules. Module A was the household roster, collecting demographic data on individual members and household characteristics. Module B asked about enrollment and experiences with Smartcards. Module C asked about payments and involvement with the welfare programs, with separate modules for SSP and NREGS samples. Module D asked about consumption, Module E about income, Module F on assets and Module G on other household balance sheet items. Modules B and C, which asked about beneficiary experience with Smartcards and the welfare programs, were asked to the individual beneficiary herself, with separate sets collected for each individual beneficiary within the household. The other modules could be answered by either the male or female head of household.

Table B.1 describes in further detail the construction of each of the main outcome variables we report in the paper.

B.2.1 Matching household records to official records

As explained in detail in the section on sampling below, we sampled NREGS jobcards and individual SSP beneficiaries. Matching SSP beneficiaries to official records is straightforward since there is only one sampled beneficiary. Below we describe the process of matching NREGS official records with our household survey.

Complications may arise in this matching process because of two reasons. First, the set of household members as listed on the sampled NREGS jobcard may be different from the set of household members living under one roof that we surveyed. This complication is relatively easy to fix, as we know the names, ages, and genders of everyone listed on sampled jobcard as well as all members of the surveyed household. Although we surveyed every beneficiary living in the household about their NREGS employment, for our main leakage regressions (Table 3) we can match individuals by name and only include survey records for those individuals listed on the officially sampled jobcard.

The second complication is that the same surveyed household may have more than one jobcard, with potentially different sets of household members listed on each jobcard. This issue is more difficult to deal with, since reverse matching individuals from the surveyed household to the full set of jobcard records is close to impossible.

The following example illustrates these complications more concretely. Suppose that Karthik, Paul, and Sandip live in one household that is surveyed. Only Karthik and Paul are listed on the officially sampled jobcard (let's call it jobcard 1). For our main leakage regressions (Table 3), we do not include Sandip's reported work. It is also possible that (with or without their knowledge) Karthik, Paul, and Sandip are listed on a different jobcard (jobcard 2) that is not sampled. Reverse matching Karthik, Paul, and Sandip by name the to full jobcard list is basically impossible. In Section E.1 below we describe how we use a scaling factor to estimate overall leakage rates given that households may hold multiple jobcards.

B.3 Worksite audits

In addition to the household surveys in which we asked NREGS beneficiaries about their work experiences on the program, we also conducted "stealth" worksite audits in which an enumerator visited active worksites on a motorcycle during work hours and simply counted up the number of workers present. These visits happened precisely during the study period - May 28 to July 15 - that we asked about at the endline survey. The visits were conducted in 6 GPs per mandal - 5 GPs which also had household surveys, and 1 additional randomly sampled GP that was not part of our household survey. Thus we have one GP that was surveyed but not audited, and one GP that was audited and not surveyed, in order to test for effects of each activity on the other (see Section E.5 below for discussion of potential Hawthorne effects).

The stealth audit process was complicated by the fact that we did not want to rely too much on local officials to conduct it, and also because there is generally at least a two week delay in digitizing records and hence being able to electronically access the list of active worksites. Our procedure was to obtain the list of active worksites in a given GP from the official website, send an enumerator on a reconnaissance mission in which he asked villagers about the location of these worksites within the GP, but then wait about a week before the actual worksite visit in order to avoid any response by local officials to the reconnaissance mission itself. Given the lag in reporting and the fact that activity on worksites is fluid, we were not able to always find all listed and sampled worksites. However, the procedure followed was exactly the same in treatment and control mandals.

C Randomization, sampling, and attrition

C.1 Randomization

Under the terms of the MoU signed with the Government of Andhra Pradesh, we assigned the mandals in our eight study districts to treatment status as follows.

Our study districts contain a total of 405 mandals. Of these, we excluded 2 which were fully urban and so had no NREGS activity, 106 in which the government had already begun rolling out Smartcards at the time the MoU was signed, and 1 for which we were unable to obtain administrative data for stratification. We then randomized the remaining 296 mandals into three groups: treatment, buffer, and control. The government agreed to roll out treatment sequentially across those three groups: first in the treatment group, then in the buffer group, and finally in the control group. We included the buffer group in the design to ensure that we would have adequate time to collect endline data after Smartcards had deployed in treatment mandals, but before they deployed in control mandals.

Because the government was eager to roll out Smartcards quickly, they limited the number of mandals we could allocate to the control group relative to treatment in each district. Specifically, the government agreed to allocate 15 mandals to treatment and 6 to control in each of Adilabad, Anantapur, Khammam, Kurnool, Nellore, and Nalgonda; 12 to treatment and 5 to control in Kadapa, and 10 to treatment and 4 to control in Vizianagaram, for a total of 112 treatment mandals and 45 controls, with the remaining 139 mandals to be allocated to the buffer group. We assigned mandals to group by lottery, stratifying on revenue division (an administrative grouping of mandals within districts) and the first principal component of a vector of mandal characteristics. Revenue divisions do not serve a major administrative function but provided a convenient way to ensure geographic balance. Since integer constraints meant that we could not ensure that every revenue division has at least one treated and one control mandal, we do not include revenue division fixed effects but rather district fixed effects in our analysis (probability of treatment and control assignment is fixed within district). Including revenue division fixed effects rather than district fixed effects does not affect any of our results qualitatively. The mandal characteristics used were population, literacy rate, number of NREGS jobcards, peak season NREGS employment rate, proportion Scheduled Caste, proportion Scheduled Tribe, proportion SSP disability recipient, and proportion other SSP pension recipient.

Table C.1 reports balance on mandal characteristics from administrative data, including both variables we included in the stratification and others we did not. Unsurprisingly, the samples are well-balanced. Table C.2 reports balance on household characteristics from our baseline survey, which were not available at the time we conducted our randomization. Again the two samples appear well-balanced, with significant differences appearing no more often that would be expected by chance.

C.2 Sampling

For data collection activities we selected a total of 880 GPs: six GPs per mandal in six districts and four GPs per mandal in the remaining two. We sampled fewer GPs per mandal in the latter group because GoAP reallocated these two districts to new banks (and told us we could include them in the study) after we had already begun planning and budgeting, and our funding was limited. We sampled GPs using probability (approximately) proportional to size (PPS) sampling without replacement. As is well known, it is not possible to guarantee strict PPS sampling of more than one unit from a group as the probabilities implied by PPS may exceed one for large units; in these cases we top-coded sampling probabilities at one. A GP typically consists of a few distinct habitations, with an average of 3 habitations per GP; for logistical convenience we selected one habitation within each selected GP using strict PPS sampling.

We selected households within these habitations in the same way for baseline and endline surveys. We sampled a repeated cross-section (rather than a panel) of households to ensure that the endline sample was representative of program participants at that time. In each round of surveys we sampled a total of 10 households in each habitation, ensuring that a field team could complete surveys in one habitation per day. Of these we sampled 6 from the frame of NREGS jobcards and 4 from the frame of SSP recipients. Sampling in fixed proportions enabled our survey enumerators to specialize in administering NREGS or SSP survey modules. Finally, of the 6 NREGS jobcards we drew 5 from the list of households in which at least one member had worked during May-June according to official records and one household in which no member had worked. We over-sampled the former group in order to increase our precision in estimating leakage, since households that were not paid according to the official records are unlikely to have in fact received funds. At the same time we included some households from the latter group to ensure we could pick up treatment effects on access to work; sampling only among households that had participated in the NREGS would have precluded this. Note that treatment did not change the probability that a household was reported as working in the official data, nor did it change the number of days reported (Table C.3). Finally, we re-weight all our regressions using inverse sampling probabilities to ensure that all estimates are representative of the full frame of jobcards.

For our baseline survey we sampled 8,527 households, of which we were unable to survey or confirm existence of 1,000, while 102 households were confirmed as ghost households, leaving us with a final set of 7,425 households. The corresponding numbers for endline were 8,774 sampled, 287 not confirmed or surveyed, 8 physically missing surveys, and 365 households confirmed as ghosts, leaving us with 8,114 usable surveys with data. Tables C.4 and C.5 show that the households not confirmed or surveyed do not differ across treatment and control from the ones that were surveyed. The relatively high count of omitted households at baseline is due mainly to surveyor errors in coding the status of hard-to-locate households – for example, not confirming status of "ghost" households by writing down names of three neighbors willing to testify that no such household/beneficiary exists. Recognizing these difficulties we simplified the flowchart for coding household status so that in the endline survey we omitted far fewer households, and the 287 we do omit were nearly all left out because we were genuinely unable to trace them. In any case, we use the baseline data only to control for village-level means of outcome variables, so that non-completion of individual baseline surveys affects only the precision and not the consistency of our estimates. Note that ghost households in whose name official payments are made *will* be included in our leakage regressions, increasing observation count in those regressions.

C.3 Sampling frame turnover

The databases of beneficiaries from which we sample (NREGS jobcards and SSP pensioners) evolve over time as new records are created and old ones removed. New jobcards are created in response to applications from eligible (i.e. rural) households; old records may be removed from the database when someone dies, migrates out of state, or when families change structure (e.g. divorce) or separate (e.g. joint household splits), in which case each new household gets a new jobcard and old ones are removed. In the case of the SSP, new pensioners are recorded as they are moved off of waiting lists onto active lists, and old pensioners are removed when they die or migrate.

Because of these sources of churn, and because we sample a repeated cross-section of households from the NREGS and SSP frames, it is possible that our estimates of treatment effects confound the effects of Smartcards on a *given* participant with effects on the *composition* of participants. To examine this we test for differences by treatment status in the rate or composition of change in each of our two sampling frames.

In control mandals, 2.4% of NREGS jobcards that were in our baseline frame drop out by endline sampling. On the other hand, 5.9% of jobcards in the endline frame are new entrants. Neither of these rates are significantly different in treatment mandals (Table C.6a) and there is also no difference in the total number of jobcards across treatment and control mandals (Table C.7). This is not particularly surprising as most potential NREGS participants likely had job cards already by the time of Smartcards rolled out: 65% of rural households in Andhra Pradesh had jobcards as of 2010 (authors calculations using National Sample Survey Round 66 (2009-2010)).

Turning to the SSP frame, churn rates are somewhat higher (9.7% dropout rate and 16% entrance rate) but again balanced across treatment and control (Table C.6b). Moreover, new entrants to both frames are similar across control and treatment on demographics (household size, caste, religion, education) and socioeconomics (income, consumption, poverty status) for both NREGS and SSP programs (Table C.8). Finally, the households surveyed at baseline are similar to households surveyed at endline on socio-demographic characteristics such as age composition, literacy, and religion (Table C.9). These results suggest that exposure to the Smartcard treatment did not affect the size or the composition of the frame of potential program participants.

D Correlates of Smartcard Implementation

This section presents and discusses the correlates of Smartcard implementation at various levels. We start with the selection of districts for the evaluation, and compare them to other districts in the state to assess the extent to which our study districts are representative. Within these districts, the introduction of Smartcards was randomized at the mandal (sub-district) level. However, not all treatment mandals actually implemented Smartcards; within implementing mandals, not all villages converted to the Smartcards-based payment system; and within converted villages, not all households obtained a Smartcard. This is why our experimental analysis focuses on the intent to treat estimates. Nevertheless, it is of independent interest to understand the correlates of program implementation, as it may help predict roadblocks in implementation elsewhere. We show these results below.

D.1 Districts

As mentioned earlier, the eight study districts were not randomly chosen. Table D.2 (extended version of previously submitted table) compares the study districts to the other rural districts of AP (since NREGS was only implemented in rural areas). Overall, we see that study districts have a slightly lower rural population, but are otherwise similar to the non-study districts on several indicators including demographics, the fraction of agricultural laborers, and village-level facilities, suggesting that our estimates are likely to generalize to all of rural Andhra Pradesh. These similarities also suggest that the main reason for the non-performance of the banks who had initially been assigned these districts was related to bank-specific factors as opposed to district-specific ones .[38]

D.2 Mandals

While mandals that were randomized into treatment status were all supposed to be converted to the Smartcard-based payment system over the course of two-years, in practice only 80% of the mandals got converted (defined as having at least one GP that had converted to the new system). Table D.3 presents correlations between baseline characteristics at the mandal-level and whether a mandal was converted to the new system for NREGS (columns 1-4) and SSP (columns 5-8). We present coefficients from both binary and multiple regressions, and look at both the extensive margin (whether a mandal had converted) and the intensive margin (the fraction of GP's converted).

Overall, we find no noticeable pattern in mandals getting converted for NREGS payments, except that mandals that got converted had slightly lower baseline levels of time to collect payments. For SSP however, we see that mandals that had a higher proportion of residents below the poverty line (BPL) and had a higher total volume of payments were more likely to get converted, and converted more GP's.

D.3 Villages (GPs)

We find a similar set of correlations with whether a village got converted to the Smartcard system and with the treatment intensity (defined as the fraction of total transactions that are conducted with carded beneficiaries). Table D.4 shows these correlations, and we see that villages with a higher fraction of BPL population were more likely to be carded for both NREGS and SSP and that villages with a larger total amount of SSP payments were more likely to be converted.

D.4 Households

Finally, we present individual-level correlates of having a Smartcard in Table D.5. A similar pattern to the village-level correlates emerges at the individual level for the NREGS, with more vulnerable (lower income, female, scheduled caste, and being more active in NREGS) beneficiaries more likely to have Smartcards. No such pattern is seen for SSP households

[38]One example of such a bank-specific challenge was the quality of the Bank-TSP partnership. An important reason for non-implementation of Smartcards in some districts was that the banks and TSP's (who were jointly awarded the Smartcard contract for the district) were not able to manage their contracts, commitments, and commissions adequately, which stalled implementation in these districts. Such challenges were more likely to be a function of the organizations rather than a function of specific districts (see Mukhopadhyay et al. (2013)for more details on implementation challenges).

(perhaps because all participants are vulnerable to begin with, whereas NREGS is a demand-driven program).

Overall, the results in this section are consistent with the idea that banks prioritized enrolling in mandals and GPs with more program beneficiaries and hence more potential commission revenue, while conditional on a village being converted the more active welfare participants were more likely to enroll. Further, since enrollment typically took place in short-duration camps (typically lasting 1-2 days) that beneficiaries had to attend to get enrolled, villages with more (potential) beneficiaries may have also had a greater incentive to make sure that beneficiaries were informed about these camps and encouraged to enroll for a Smartcard.

E Further leakage results and robustness

E.1 Estimating average leakage

As discussed in the text, we cannot estimate average levels of leakage in our data by simply comparing receipts per *household* with official disbursements per *jobcard*, since there are many more jobcards in Andhra Pradesh than there are households with at least one jobcard. In this section we illustrate with an example how this affects our calculations, and explain in detail how we correct for it.

To illustrate the problem, return to the example introduced earlier in Section B.2.1, where Karthik, Paul, and Sandip form one surveyed household that has two jobcards. Figure E.1 depicts a situation where we sampled Jobcard 1, which only has partial records of payments to Karthik and Paul, but not Jobcard 2, which has additional details of payments made to Paul and Sandip. Actual leakage is the sum of all payments made to the household (Jobcard 1 + Jobcard 2 = 30 + 35 + 50 = 115) minus total receipts by the household ($30 + 20 + 40 = 90$), which equals Rs. 25. If we naively compared household earnings to jobcard disbursements, however, our estimate of leakage would be Rs. -60. Even if we matched workers by name (as we do for all the analysis in the main paper) and removed Sandip, who is not listed on Jobcard 1, we would still under-estimate leakage at Rs. -20.

In principle one possible solution to this problem would be to find Jobcard 2 in the official data, but in practice this is infeasible as it would involve trying to reverse match by name across a very large number of records. Reliably making such matches is particularly difficult given the frequency of misspellings, alternative spellings, errors in transliteration, and similarities between names that are actually different, and by the fact that we do not know what (sub)set of family members may be listed on any given jobcard. We therefore focus instead on adjusting our estimates for the rate at which we under-sample jobcards relative to households. If we knew that the household in this example had two jobcards, we could simply multiply our estimate of official disbursements by 2 to obtain a corrected estimate of total disbursements to the household. While this would not necessarily calculate the correct amount disbursed given that we sampled Jobcard 1, it does yield the correct amount in expectation since we are equally likely to sample Jobcard 1 or Jobcard 2.

The challenge with this approach is that we do not know how many jobcards are associated with any given household. There are two ways we can potentially deal with this: we can estimate the average number of jobcards per household, or ask households directly how many jobcards they have. The latter approach gives us household-specific answers and so is likely to be more precise, but this comes at the cost of three sources of bias. First, households need not know about all the job cards issued in their name, especially cards created by officials for the express purpose of stealing money. Second, households that do have multiple jobcards would possibly be uncomfortable reporting this, as by law each household should have a single jobcard. Finally, our survey methodology may have led to undercounting jobcards; the question that asked about the number of jobcards accompanied instructions to produce jobcards in order to write down the jobcard number, and if all household jobcards were not physically available at the time of the survey, it is possible that enumerators may not have counted them.

Given these biases, a more reliable way of estimating the ratio of jobcards to households is to use independent, representative records from the National Sample Survey, which we can use to estimate the number of jobcards per household at more aggregate levels. We do this at the district level and estimate an average ratio of 1.9 jobcards per household holding at least one jobcard. (In contrast, surveyed households reported 1.2 jobcards on average to us.) We then scale up official payments to each household using the scaling factor specific to their district. For comparison we calculate the earnings reported by *all* workers in the same household (not just those matched to sampled jobcards, as we do in the main analysis).

The downside of this approach is of course that it introduces a substantial source of noise into the dependent variable and our estimates in order to achieve consistency.[39] To see why, consider a typical household with two jobcards, A and B, on which amounts Y_A and Y_B are paid out. Suppose for purposes of illustration that these variables are iid. If we observed both then the variance of our estimate of the total would be $Var(Y_A + Y_B) = 2Var(Y_A)$. But since we only observe Y_A and have to estimate $Y_A + Y_B$ using $2 \times Y_A$, the variance of

[39]Note that this procedure is not mechanically affected by treatment, as the introduction of Smartcards did not affect the number of jobcards (Table C.7). While the biometric data collected during Smartcard enrollment was intended to be used to de-duplicate the beneficiary database, this was never done as Smartcard enrollment was still far from complete and many jobcards could not be linked to a Smartcard.

our estimate is now $Var(2 \times Y_A) = 4Var(Y_A)$. In other words, our precision is half what it would be if we know both the jobcards associated with the household, as opposed to just one of them.

Using this method, we estimate an average leakage rate of Rs. 80 per household, or 30.7% of average official outlays (Table E.1). We also estimate treatment effects on official and actual payments as well as leakage which are similar to the main results, albeit noisier, with the p-value of the treatment effect on leakage equal to 0.18 (column 7). This is unsurprising given that scaling gives us an unbiased estimate of average leakage, but an inefficient test for changes in leakage relative to the test in Table 3a. We can improve the precision by exploiting the fact that for official payments we observe the jobcard-specific baseline value, and not just the GP average (as we do for actual payments). Since auto-correlation in official payments over time is clearly higher at the jobcard level than at the GP level, this provides a meaningful increase in precision. Controlling for these jobcard specific values reduces the p-value on our leakage estimates to 0.11 (column 8) and increases the magnitude of the estimated coefficient.[40]

E.2 Collusion and recall

The main threat to the validity of the leakage results is differential mis-reporting on our survey across treatment and control areas. This may be possible for a number of reasons. First, survey respondents might collude with officials and thus report higher payments than they should have received, and this collusion increases with treatment. Second, treatment may differentially affect recall, if for example respondents in treatment areas are able to better remember payment amounts, or pay more attention because the Smartcards intervention makes payments more salient.

We assuage both concerns through a number of methods. We first report results that suggest both collusion or recall bias are unlikely, and then point to indicators that separately rule out either collusion or recall bias.

Our first piece of evidence comes from the quantile plot of survey payments. As Figure 3 shows, we see a significant increase only in payments received by those who would have otherwise received no payments (relative to the control group). Since there is no reason to expect collusion only with this sub-group (if anything, it would arguably be easier for officials to collude with workers with whom they were already transacting), this pattern

[40]Note that controlling for the jobcard-specific baseline value makes no difference to our main results. While it reduces magnitude and increases precision of impact on official payments (so that there is an even more precise zero result), it does not meaningfully change leakage results. We therefore stick with standard specification that uses baseline GP-level means in Table3a for simplicity and consistency with the rest of the main tables.

seems harder to reconcile with a collusion-based explanation. Similarly, it is highly unlikely the recall bias takes the form of respondents in treatment areas suddenly remembering that they had worked some versus not worked at all, given how salient NREGS is in the lives of these workers; a more plausible explanation involving recall bias would suggest respondents remember the actual payment more accurately.

Second, we conducted independent audits of NREGS worksites in treatment and control mandals during our endline surveys, and counted the number of workers who were present during unannounced visits to worksites. As described in Section B.3 above, these measures are somewhat noisy. However, we do see an insignificant 39.3% increase in the number of workers found on worksites in treatment areas during our audits (Table E.2), and cannot reject that this is equal to the 24% increase in survey payments reported in Table 3a. Thus, the audits suggest that the increase in survey payments reported are proportional to the increase in workers found at the worksites during our audits, indicating that misreporting either because of collusion or recall bias is unlikely.

Next, we directly test for differential rates of false survey responses by asking survey respondents to indicate whether they had ever been asked to lie about NREGS participation, using the "list method" to elicit mean rates of being asked to lie without forcing any individual to reveal their answer. The list method is a standard device for eliciting sensitive information and allows the researcher to estimate population average incidence rates for the sensitive question, though the answers cannot be attributed at the respondent level (Raghavarao and Federer, 1979; Coffman et al., 2013). We present a subset of respondents with the following statement - "Members of this household have been asked by officials to lie about the amount of work they did on NREGS") - but respondents do not respond directly about whether the agree with the statement; instead they are also presented with five other statements, and asked to tell us how many of the statements they would agree with. A second subset of respondents is presented with the other five statements, but not the sensitive statement. A third subset is presented with the other five statements along with a statement they would certainly disagree with (in order to determine whether simply presenting more statements leads to more "yes" responses). This statement says "Members of this household have been given the chance to meet with the CM of AP to discuss problems with NREGS." We can then compare the differences in numbers between the first and second groups in treatment and control areas, while adjusting for any increases coming purely from the increase in question numbers. Using simply the differences in numbers between the first and second subsets, we find that at most 15% of control group respondents report having been asked to lie and find no significant difference between the treatment and control groups on this measure (Table E.3). However, data from the third subset suggests that simply asking more questions leads to more "yes" responses, so it is possible that no one in the control group may have been asked to lie.

Other indicators also rule out differential collusion. We saw that beneficiaries overwhelmingly prefer the new payment system to the old, which would be unlikely if officials were capturing most of the gains. Finally, we find evidence that Smartcards increased wages in the *private* sector, consistent with the interpretation that it made NREGS employment a more remunerative alternative, and a more credible outside option for workers (see section 5).

With respect to differential recall, we paid close attention to the measurement of data on NREGS employment, learning from and improving on our previous work on this issue (Niehaus and Sukhtankar, 2013a,b). One of the main methods through which we helped respondents recall is the recording of work in the physical jobcard. Neither the format nor the recording of jobcard entries were affected by treatment, and hence differential recall bias appears a priori unlikely. Moreover, the average treatment GP had been treated for 14.5 months (or 2 full NREGS seasons), hence the Smartcards intervention was not that new. Most concretely, we can use the fact that our survey was spread over two months to check whether there was indeed differential recall. If differential recall is driving our results, then, holding constant the week in which work was actually done, the estimated treatment effect on leakage should be more negative (higher in magnitude) if the survey was conducted with a greater lag as opposed to a shorter lag after actual work. Table E.4 shows that there is no consistent pattern across survey weeks, suggesting that survey lag and differential recall bias do not affect our results.

E.3 Spillovers

E.3.1 Geographic and strategic spillovers

While the main estimates in the paper assume that program performance in a given mandal depends only on that mandal's treatment status, it is possible that our outcomes are also affected by the treatment status in adjacent mandals. Spillovers effects that are "positive" (i.e. have the same sign as direct treatment effects) will simply lead us to under-estimate the direct effects, but spillovers that are "negative" (i.e. opposite sign as direct effects) could lead us to over-estimate the direct effects. For example, if officials in control mandals hear about Smartcards and try to steal more in anticipation of future rollout, we could over-estimate effects on corruption.

First, we note that we see no reallocation of funds away from treated mandals towards control mandals – average official outlays in the two track each other closely from baseline to endline (Figure 1). This is inconsistent with spillover effects in which senior officials route funds to the places where they are easiest to steal.

In addition, we test for spatial spillovers. We first construct a measure of exposure to treatment in the neighborhood of each GP. Specifically, we calculate the fraction of neighboring GPs that are (i) within a radius R of the given GP, and (ii) located in a different mandal, that are treated. We impose condition (ii) because the treatment status of neighboring GPs in the same mandal is identical to own treatment status, so we cannot separately identify their effects.

Tables E.5, E.6, and E.7 report results from this estimation for the payment process and leakage, with NREGS and SSP outcomes separately. Consistent with the fact that the main unit of program implementation is the village (GP), there are no spillovers on the payment process, while the treatment effect remains invariant to the inclusion of our measure of exposure. Moreover, there is no evidence of an effect of neighbors' treatment status on leakage in either NREGS or SSP.

E.3.2 Spillovers to other parts of program budget

Our estimates of leakage are entirely focused on the NREGS labor budget, since Smartcards affected wage payments. It is possible that while leakage from the labor budget is reduced, leakage is displaced to other parts of the overall NREGS budget. In order to test for this possibility, we collected NREGS budget data disaggregated by category for the months of May, June, and July 2010 and 2012.

To begin with, the data support our decision to focus on the labor budget, as the labor budget is over 91% of the overall budget. This suggests that displacement effects, if any, will be limited. There are no statistically significant effects of treatment on other areas of the budget such as materials or contingency expenses (Table E.8). While we cannot directly measure leakage, since we do not measure actual materials expenditure, the fact that official material expenses did not increase suggests that there was no large-scale displacement.

E.4 Payment timing

A further concern is that survey reports simply reflect the fact that treatment reduced payment delays, so more respondents in treated areas would have been paid at the time of survey, rather than a reduction in leakage. While we minimized this risk by surveying households an average of ten weeks after NREGS work was completed (while the mean payment delay is five weeks), it is still possible that some households had not been paid by the time we surveyed. Since we asked respondents when exactly they got paid for each spell of work, as well as whether they have been paid yet for the spell in question, we can simply verify that the rate of completed payments was identical across treatment and control mandals (Table E.2).

E.5 Hawthorne effects

A final concern might be that the various types of data collection activities affect the reporting of survey or official data. For example, it is possible that officials or workers noticed our stealth auditors, and somehow connected them to our survey (which took place an average of ten weeks after NREGS work was completed), and adjusted their reporting of official quantities or survey responses. We carefully designed our data collection procedures to test for this possibility. First, we can check using the full set of official records whether official payment quantities are affected by the presence of our auditors or surveyors in the village (by comparing villages sampled for these activities to those not sampled). As Table E.9 shows, there is no evidence of effects on official reports. Note that each cell in the table reports results from a separate regression, testing whether conducting audits or surveys overall in a GP affected official records, as well as separately whether reports from that particular week were affected (in case there was only a short-term response). Since these regressions include the full set of official muster data, we can see that the effects are precisely measured and close to zero.

Second, as Section B.3 described we conducted audits in 5 out of 6 surveyed GPs, and conducted surveys in 5 out of 6 audited GPs, allowing us a comparison GP in each case. Again, Table E.9 shows that there is no evidence of either activity affecting the other. Admittedly the results here are somewhat noisy given limited power, but we have no evidence - quantitative or anecdotal - to suggest that our data collection itself affected measurement.

F Heterogeneity by baseline characteristics

The two main dimensions of heterogeneous impacts we focus on in the text are the non-parametric plots of quantile treatment effects, and linear interactions between the treatment and the baseline value of the outcome for each outcome studied (4.4). In addition, we also examine heterogeneity of impact along other measures of vulnerability such as consumption, measures of socio-economic disadvantage (fraction of the BPL population and belonging to historically-disadvantaged scheduled castes (SC)), as well as the importance of the program to the village (official amounts paid).

Overall, we find little consistent evidence of heterogeneity of program impact (Table F.1).

Two out of 20 tests in Panel A (NREGS) are significantly different from zero at the 10% level, which is the expected rate of rejection under a null hypothesis of no significant heterogeneity of impacts. Similarly, for SSP we find no evidence of heterogeneous impacts for either official or survey payments. The only suggestive evidence of heterogeneity is for the time to collect SSP payments but there is no clear pattern here. Time to collect appears to have gone down more in villages that had higher consumption, but also in villages with a greater BPL proportion. We also plot the quantile treatment effects on the time take to collect SSP payments in Figure F.1 and see no significant impact at any percentile of the endline distribution of time to collect payments, which is not surprising given the lack of impact on the mean time to collect SSP payments.

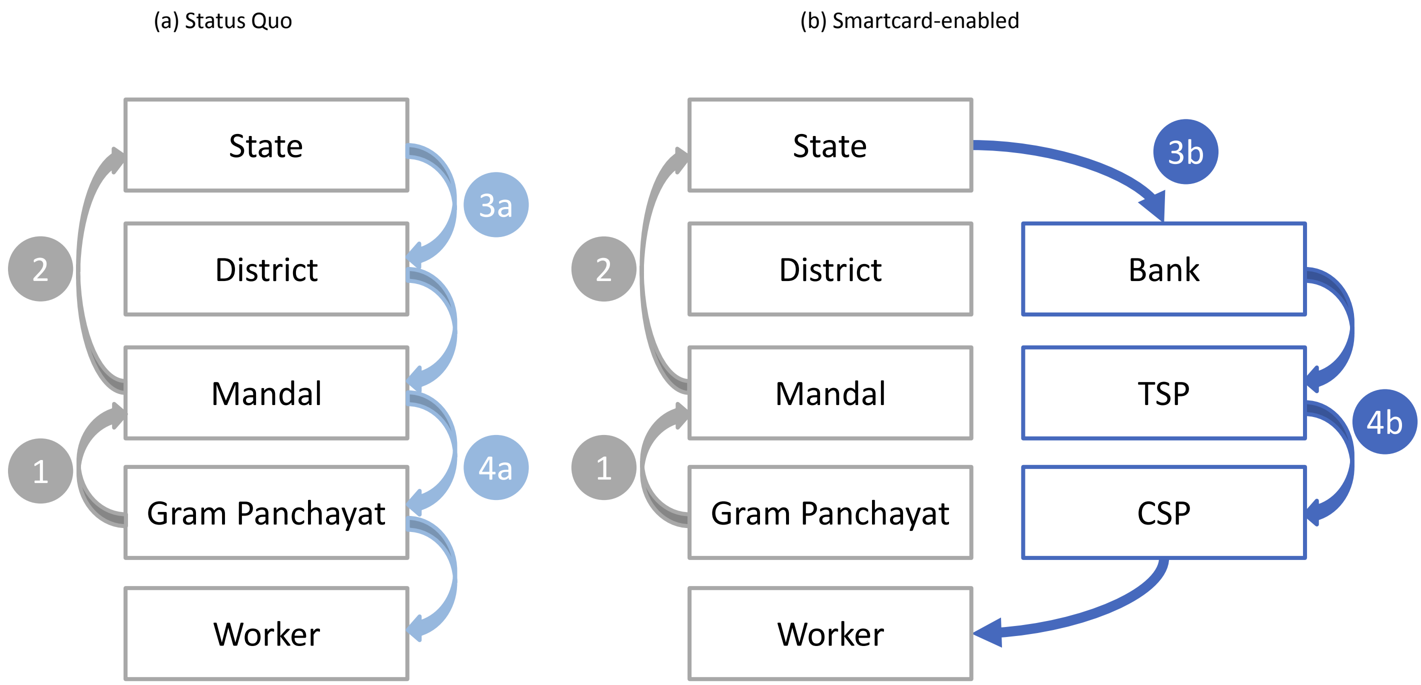


Figure A.1: Comparison of treatment and control payment systems

"TSP" is a Technology Service Provider, a firm contracted by the bank to handle details of electronic transfers. "CSP" is a Customer Service Provider, from whom beneficiaries receive cash payments after authentication. In both systems, (1) paper muster rolls are maintained by the GP and sent to the mandal computer center, and (2) the digitized muster roll data is sent to the state financial system. In the status quo model, (3a) the money is transferred electronically from state to district to mandal, and (4a) the paper money is delivered to the GP (typically via post office) and then to the workers. In the Smartcard-enabled system, (3b) the money is transferred electronically from the state to the bank to the TSP, who transfers cash to the CSP, and (4b) the CSP delivers the cash and receipts to authenticated recipients.

(a) Sample Smartcard

(b) Point-of-Service device

Figure A.2: The technology

Table B.1: Data sources and variable construction for main outcomes

Variable used in Table(s)	Description	Source
Carded GP (T1)	A GP is considered "carded" or "converted" when payments are moved to the Smartcards-based payment system run by Banks/TSPs. This happens separately for NREGS and SSP, usually when 40% of beneficiaries of each program are issued Smartcards. The outcome for SSP is based on May 2012 data as June 2012 data were not available.	DoRD; some missing data downloaded from program websites
Mean fraction carded payments (T1)	This refers to the fraction of payments to NREGS beneficiaries/pension recipients who had Smartcards, averaged over May and June 2012 (NREGS) or May 2012 (SSP, see above).	Program websites
Payments generally carded (village mean) (T1)	Hhd survey section B asks individuals whether they usually swipe their Smartcards or use their fingerprints as ID to collect NREGS/SSP payments. Using habitation and GP sampling probabilities, we construct a weighted GP average for the proportion of payments generally carded.	Hhd survey, section B
Most recent payments carded (village mean) (T1)	Analogous to the question above, only now the outcome question is whether an individual swiped the card or used finger prints when they *last* collected a payment.	Hhd survey, section B
Time to Collect (Min) (T2, T8, F3)	The average duration of time taken to collect payments (including unsuccessful trips). This question is asked once per survey to individual NREGS/SSP beneficiaries. We discarded missing values (not replacing them with zeros).	Hhd survey, section C
Payment Lag: Ave Payment Delay (T2, T8, F3)	This outcome is available for NREGS only, and is constructed for each week that each individual beneficiary worked in the endline study period. We ask for the date (e.g. 7/17/12) that the individual collected her NREGS payment for each study week. We then calculate the # days between the end of the respective study week (e.g. 6/24/12) and the date of the payment. In this example the payment delay is 22 days.	Hhd survey, section C
Payment Lag: Deviation (T2)	Using the average payment delay at the individual-week level, we calculate the mandal-week median payment delay. The outcome is then absolute value of the difference between individual payment delay in week w and the mandal median delay in week w.	Hhd survey, section C
Official (NREGS) (T3, T8, F3)	Our data include start date, end date, and amount paid for every work spell in our study mandals for baseline and endline years. We assign officially recorded spells to correspond to survey study weeks, obtain average weekly payment by dividing by the # of endline study weeks (7), and aggregate data at the household level. We include in our official measure payments made to ghost households, but do not include sampled jobcards who we were unable to find in our household survey exercise.	Tata Consultancy Services
Survey (NREGS) (T3, T8, F3)	We ask every individual NREGS beneficiary in the household about details of work done and payment received for each of the study weeks, generate average weekly receipts and aggregate data at the household level. We only include payments received by individuals who are listed on the sampled official jobcard. We include payments made to ghost households as 0.	Hhd survey, section C
Official (SSP) (T3, F3)	The SSP data lists every monthly disbursement made to beneficiaries. We take the average disbursement across the months of May, June and July 2012 as the outcome variable. We include in our official measure payments made to ghost beneficiaries.	DoRD

Survey (SSP) (T3, F3)	We ask SSP receipts how much their pension is supposed to pay every month, and subtract from this amount their reports of bribes they paid to obtain the payment or reductions from the payment amount. We only include data for the sampled SSP beneficiary corresponding to disbursement records.	Hhd survey, section C
Leakage (T3, T8, F3)	Leakage is defined as the difference between "official" and "survey" payments for both NREGS and SSP.	
Ghost hhds (T5)	Ghost households are households (or sampled individual beneficiaries within households) who either do not exist or had permanently migrated before the start of the NREGS study period/ SSP disbursement period. The non-existence, death, or permanent migration must be confirmed by 3 neighbors. Households who had temporarily migrated, were confirmed by neighbors to be "in the area" but could not be found, or whose status could not be confirmed are not considered ghost households and are simply excluded from the analysis.	Hhd survey
Other overreporting: NREGS (T5)	Consider the set of complete surveys, i.e. exclude ghosts, and the outcome variables "official" and "survey" as described above. If "official" is positive but "survey" is 0, the overreporting indicator is 1; otherwise it is 0. We construct this variable at the household level, and exclude any households with 0 official payments in the study period.	
Other overreporting: SSP (T5)	The difference between what the official records ("official") list as disbursed and what a pensioner thinks she is entitled to is the amount overreported	
Bribe to Collect (NREGS) (T5)	We ask each individual beneficiary if she had to pay anything in order collect her earnings for the most recent work spell. We limit estimation to beneficiaries who worked in the study period.	Hhd survey, section C
Underpayment (SSP) (T5)	We ask the sampled pensioner whether she usually has to pay a bribe to receive benefits, or equivalently whether any amount is deducted from the pension amount.	Hhd survey, section C
Proportion of Hhds doing NREGS work (T6)	We classify a individual as working if "survey" (see above) is positive. A household is classified as working if *any* member is classified as working.	Hhd survey, section C
Was any Hhd member unable to get NREGS work in May/December (T6)	This question is asked only once per household to the head of the household. The question asks whether any member of the household wanted to work on NREGS during the month of May (or December) but was unable to obtain employment.	Hhd survey, section C
Is NREGS work available when anyone wants it (T6)	This question is also asked only once per household. It ask whether "in general, can anyone in this village who wants work on NREGS get it?"	Hhd survey, section C
Did you have to pay anything to get this NREGS work (T6)	We ask individuals who worked in the study period whether they had to pay anything to obtain work in the study period.	Hhd survey, section C
Did you have to pay anything to receive this pension (T6)	We ask pension beneficiaries who started received their pension after baseline whether they had to pay anything to start receiving their pension.	Hhd survey, section C
Official disbursement trends in NREGS (F1)	Using the comprehensive muster roll dataset that contains information on all work spells done on NREGS, we aggregate the data to the GP-week level. We take weekly totals (weeks defined by the start of work spell) across all work spells within a GP. The time series plotted in the Figure is the sum across all GPs in a given week.	Tata Consultancy Services

% (converted) Mandals % (converted) GPs (F2)	A mandal is classified as converted if at least one GP within that mandal is listed as converted. A GP is classified as converted when payments are moved to the Smartcards-based payment system run by Banks/TSPs, usually when 40 % of beneficiaries have been issued a Smartcard. Data are restricted to treatment mandals for both lines	DoRD; some missing data downloaded from program websites
% Carded Payments (F2)	Within converted GPs, we obtain the ratio of carded payments to total payments each month, and create an average of the proportion of carded payments across treatment mandals in the 8 study districts. We multiply this average by the fraction of converted GPs in converted mandals, and by the fraction of converted mandals (see above row), to obtain the overall % of carded payments.	Program websites

"DoRD": Department of Rural Development, Government of Andhra Pradesh

Program websites: `www.nrega.ap.gov.in` and `www.rd.ap.gov.in`

Table C.1: Balance on baseline characteristics

	Treatment	Control	Difference	p-value
	(1)	(2)	(3)	(4)
	Numbers based on official records from GoAP in 2010			
% population working	.53	.52	.0062	.47
% male	.51	.51	.00023	.82
Literacy rate	.45	.45	.0043	.65
% SC	.19	.19	.0025	.81
% ST	.1	.12	-.016	.42
Jobcards per capita	.54	.55	-.0098	.63
Pensions per capita	.12	.12	.0015	.69
% old age pensions	.48	.49	-.012	.11
% weaver pensions	.0088	.011	-.0018	.63
% disabled pensions	.1	.1	.0012	.72
% widow pensions	.21	.2	$.013^{**}$	.039
	Numbers based on 2011 census rural totals			
Population	45580	45758	-221	.91
% population under age 6	.11	.11	-.00075	.65
% agricultural laborers	.23	.23	-.0049	.59
% female agri. laborers	.12	.12	-.0032	.52
% marginal agri. laborers	.071	.063	.0081	.14
	Numbers based on 2001 census village directory			
# primary schools per village	2.9	3.2	-.28	.3
% village with medical facility	.67	.71	-.035	.37
% villages with tap water	.59	.6	-.007	.88
% villages with banking facility	.12	.16	$-.034^{**}$	.021
% villages with paved road access	.8	.81	-.0082	.82
Avg. village size in acres	3392	3727	-336	.35

This table compares official data on baseline characteristics across treated and control mandals. Column 3 reports the difference in treatment and control means, while column 4 reports the p-value on the treatment indicator from simple regressions of the outcome with district fixed effects as the only controls. A "jobcard" is a household level official enrollment document for the NREGS program. "SC" ("ST") refers to Scheduled Castes (Tribes), historically discriminated-against sections of the population now accorded special status and affirmative action benefits under the Indian Constitution. "Old age", "weaver", "disabled" and "widow" are different eligibility groups within the SSP administration. "Working" is defined as the participatin in any economically productive activity with or without compensation, wages or profit. "Main" workers are defined as those who engaged in any economically productive work for more than 183 days in a year. "Marginal" workers are those for whom the period they engaged in economically productive work does not exceed 182 days. The definitions are from the official census documentation. Standard errors are clustered at the mandal level. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.2: Balance on baseline characteristics: household survey

	NREGS				SSP			
	Treatment	Control	Difference	p-value	Treatment	Control	Difference	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hhd members	4.8	4.8	.022	.89	4.1	4.2	-.15	.41
BPL	.98	.98	.0042	.73	.98	.97	.0039	.65
Scheduled caste	.22	.25	-.027	.35	.19	.23	-.038*	.08
Scheduled tribe	.12	.11	.0071	.81	.097	.12	-.022	.46
Literacy	.42	.42	.0015	.93	.38	.39	-.012	.42
Annual income	41,482	42,791	-1,290	.52	33,622	35,279	-2,078	.34
Annual consumption	104,717	95,281	8,800	.39	74,612	77,148	-3,342	.56
Pay to work/enroll	.011	.0095	.00099	.82	.054	.07	-.016	.26
Pay to collect	.058	.036	.023	.13	.06	.072	-.0078	.81
Ghost Hhd	.03	.017	.013	.14	.012	.0096	.0019	.75
Time to collect	156	169	-7.5	.62	94	112	-18**	.03
Owns land	.65	.6	.058*	.06	.52	.48	.039	.18
Total savings	5,863	5,620	3.7	1.00	4,348	3,670	729	.30
Accessible (in 48h) savings	800	898	-105	.68	704	9,576	-9,211	.29
Total loans	62,065	57,878	5,176	.32	43,161	43,266	-813	.81
Owns business	.21	.16	.048**	.02	.16	.19	-.025	.29
Number of vehicles	.11	.12	-.014	.49	.1	.093	.0039	.83
Average Payment Delay	28	23	.036	.99				
Payment delay deviation	11	8.8	-.52	.72				
Official amount	167	159	12	.51				
Survey amount	171	185	-13	.55				
Leakage	-3.8	-26	25	.14				
NREGS availability	.47	.56	-.1**	.02				
Hhd doing NREGS work	.41	.41	.000024	1.00				
NREGS days worked, June	8.3	8	.33	.65				
NREGS hourly wage, June	13	14	-1.3	.13				
NREGS overreporting	.15	.17	-.015	.55				
# addi. days hhd wanted NREGS work	15	16	-.8	.67				

This table compares household survey data on baseline characteristics across treatment and control mandals. Columns 3 and 6 report the difference in treatment and control means, while columns 4 and 8 report the p-value on the treatment indicator, all from simple regressions of the outcome with district fixed effects as the only controls. "BPL" is an indicator for households below the poverty line. "Pay to work/enroll" refers to bribes paid in order to obtain NREGS work or to start receiving SSP pension. "Pay to Collect" refers to bribes paid in order to receive payments. "Ghost HHD" is a household with a beneficiary who does not exist (confirmed by three neighbors) but is listed as receiving payment on official records. "Time to Collect" is the time taken on average to collect a benefit payment, including the time spent on unsuccessful trips to payment sites, in minutes. "Accessible (in 48h) savings" is the amount of savings a household could access within 48h. "Payment delay deviation" is the absolute value of the difference between an individuals payment delay and the mandal median. "NREGS availability" is an indicator for whether a household believes that anybody in the village could get work on NREGS when they want it. "NREGS overreporting" is the incidence of jobcards that had positive official payments reported but zero survey amounts (not including ghosts). Standard errors are clustered at the mandal level. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.3: Impacts on official records of NREGS participation

	Worked on NREGS (%)		Days worked on NREGS	
	(1)	(2)	(3)	(4)
Treatment	.015	.016	.32	.39
	(.016)	(.016)	(.32)	(.32)
District FE	Yes	Yes	Yes	Yes
Adj R-squared	.03	.03	.04	.02
Control Mean	.4	.36	5.9	4.9
N. of cases	2116302	900404	2116302	900404
Level	Hhd	Hhd	Hhd	Hhd
Data used	All GPs	Survey GPs	All GPs	Survey GPs

This table analyzes whether treatment affected the extensive margin of work reported in official records. The unit of analysis is the jobcard. The outcome in columns 1 and 2 is a binary variable equal to 1 if any household member listed on the jobcard is reported to have worked in the endline study period between May 28 and July 15, 2012. The outcome in columns 3 and 4 is the number of household-days worked during the same period as recorded on the official jobcard. Columns 2 and 4 restrict the sample to the 880 GPs sampled for the household survey. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization as control variable as well as district fixed effects. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.4: Comparing surveyed and non-surveyed sampled households - NREGS

	# Members	% Female	Ave. age	ST/SC	Worked in May	#BL spells per member	#EL spells per member	Avg. implied daily wage BL	Avg. implied daily wage EL
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment	-.038	-.011	-.57**	-.017	-.024	-.0026	.0053	-3.1	-3.7
	(.052)	(.01)	(.28)	(.028)	(.03)	(.01)	(.013)	(1.9)	(2.5)
Non-surveyed hhd	-.31*	-.00076	-.095	-.023	-.27***	-.048	-.064	2.5	1.2
	(.17)	(.052)	(1.1)	(.057)	(.068)	(.035)	(.047)	(3.9)	(5.3)
Treatment X non-surveyed hhd	-.35*	.014	1.4	.029	.025	.058	-.027	2.4	-7.5
	(.19)	(.061)	(1.4)	(.079)	(.084)	(.044)	(.052)	(5.6)	(6.1)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.01	.01	.01	.05	.03	.01	.01	.10	.10
Control Mean	2.6	.25	37	.4	.6	.15	.21	97	106
N. of cases	5078	5078	5078	5078	5078	5078	5078	1716	2450
Level	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd
Unit	Number	%	Years	%	%	Number	Number	Rs.	Rs.

This table compares sampled NREGS households who were surveyed to sampled NREGS households who could not be surveyed (excluding confirmed ghost households), using official data. Reasons for missing surveys could be temporary migration, repeated absence on survey dates or refusal to participate in the survey. There were 4943 completed and 135 unsuccessful surveys. "Treatment X non-surveyed hhd" is an interaction term. All outcomes are taken from official jobcard records (demographics of workers listed on jobcard) and muster rolls (information on work spells completed by members on the jobcard). "Worked in May" is an indicator for whether work was reported on the jobcard for May 2012. The periods "BL" and "EL" refer to May 31 - July 4, 2010 and May 28 - July 15, 2012 respectively. "Work spells per member" is the total number of distinct work spells reported on a jobcard divided by the number of members listed on the jobcard. "Avg. implied daily wage" is the total amount earned on a jobcard during the respective period divided by the total number of work days during the respective period. Note that in column 8-9 only jobcards with positive numbers of work days in the respective period were used. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization as control variable. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.5: Comparing surveyed and non-surveyed sampled households - SSP

	% Female	Age	% ST/SC	% Old age	% Widow	% Disabled	% Abhayahastam or Toddy Tappers	Avg. disburs. in 2010	Avg. disburs. in 2011	Avg. disburs. in 2012	Avg. disburs. during BL	Avg. disburs. during EL
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	.043**	.36	.012	.005	-.00037	.0038	-.0085	-3.5	-2.6	10**	-4.5	8.1*
	(.019)	(.6)	(.028)	(.015)	(.014)	(.012)	(.013)	(5.2)	(4.2)	(4.1)	(5.8)	(4.1)
Non-surveyed hhd	-.1	4.6**	.19**	.16***	-.041	-.11***	-.014	-28**	-36***	-36**	-26*	-44***
	(.077)	(1.8)	(.097)	(.059)	(.057)	(.031)	(.04)	(14)	(14)	(15)	(15)	(16)
Treatment X non-surveyed hhd	.14	-.58	-.18*	-.047	.075	.002	-.03	19	-2	8.3	23	-3.1
	(.088)	(2.2)	(.11)	(.077)	(.073)	(.036)	(.044)	(16)	(15)	(17)	(17)	(20)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.01	.00	.06	.01	.00	.00	.01	.00	.01	.01	.00	.01
Control Mean	.58	58	.32	.52	.27	.12	.088	204	258	256	197	253
N. of cases	3317	3317	3317	3317	3317	3317	3317	3317	3317	3317	3317	3317
Level	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd

This table compares sampled SSP households who were surveyed to sampled SSP households who could not be surveyed (excluding confirmed ghost households), using official data. Reasons for missing surveys could be temporary migration, repeated absence on survey dates or refusal to participate in the survey. There were 3171 completed surveys and 152 unsuccessful surveys (another 6 surveys were dropped since no beneficiary in the household could be name-matched to name on the pension card). Outcomes in columns 1-3 are taken from the official database of registered pension beneficiaries. Columns 4-7 compare the proportion of pensioners within a certain eligibility category across groups. Column 7 in particular compares the prevalence of Abhayatstham pension - a pension scheme for women active in self-help groups - and "Toddy Tappers" - paid to the historic trade of palm wine producers. Columns 8 to 10 compare official disbursements averaged across all 12 months of the respective year while columns 11 to 12 compare average disbursements during months May, June and July of the respective year (where "BL" refers to 2010 and "EL" to 2012). All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization as control variable. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.6: Attrition from and entry into sample frames

(a) NREGS

	Treatment	Control	Difference	p-value
	(1)	(2)	(3)	(4)
Attriters from Baseline	.013	.024	-.012	.19
Entrants in Endline	.06	.059	.0018	.74

(b) SSP

	Treatment	Control	Difference	p-value
	(1)	(2)	(3)	(4)
Attriters from Baseline	.097	.097	-.000016	1
Entrants in Endline	.17	.16	.0056	.37

These tables compare the entire NREGS sample frame – i.e., all jobcard holders – and the entire SSP beneficiary frame across treatment (column 1) and control (column 2) mandals. Column 3 reports the difference in treatment and control means, while column 4 reports the p-value on the treatment indicator, both from simple regressions of the outcome with district fixed effects as the only controls. Row 1 presents the proportion of NREGS jobcards and SSP beneficiaries that dropped out of the sample frame between baseline and endline. Row 2 presents the proportion that entered the sample frame between baseline and endline. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors are clustered at the mandal level. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.7: Endline number of jobcards

	Endline # of JCards	
	(1)	(2)
Treatment	8.5	5.6
	(7.5)	(7.3)
District FE	Yes	Yes
Baseline Level	Yes	Yes
Adj R-squared	.97	.97
Control Mean	664	675
N. of cases	2897	874
Level	GP	GP

This table examines whether treatment led to any changes in the number of NREGS jobcards at the GP-level between baseline (2010) and endline (2012). It uses data from the full jobcard data frame in treatment and control mandals. Column 1 includes all GPs within study mandals. Column 2 shows only GPs sampled for our household survey. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors are clustered at the mandal level. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.8: Compositional changes in sample at endline

(a) NREGS

	(1) N. of Members	(2) Hindu	(3) SC	(4) Any Hhd Mem Reads	(5) BPL	(6) Total Consump	(7) Total Income	(8) Own Land
Treatment	.042	-.024	.022	-.031	-.0022	-767	7201*	.054**
	(.11)	(.018)	(.022)	(.027)	(.023)	(4653)	(3839)	(.024)
EL Entrant	-.16	.0094	.03	.065	.067	-10564	-3281	-.052
	(.25)	(.047)	(.077)	(.049)	(.043)	(6874)	(10373)	(.12)
Treat X EL Entrant	.12	-.024	-.082	-.089	-.049	5029	16803	.056
	(.34)	(.058)	(.088)	(.071)	(.057)	(9075)	(14119)	(.14)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.02	.07	.03	.01	.01	.01	.04	.01
Control Mean	4.3	.93	.19	.85	.89	90317	69708	.59
N. of cases	4909	4909	4909	4869	4887	4902	4875	4887
Level	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd

(b) SSP

	(1) N. of Members	(2) Hindu	(3) SC	(4) Any Hhd Mem Reads	(5) BPL	(6) Total Consump	(7) Total Income	(8) Own Land
Treatment	-.0067	.02	-.027	-.047*	-.0014	-1511	4607	.0029
	(.12)	(.021)	(.021)	(.027)	(.018)	(4006)	(4010)	(.032)
EL Entrant	-.035	.0083	-.08**	-.02	.076***	-1883	-1695	.095*
	(.27)	(.041)	(.034)	(.043)	(.026)	(4010)	(4565)	(.056)
Treat X EL Entrants	-.081	.0008	.049	.067	-.051	7688	6028	-.051
	(.3)	(.046)	(.04)	(.055)	(.034)	(5569)	(5664)	(.068)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.04	.05	.02	.02	.01	.02	.07	.02
Control Mean	3.5	.89	.21	.64	.87	63792	52763	.52
N. of cases	3152	3152	3152	3113	3131	3150	3137	3142
Level	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd

These tables show that new entrants to the NREGS and SSP samples are no different across treatment and control groups. “EL Entrant” is an indicator for a household that entered the sample for the endline survey but was not in the baseline sample frame. “Treat X EL Entrant” is the interaction between the treatment indicator and the endline entrant indicator, and the coefficient of interest in these regressions. “N. of Members” is the number of household members. “Hindu” is an indicator for the household belonging to the hindu religion. “SC” is an indicator for the household belonging to a “Scheduled Caste” (historically discriminated-against caste). “Any Hhd Mem Reads” is a proxy for literacy. “BPL” is an indicator for the household being below the poverty line. “Total Consump” is total consumption. “Total Income” is total household income with the top .5% percentile of observations censored.“Own land” is an indicator for whether the household owns any land. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.9: Comparing characteristics of surveyed households at baseline and endline

(a) NREGS

	# hhd members	% non-working age	% children	% female members	% Hindu	% Muslim	% Christian	% SC	% ST	% Hhd head is widow	% members can read	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
EL survey	-3.8***	1***	.28***	1.6***	.017*	-.0073	-.0089	-.0035	.018	-.032	-.17***	-.002
	(.09)	(.04)	(.019)	(.046)	(.0095)	(.0069)	(.0086)	(.019)	(.012)	(.021)	(.013)	(.0042)
Treatment	-.11	-.0063	-.0011	-.0012	-.0053	.0078	-.0056	-.014	.0067	.012	.0025	-.0003
	(.1)	(.017)	(.0087)	(.016)	(.012)	(.008)	(.01)	(.024)	(.025)	(.02)	(.013)	(.0037)
EL survey X treatment	.11	.068	.043*	-.00071	-.0062	.0027	.007	.0043	-.013	-.017	-.0054	-.0031
	(.1)	(.049)	(.026)	(.055)	(.013)	(.0081)	(.011)	(.022)	(.013)	(.025)	(.016)	(.0048)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BL Control Mean	4.8	.35	.098	.51	.9	.039	.052	.26	.12	.15	.61	.014
N. of cases	9555	9555	9555	9555	9555	9555	9555	9532	9532	8104	9512	9555
Level	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd

(b) SSP

	# hhd members	% non-working age	% children	% female members	% Hindu	% Muslim	% Christian	% SC	% ST	% Hhd head is widow	% members can read	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
EL survey	-3.3***	.93***	.18***	1.3***	.0037	.014	-.012	-.032*	.0021	-.072**	-.16***	-.0087
	(.16)	(.063)	(.023)	(.055)	(.012)	(.011)	(.0095)	(.018)	(.014)	(.03)	(.016)	(.0056)
Treatment	-.23	.0037	-.004	.032	.013	.0055	-.017*	-.028	-.022	.054*	-.03*	-.00022
	(.17)	(.023)	(.01)	(.023)	(.014)	(.011)	(.0094)	(.022)	(.025)	(.027)	(.016)	(.0055)
EL survey X treatment	.22	.033	.013	-.049	.011	-.024**	.0077	.03	.0096	-.037	.03	-.0013
	(.17)	(.071)	(.027)	(.066)	(.014)	(.012)	(.011)	(.022)	(.016)	(.036)	(.02)	(.0064)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BL Control Mean	4.3	.39	.065	.57	.89	.049	.055	.23	.14	.43	.48	.015
N. of cases	5931	5931	5931	5931	5931	5931	5931	5922	5922	5327	5886	5931
Level	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd	Hhd

The regressions above compare households surveyed at baseline versus those surveyed at endline on basic socio-economic characteristics. The dependent variable are: the number of members per household, the percentage of members younger than 18 or older than 65, the percentage of members younger than 7, the percentage of households of the respective religion or of the respective category (columns 5 to 9), the percentage of households whose head is a widow and finally the percentage of household members who can read. "EL survey" is a binary variable indicating an observation from the endline survey. "EL survey X treatment" is an interaction effect of being surveyed at endine and being in treatment. "BL control mean" is the mean of the outcome within the control group at baseline. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization as control variable. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

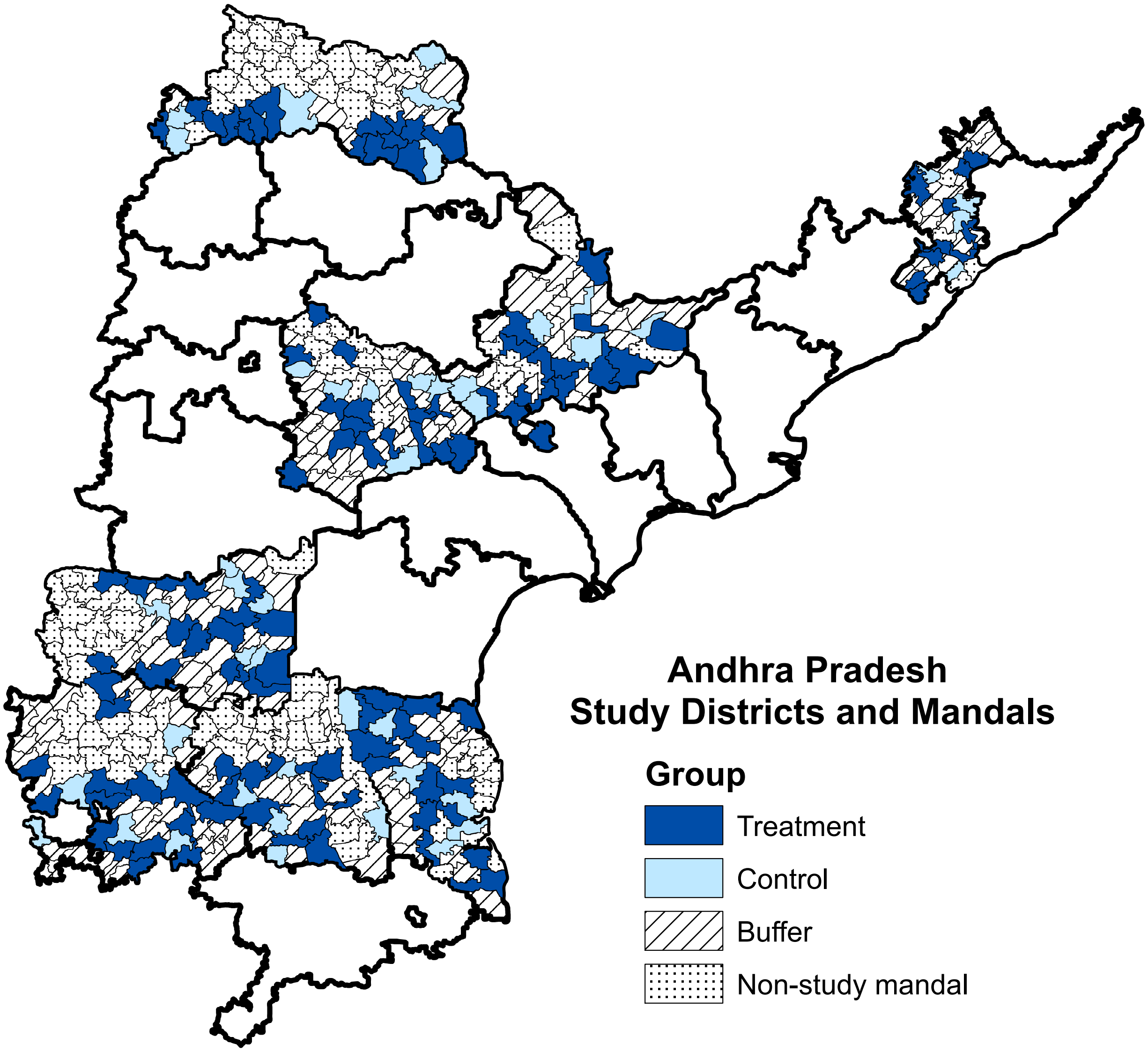


Figure C.1: Study districts with treatment and control mandals

This map shows the 8 study districts - Adilabad, Anantapur, Kadapa, Khammam, Kurnool, Nalgonda, Nellore, and Vizianagaram - and the assignment of mandals (sub-districts) within those districts to one of four study conditions. Mandals were randomly assigned to one of three waves: 112 to wave 1 (treatment), 139 to wave 2, and 45 to wave 3 (control). Wave 2 was created as a buffer to maximize the time between program rollout in treatment and control waves; our study did not collect data on these mandals. A "non-study mandal" is a mandal that did not enter the randomization process because the Smartcards initiative had already started in those mandals (109 out of 405). Randomization was stratified by district and by a principal component of mandal characteristics including population, literacy, Scheduled Caste and Tribe proportion, NREGS jobcards, NREGS peak employment rate, proportion of SSP disability recipients, and proportion of other SSP pension recipients.

Table D.1: Comparison of study districts and other AP districts

	Study Districts	Other AP	Difference	p-value
	(1)	(2)	(3)	(4)
	Numbers based on 2011 census rural totals			
% population rural	.74	.73	.0053	.89
Total rural population	2331398	2779458	-448060*	.067
% male	.5	.5	.0026	.22
% population under age 6	.11	.11	.0047	.35
% ST	.18	.19	-.0094	.69
% SC	.13	.083	.045	.25
Literacy rate	.52	.54	-.022	.37
% working population	.53	.51	.016	.23
% female working population	.24	.22	.015	.34
% main agri. laborers	.23	.22	.0094	.65
% main female agri. laborers	.12	.1	.014	.29
% marginal agri. laborers	.067	.064	.0032	.64
	Numbers based on 2001 census village directory			
# primary schools per village	2.3	2.4	-.14	.68
% villages with medical facility	.56	.67	-.11	.13
% villages with tap water	.53	.56	-.037	.76
% villages with banking facility	.11	.2	-.094	.32
% villages with paved road access	.72	.78	-.06	.39

This table compares characteristics of our 8 study districts and the remaining 13 non-urban (since NREGS is restricted to rural areas) districts in erstwhile Andhra Pradesh, using data from the 2001 and 2011 censuses. Column 3 reports the difference in means, while column 4 reports the p-value on a study district indicator, both from simple regressions of the outcome with no controls. "SC" ("ST") refers to Scheduled Castes (Tribes), historically discriminated-against sections of the population now accorded special status and affirmative action benefits under the Indian Constitution. "Working" is defined as participating in any economically productive activity with or without compensation, wages or profit. "Main" workers are defined as those who engaged in any economically productive work for more than 183 days in a year. "Marginal" workers are those for whom the period they engaged in economically productive work does not exceed 182 days. Note that the difference in "main" and "marginal" workers only stems for different periods of work. An "agricultural laborer" is a person who works for compensation on another person's land (compensation can be paid in money, kind or share). The definitions are from the official census documentation. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table D.2: Comparison of study mandals and dropped mandals

	Mandals considered for randomization	Mandals not considered	Difference	p-value
	(1)	(2)	(3)	(4)
	Numbers based on 2011 census rural totals			
% population rural	.89	.89	-.015	.58
Total rural population	46380	45582	-1580	.27
% male	.5	.5	.00039	.64
% population under age 6	.11	.12	-.005***	.00028
% SC	.19	.18	.014**	.031
% ST	.12	.14	-.026*	.095
Literacy rate	.53	.51	.01*	.061
% working population	.53	.53	-.0011	.8
% female working population	.24	.24	-.0039	.28
% main agri. laborers	.23	.21	.0019	.77
% female main agri. laborers	.12	.11	-.0019	.59
% marginal agri. laborers	.069	.066	.0043	.24
	Numbers based on 2001 census village directory			
# primary schools per village	2.9	2.6	.31*	.052
% village with medical facility	.68	.62	.044*	.082
% villages with tap water	.6	.62	-.052*	.081
% villages with banking facility	.13	.12	.0015	.87
% villages with paved road access	.78	.76	.018	.49
Avg. village size in acres	3404	3040	298	.12

This table compares characteristics of the 296 mandals that entered the randomization (and were randomized into treatment, control and buffer) to the 108 rural mandals in which the Smartcard initiative had begun prior to our intervention, using data from the 2001 and 2011 censuses. One mandal (Kadapa mandal in Kadapa district, i.e. the district's capital) is excluded since it is fully urban (hence has no NREGS). Column 3 reports the difference in means, while column 4 reports the p-value on an indicator for a mandal that was randomized, both from simple regressions of the outcome with district fixed effects. "SC" ("ST") refers to Scheduled Castes (Tribes), historically discriminated-against sections of the population now accorded special status and affirmative action benefits under the Indian Constitution. "Working" is defined as the participating in any economically productive activity with or without compensation, wages or profit. "Main" workers are defined as those who engaged in any economically productive work for more than 183 days in a year. "Marginal" workers are those for whom the period they engaged in economically productive work does not exceed 182 days. Note that the difference in "main" and "marginal" workers only stems for different periods of work. An "agricultural laborer" is a person who works for compensation on another person's land (compensation can be paid in money, kind or share). The definitions are from the official census documentation. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table D.3: Baseline covariates and program implementation at mandal level

	NREGS				SSP			
	Mandal converted		Intensity		Mandal converted		Intensity	
	(1) Binary	(2) Multiple	(3) Binary	(4) Multiple	(5) Binary	(6) Multiple	(7) Binary	(8) Multiple
Time to Collect (1 hr)	-.093**	-.1**	-.0085	-.013	.061	.072	.012	.022
	(.043)	(.044)	(.026)	(.027)	(.058)	(.057)	(.047)	(.047)
Official Amount (Rs. 100)	.019	.011	.016	.012	.22***	.26***	.14**	.17**
	(.028)	(.038)	(.017)	(.023)	(.084)	(.088)	(.069)	(.072)
Survey Amount (Rs. 100)	.023	.015	.016	.0042	-.021	-.092*	-.021	-.071*
	(.033)	(.044)	(.02)	(.027)	(.045)	(.047)	(.036)	(.039)
SC Proportion	-.032	.009	-.085	-.058	-.072	.083	-.027	.072
	(.22)	(.22)	(.13)	(.14)	(.2)	(.19)	(.16)	(.16)
BPL Proportion	.96	1.2	.71	.64	1.1*	1.5**	.94*	1.2**
	(1.2)	(1.2)	(.69)	(.72)	(.68)	(.69)	(.55)	(.57)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared		.14		.57		.22		.44
N. of cases	112	112	112	112	112	112	112	112

This tables analyzes the effects of baseline covariates on endline program implementation in treatment areas. The columns labeled "binary" show coefficients from regressions with each covariate regressed separately. Hence every cell in columns 1, 3, 5 and 7 shows the result from a separate regression. In contrast, the columns labeled "multiple" run one single regression with all covariates. A "converted mandal" is a mandal in which at least one GP has converted to Smartcard based payments. As of July 2012, 92 of 112 (82%) mandals were converted for NREGS payments, while 100 of 112 (93%) were converted for SSP payments. "Treatment intensity" is the mandal mean of the proportion of transactions done with carded beneficiaries in carded GPs. All regressors are mandal-level averages. "Time to collect (1 hr)" is the average time taken to collect a payment (in hours), including the time spent on unsuccessful trips to payment sites. "Official amount (Rs. 100)" refers to amounts paid as listed in official records. "Survey amount (Rs. 100)" refers to payments received as reported by beneficiaries. "SC proportion" is GP proportion of Scheduled Caste households. "BPL proportion" is GP proportion of households below the poverty line. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table D.4: Baseline covariates and program implementation at GP level

	NREGS				SSP			
	Carded GP		Intensity		Carded GP		Intensity	
	(1) Binary	(2) Multiple	(3) Binary	(4) Multiple	(5) Binary	(6) Multiple	(7) Binary	(8) Multiple
Time to Collect (1 hr)	-.018	-.022	-.0007	-.0035	-.021	-.022	-.016	-.016
	(.015)	(.014)	(.01)	(.01)	(.03)	(.029)	(.024)	(.024)
Official Amount (Rs. 100)	-.0093	-.0032	.0053	.0065	.056*	.095***	.035	.059**
	(.014)	(.017)	(.0094)	(.011)	(.03)	(.035)	(.022)	(.027)
Survey Amount (Rs. 100)	-.011	-.011	.0034	-.0034	-.0092	-.023*	-.0045	-.013
	(.014)	(.017)	(.011)	(.012)	(.014)	(.013)	(.0093)	(.0091)
SC Proportion	-.067	-.037	-.061	-.041	-.069	-.046	-.035	-.023
	(.078)	(.077)	(.054)	(.054)	(.059)	(.059)	(.045)	(.047)
BPL Proportion	.81**	.91**	.51*	.55	.38**	.43**	.25**	.27**
	(.37)	(.45)	(.3)	(.34)	(.17)	(.17)	(.12)	(.11)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared		.26		.45		.35		.43
N. of cases	627	625	627	625	586	584	586	584

This tables analyzes the effects of baseline covariates on endline program implementation at the GP-level. The columns labeled "binary" show coefficients from regressions with each covariate regressed separately. Hence every cell in columns 1, 3, 5 and 7 shows the result from a separate regression. In contrast, the columns labeled "multiple" run one single regression with all covariates. "Carded GP" is a gram panchayat that has converted to Smartcard based payment, which usually happens once 40% of beneficiaries have been issued a card. "Treatment intensity" is the proportion of transactions done with carded beneficiaries in carded GPs. All regressors are GP-level averages. "Time to collect (1 hr)" is the average time taken to collect a payment (in hours), including the time spent on unsuccessful trips to payment sites. "Official amount (Rs. 100)" refers to amounts paid as listed in official records. "Survey amount (Rs. 100)" refers to payments received as reported by beneficiaries. "SC proportion" is GP proportion of Scheduled Caste households. "BPL proportion" is GP proportion of households below the poverty line. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table D.5: Correlates of owning a Smartcard

	NREGS		SSP	
	(1) Binary	(2) Multiple	(3) Binary	(4) Multiple
Income (Rs. 10,000)	-.0043**	-.0039**	.0015	.0010
	(.0020)	(.0020)	(.0020)	(.0019)
Consumption (Rs. 10,000)	-.0014	-.00088	.0013	.00096
	(.0012)	(.0012)	(.0021)	(.0021)
Official amount (Rs. 100)	.0041***	.0043***	.00024	.000075
	(.00083)	(.00082)	(.0028)	(.0028)
SC	.070*	.074**	.017	.018
	(.037)	(.036)	(.029)	(.029)
Female	.039**	.042**	-.021	-.022
	(.017)	(.017)	(.024)	(.024)
District FE	Yes	Yes	Yes	Yes
Adj R-squared		.27		.21
Dep Var Mean	.47	.47	.73	.73
N. of cases	5200	5164	1872	1862
Level	Indiv.	Indiv.	Indiv.	Indiv.

This tables analyzes how endline covariates predict which individuals have or use a Smartcard within villages that have moved to Smartcard based payments ("Carded GPs"). The columns labeled "binary" show coefficients from regressions with each covariate regressed separately. Hence every cell in columns 1 and 3 shows the result from a separate regression. In contrast, the columns labeled "multiple" run one single regression with all covariates. "Income (Rs. 10,000)" is household income with units as 1 = Rs. 10,000. "Consumption (Rs. 10,000)" is household consumption. "Land value (Rs. 10,000)" is household land value. "NREGS amount (Rs. 1,000)" is household NREGS income during the study period. "SC" is a dummy for whether household is Scheduled Caste. "Total Income" is total household income with the top .5% percentile of observations censored. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table E.1: Scaled NREGS earnings and leakage regressions

	Official			Survey		Leakage		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	9.7	4.6	-7.3	33	32	-23	-27	-33
	(25)	(24)	(23)	(21)	(20)	(21)	(20)	(21)
BL GP Mean		.16***			.1***		.14***	
		(.025)			(.038)		(.034)	
BL jobcard payment			.24***					.16***
			(.048)					(.053)
BL jobcard payment > 0			185***					86**
			(32)					(34)
BL GP Mean survey payment								-.1**
								(.047)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.03	.05	.19	.06	.07	.06	.07	.12
Control Mean	260	260	260	180	180	80	80	80
N. of cases	5143	5107	5107	5143	5107	5143	5107	5107

This table reports regressions of program benefits (in Rupees) as reported in official or survey records. Regressions include all sampled NREGS households who were a) found by survey team to match official records or b) listed in official records but confirmed as "ghosts". "Ghosts" refer to households or beneficiaries within households that were confirmed not to exist, or who had permanently migrated before the study period started on May 28, 2012. Each outcome observation refers to household-level average weekly amounts for NREGS work done during the study period (May 28 to July 15 2012). "Official" refers to amounts paid as listed in official muster records, *scaled by the average number of jobcards per household in the district*. "Survey" refers to payments received as reported by beneficiaries. "Leakage" is the difference between these two amounts. "BL GP Mean" is the GP average of household-level weekly amounts for NREGS work done during the baseline study period (May 31 to July 4 2010). The "BL GP Mean" for "Official" was scaled the same way the dependent variable was. "BL jobcard payment" was the official weekly disbursement on the sampled jobcard during the baseline study period; "BL jobcard payment > 0" is an indicator for this payment being positive. Note that the regressions no longer include only individuals listed on sampled jobcards but rather household-level average weekly amounts using data from all working household members. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table E.2: Other leakage robustness results

	# of workers found in audit		Paid yet for a given period			
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	13	11	.029	.032		
	(12)	(10)	(.033)	(.035)		
Treatment X First 4 weeks					.04	.044
					(.034)	(.036)
Treatment X Last 3 weeks					-.035	-.034
					(.059)	(.063)
District FE			Yes	Yes	Yes	Yes
Week FE	No	Yes	Yes	Yes	Yes	Yes
BL GP Mean	No	No	No	Yes	No	Yes
p-value: first 4 weeks = last 3 weeks					.19	.21
Adj R-squared	.097	.14	.085	.085	.087	.087
Control Mean	28	28	.9	.9	.9	.9
N. of cases	508	508	11854	11174	11854	11174
Level	GP	GP	Indiv-Week	Indiv-Week	Indiv-Week	Indiv-Week

In columns 1 and 2, units represent estimated number of NREGS workers on a given day, found in an independent audit of NREGS worksites in GPs. In columns 3-6, the outcome is an indicator for whether an NREGS respondent had received payment for a given week's work at the time of the survey, weighted by the official payment amount. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table E.3: Summary statistics and treatment effects from the list experiment

(a) Summary statistics

	Treatment	Control	Difference	p-value	N
	(1)	(2)	(3)	(4)	(5)
Version 1	2.13	2.19	-.06	.54	1601
Version 2	2.21	2.34	-.13	.25	1616
Version 3	2.34	2.46	-.11	.32	1572

(b) Regression-adjusted treatment effects

	All versions	Versions 1 & 2
	(1)	(2)
Treatment	-.057	-.054
	(.11)	(.11)
Version 2	.15	.16
	(.11)	(.11)
Version 3	.27***	
	(.1)	
Version 2 X treatment	-.089	-.095
	(.13)	(.13)
Version 3 X treatment	-.056	
	(.12)	
District FE	Yes	Yes
p-val: Version 2 X Tr. = 0	.49	.46
p-val: Version 3 X Tr. = 0	.63	
Adj R-squared	.14	.12
Version 1 control mean	2.19	2.19
N. of cases	4789	3217
Level	Hhd	Hhd

This table presents results of the "list experiment" conducted within the survey to determine whether officials asked households to lie about their NREGS participation and payments. Columns 1-2 in panel a) show means for the treatment and control group respectively. Column 3 shows the regression-adjusted difference from a regression with the district FE and a vector of mandal characteristics used to stratify randomization as covariates. The p-value in column 4 is from a two-sided test in which the null hypothesis is that the difference in column 3 is equal to 0. "Version 1" denotes respondents who were asked how many of 5 statements they would agree with. "Version 2" denotes those were presented with the same 5 statements as Version 1 as well as an additional sensitive statement: "Members of this household have been asked by officials to lie about the amount of work they did on NREGS". "Version 3" denotes those were presented with the same 5 statements as Version 1 "Members of this household have been given the chance to meet with the CM of AP to discuss problems with NREGS?"). Panel b) reports regression-adjusted treatment effects. Column 1 compares version 1 to version 2 and version 3 while column 2 only compares version 1 and 2. "Version 2 X treatment" and "Version 3 X treatment" are interaction terms of having faced the respective survey version and being in the treatment group. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at the mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table E.4: Analyzing potential recall bias in leakage results

	Survey		Leakage	
	(1)	(2)	(3)	(4)
Treatment X surveyed in week 1	54	38	35	55
	(86)	(78)	(88)	(83)
Treatment X surveyed in week 2	77*	88*	-92**	-97**
	(44)	(45)	(44)	(44)
Treatment X surveyed in week 3	35	33	-42	-52
	(41)	(40)	(33)	(33)
Treatment X surveyed in week 4	35	35	-37	-42
	(44)	(45)	(47)	(44)
Treatment X surveyed in week 5	70**	77**	-37	-48
	(35)	(38)	(31)	(31)
Treatment X surveyed in week 6	46	35	-34	-37
	(37)	(36)	(35)	(36)
Treatment X surveyed in week 7	-43	-29	42	38
	(69)	(66)	(54)	(54)
Treatment X surveyed in week 8	19	11	12	24
	(24)	(30)	(24)	(20)
Treatment X surveyed in week 9	106***	105***	-28	-23
	(28)	(27)	(25)	(25)
Treatment X surveyed in week 10	-52	-58	-28	-29
	(48)	(42)	(42)	(43)
BL GP Mean		.13***		.12***
		(.041)		(.044)
District FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Adj R-squared	.07	.07	.05	.05
Control Mean	165	165	-21	-21
N. of cases	4803	4769	4803	4769
Level	Hhd	Hhd	Hhd	Hhd

The regressions include all sampled households who were a) found by survey team to match official record or b) listed in official records but confirmed as "ghosts". "Ghosts" refer to households or beneficiaries within households that were confirmed not to exist, or who had permanently migrated before the study period started on May 28, 2012. In panel (a), each outcome observation refers to household-level average weekly amounts for NREGS work done during the study period (May 28 to July 15 2012). "Treatment X surveyed in week x" is an interaction term of treatment and the household survey taking place in week x. Note that the household surveys took place in August, September and the early weeks of October 2012. Note all regressions include week fixed effects. The number of observations is different compared to Table 3a because for some surveys the survey date information was corrupted or missing. "Survey" refers to payments received as reported by beneficiaries. "Leakage" is the difference between the survey amount and the offical amount disbursed. "BL GP Mean" is the GP average of household-level weekly amounts for NREGS work done during the baseline study period (May 31 to July 4 2010). All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table E.5: Geographical spillovers on payment collection time

(a) NREGS

	Time to collect (Min)				
	(1)	(2)	(3)	(4)	(5)
Treatment	-23**	-23***	-23***	-23***	-22**
	(8.9)	(8.4)	(8.5)	(8.6)	(8.6)
Fraction GPs treated within 10km	-3.6				
	(7.6)				
Fraction GPs treated within 15km		-6.2			
		(9.8)			
Fraction GPs treated within 20km			-11		
			(13)		
Fraction GPs treated within 25km				-15	
				(15)	
Fraction GPs treated within 35km					-13
					(24)
BL GP Mean	.067*	.074*	.072*	.076*	.077*
	(.04)	(.04)	(.04)	(.041)	(.041)
District FE	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.07	.08	.08	.08	.08
Control Mean	112	112	112	112	112
N. of cases	9081	9781	9978	10045	10074
Level	Indiv.	Indiv.	Indiv.	Indiv.	Indiv.

(b) SSP

	Time to Collect (Min)				
	(1)	(2)	(3)	(4)	(5)
Treatment	-5	-5.6	-5	-4.9	-4.3
	(5.7)	(5.3)	(5.3)	(5.3)	(5.4)
Fraction GPs treated within 10km	-12*				
	(6.6)				
Fraction GPs treated within 15km		-12			
		(7.8)			
Fraction GPs treated within 20km			-11		
			(9.5)		
Fraction GPs treated within 25km				-12	
				(12)	
Fraction GPs treated within 35km					-15
					(16)
BL GP Mean	.28***	.23***	.22***	.22***	.22***
	(.08)	(.074)	(.073)	(.072)	(.072)
District FE	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.13	.11	.11	.11	.11
Control Mean	77	77	77	77	77
N. of cases	3209	3462	3524	3545	3551
Level	Indiv.	Indiv.	Indiv.	Indiv.	Indiv.

The dependent variable in all columns in both panels is the average time taken to collect a payment (in minutes), including the time spent on unsuccessful trips to payment sites, with observations at the beneficiary level. The "fraction GPs treated within x" is the ratio of the number of GPs in treatment mandals within radius x km over the total GPs within wave 1, 2 or 3 mandals. Note that wave 2 mandals are included in the denominator, and that same-mandal GPs are excluded in both the denominator and numerator. "BL GP Mean" is the GP-average of the dependent variable at baseline. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table E.6: Geographical spillovers on timeliness of payments

(a) Payment lag (Days)

	Ave Payment Delay				
	(1)	(2)	(3)	(4)	(5)
Treatment	-12***	-10***	-9.7***	-10***	-10***
	(3.8)	(3.8)	(3.5)	(3.5)	(3.3)
Fraction GPs treated within 10km	-1.3				
	(4.7)				
Fraction GPs treated within 15km		-.34			
		(5.1)			
Fraction GPs treated within 20km			2		
			(5.6)		
Fraction GPs treated within 25km				-.81	
				(6.1)	
Fraction GPs treated within 35km					-.17
					(8.7)
BL GP Mean	.023	.012	.0075	.0096	.013
	(.089)	(.087)	(.081)	(.08)	(.079)
District FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.34	.33	.33	.33	.33
Control Mean	34	34	34	34	34
N. of cases	6446	6897	7087	7169	7201
Level	Indiv-Week	Indiv-Week	Indiv-Week	Indiv-Week	Indiv-Week

(b) Absolute deviation from Mandal median (Days)

	Payment Lag Deviation				
	(1)	(2)	(3)	(4)	(5)
Treatment	-5.4***	-4.5***	-4.5***	-4.7***	-4.6***
	(1.6)	(1.7)	(1.6)	(1.6)	(1.5)
Fraction GPs treated within 10km	.19				
	(1.5)				
Fraction GPs treated within 15km		.67			
		(2.2)			
Fraction GPs treated within 20km			.72		
			(2.6)		
Fraction GPs treated within 25km				-.22	
				(2.9)	
Fraction GPs treated within 35km					1
					(4.3)
BL GP Mean	.059	.034	.036	.037	.041
	(.054)	(.051)	(.051)	(.051)	(.052)
District FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.18	.17	.17	.17	.17
Control Mean	12	12	12	12	12
N. of cases	6446	6897	7087	7169	7201
Level	Indiv-Week	Indiv-Week	Indiv-Week	Indiv-Week	Indiv-Week

The dependent variable in panel a) is the average lag (in days) between work done and payment received on NREGS. The outcome in panel b) is the absolute deviation from the week-specific median mandal-level lag. Since the data for columns 5-8 are at the individual-week level, we include week fixed effects to absorb variation over the study period. The "fraction GPs treated within x" is the ratio of the number of GPs in treatment mandals within radius x km over the total GPs within wave 1, 2 or 3 mandals. Note that wave 2 mandals are included in the denominator, and that same-mandal GPs are excluded in both the denominator and numerator. "BL GP Mean" is the GP-average of the dependent variable at baseline. Each regression includes the stratification principal component as a control variable. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table E.7: Geographical spillover effects on leakage

(a) NREGS

	Leakage				
	(1)	(2)	(3)	(4)	(5)
Treatment	-27*	-25*	-27**	-28**	-27**
	(14)	(13)	(13)	(13)	(13)
Fraction GPs treated within 10km	1.7				
	(15)				
Fraction GPs treated within 15km		-5.2			
		(22)			
Fraction GPs treated within 20km			-15		
			(27)		
Fraction GPs treated within 25km				-26	
				(29)	
Fraction GPs treated within 35km					-22
					(35)
BL GP Mean	.1**	.087**	.1**	.098**	.098**
	(.044)	(.039)	(.041)	(.039)	(.039)
District FE	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.04	.04	.04	.04	.04
Control Mean	-20	-20	-20	-20	-20
N. of cases	4564	4932	5032	5062	5073
Level	Hhd	Hhd	Hhd	Hhd	Hhd

(b) SSP

	Leakage				
	(1)	(2)	(3)	(4)	(5)
Treatment	-5.3	-5.3	-5.3	-5.5	-5.7
	(4.3)	(3.8)	(3.8)	(3.9)	(4)
Fraction GPs treated within 10km	2.2				
	(4.5)				
Fraction GPs treated within 15km		3			
		(5.2)			
Fraction GPs treated within 20km			1.5		
			(6.1)		
Fraction GPs treated within 25km				.032	
				(7.5)	
Fraction GPs treated within 35km					-4.2
					(11)
BL GP Mean	-.049	-.044	-.044	-.044	-.044
	(.031)	(.032)	(.032)	(.032)	(.032)
District FE	Yes	Yes	Yes	Yes	Yes
Adj R-squared	.01	.01	.01	.01	.01
Control Mean	15	15	15	15	15
N. of cases	2806	3034	3086	3106	3112
Level	Hhd	Hhd	Hhd	Hhd	Hhd

The regressions in both panels include all sampled households (NREGS)/beneficiaries (SSP) who were a) found by survey team to match official record or b) listed in official records but confirmed as "ghosts". "Ghosts" refer to households or beneficiaries within households that were confirmed not to exist, or who had permanently migrated before the study period started on May 28, 2012. In panel (a), each outcome observation refers to household-level average weekly amounts for NREGS work done during the study period (May 28 to July 15 2012). In panel (b), each outcome observation refers to the average SSP monthly amount for the period May, June, and July 2012. "Leakage" is the difference between the amount disbursed as indicated by ofifcial records and the amount a household reported as received in the survey. The "fraction GPs treated within x" is the ratio of the number of GPs in treatment mandals within radius x km over the total GPs within wave 1, 2 or 3 mandals. Note that wave 2 mandals are included in the denominator, and that same-mandal GPs are excluded in both the denominator and numerator. In panel a) "BL GP Mean" is the GP average of household-level weekly, amounts for work done during the baseline study period (May 31 to July 4 2010). In panel b) "BL GP Mean" is the GP average monthly amounts based on official disbursements from baseline period of May, June, and July 2010 and baseline survey information on pension proceedings (see Table 3b. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table E.8: Impacts on NREGS projects and budget categories

	Works		Person-days		Wages paid		Material expend.		Contingent expend.		Total expend.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	-4.1	-.55	-67	-114	-11299	-14924	7495	7465	-81	-154	-3885	-7534
	(7.4)	(6.9)	(306)	(221)	(32456)	(24308)	(7744)	(7589)	(219)	(197)	(36238)	(28525)
Baseline		.84***		.47***		.54***		.061***		.11***		.44***
		(.078)		(.04)		(.048)		(.016)		(.013)		(.037)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R-squared	0.10	0.26	0.21	0.43	0.20	0.42	0.02	0.02	0.09	0.14	0.17	0.36
Control Mean	88	88	3940	3940	420493	420493	34459	34459	5030	5030	459981	459981
N. of cases	8736	8730	8736	8730	8736	8730	8736	8730	8736	8730	8736	8730
Level	GP-Month	GP-Month	GP-Month	GP-Month	GP-Month	GP-Month	GP-Month	GP-Month	GP-Month	GP-Month	GP-Month	GP-Month

This table analyzes impacts on NREGS projects and various budget categories at the GP level for the months of May, June and July 2012. “Works” is the number of projects that official data sources placed in a GP in a month. A “personday” means one day of work one person. “Wages paid” is the total amount of wage outlays on NREGS for a GP within the given month. “Material expen.” and “contingent expen.” are monthly totals for NREGS work in a given GP. “Total expend.” is the sum of wage outlays, material and contingent expenditure within a GP in a given month. “Baseline” is the lagged dependent variable as observed during the respective months in 2010. All regressions include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table E.9: Hawthorne effects

	Audit	Official				Survey
	(1)	(2)	(3)	(4)	(5)	(6)
Survey in GP	-3.7	10	-4.8			
	(8)	(34)	(33)			
Audit in GP		7.5	-13	6.8	-12	116
		(31)	(28)	(42)	(37)	(106)
Audit in Week		-52	-71	-26	-34	40
		(51)	(52)	(39)	(39)	(84)
Recon in Week		12	-.8	49	44	45
		(69)	(68)	(53)	(52)	(90)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
BL GP Value	No	No	Yes	No	Yes	Yes
GP Size FE	No	Yes	Yes	No	No	No
Adj R-squared	.18					
Control Mean	49	758	758	756	756	1175
Level	Week	Week	Week	Week	Week	Week
Sample	Audit	All	All	Survey & Audit	Survey & Audit	Survey
N. of cases	676	52311	52311	7679	7679	6111

This table analyzes possible Hawthorne effects from various data collection activities. Each cell represents a separate regression of the effect on the data source (column) from the survey type (row). Units are number of days worked in a GP per week. "Survey in GP" is an indicator for whether a GP was part of the household survey. "Audit in GP" is a binary variable equal to 1 if the GP was sampled for work site audits while "Audit in week" indicates that the work site audit happened in a specific week. "Recon in week" is an indicator for whether an enumerator went to map the worksites in a specific week. "All", "Audit", and "Survey" indicate that the data came from all mandals in the study district, the GPs sampled for the work site audits or from the GPs sampled for the household survey respectively. The regressions in column 1 as well as columns 4 to 6 include the first principal component of a vector of mandal characteristics used to stratify randomization. Standard errors clustered at mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Household (surveyed)	
Name	Payment
Karthik	30
Paul	20
Sandip	40

Jobcard (sampled)	
Name	Payment
Karthik	30
Paul	0

Jobcard *not* sampled)	
Name	Payment
Paul	35
Sandip	50

Figure E.1: Illustrating multiple jobcards

Table F.1: Heterogeneity in impacts by baseline characteristics

(a) NREGS

	Time to Collect	Payment Lag	Official Payments	Survey Payments
	(1)	(2)	(3)	(4)
BL GP Mean	.024	.19	.012	.048
	(.08)	(.25)	(.042)	(.074)
Consumption (Rs. 1,000)	-.085	-.0045	-.0024	-.044
	(.16)	(.025)	(.2)	(.26)
GP Disbursement, NREGS (Rs. 1,000)	.015*	.00014	.014	.0044
	(.0078)	(.0013)	(.01)	(.016)
SC Proportion	.31	25*	3.6	13
	(48)	(13)	(49)	(51)
BPL Proportion	-61	-24	-64	-171
	(127)	(22)	(111)	(114)
District FE	Yes	Yes	Yes	Yes
Week FE	No	Yes	No	No
Control Mean	112	34	127	146
Level	Indiv.	Indiv-Week	Hhd	Hhd
N. of cases	10143	12334	4999	4999

(b) SSP

	Time to Collect	Official Payments	Survey Payments
	(1)	(2)	(3)
BL GP Mean	.21**	-.012	.029
	(.1)	(.086)	(.095)
Consumption (Rs. 1,000)	-.26**	-.014	-.096
	(.11)	(.099)	(.23)
GP Disbursement, SSP (Rs. 1000)	-.083	.059	.11
	(.094)	(.072)	(.12)
SC Proportion	18	-29	-20
	(16)	(23)	(37)
BPL Proportion	-66*	126**	95
	(35)	(53)	(83)
District FE	Yes	Yes	Yes
Control Mean	77	257	298
Level	Indiv.	Indiv.	Indiv.
N. of cases	3573	2981	2981

This table shows heterogeneous effects on major endline outcomes from GP-level baseline characteristics. Each cell shows the coefficient on the baseline characteristic interacted with the treatment indicator in separate regressions. “BL GP Mean” is the baseline GP-level mean for the outcome variable. “Consumption (Rs. 1,000)” is annualized consumption. “GP Disbursement (Rs. 1000)” is total NREGS/SSP payment amounts for the period Jan 1, 2010 to July 22, 2010. “SC Proportion” is the proportion of NREGS workspells performed by schedule caste workers/SSP beneficiaries in the period from Jan 1, 2010 to July 22, 2010. “BPL Proportion” is the proportion of households with a BPL card in the baseline survey. Standard errors clustered at the mandal level in parentheses. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

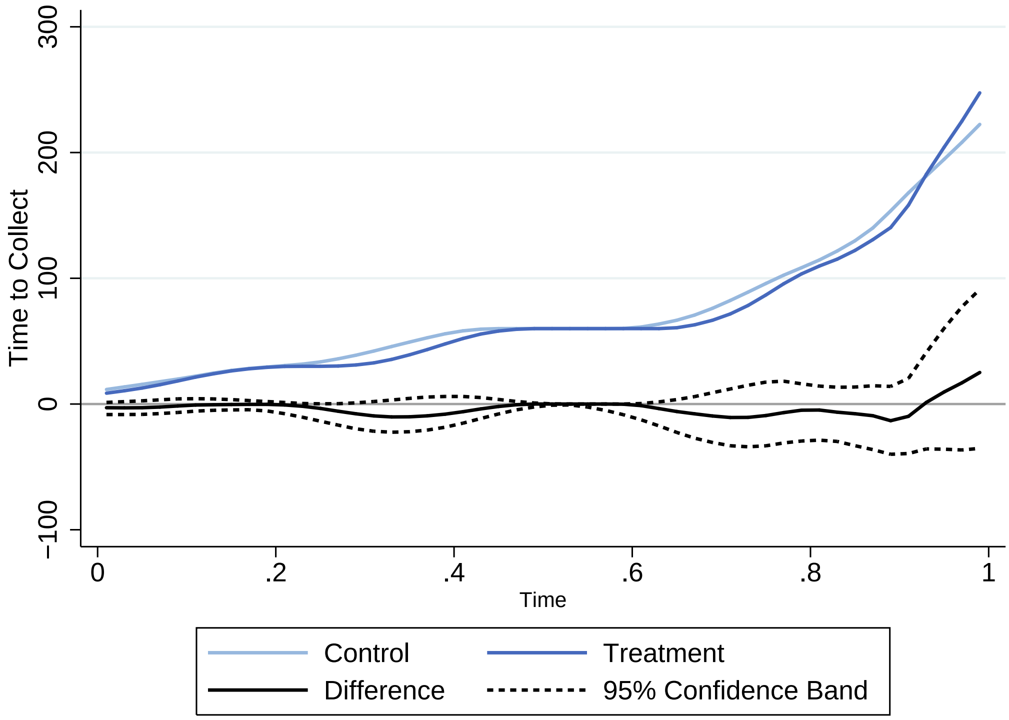


Figure F.1: Quantile treatment effect on payment collection time - SSP

This figure shows non-parametric treatment effects. “Time to collect: SSP” is the average time taken to collect a payment, including the time spent on unsuccessful trips to payment sites. All lines are fit by a kernel-weighted local polynomial smoothing function with Epanechnikov kernel and probability weights, with bootstrapped standard errors. The dependent variable is the vector of residuals from a linear regression of the respective outcome with the first principal component of a vector of mandal characteristics used to stratify randomization and district fixed effects as regressors.