NBER WORKING PAPER SERIES

IDENTITY VERIFICATION STANDARDS IN WELFARE PROGRAMS: EXPERIMENTAL EVIDENCE FROM INDIA

Karthik Muralidharan
Paul Niehaus
Sandip Sukhtankar

Working Paper 26744
http://www.nber.org/papers/w26744

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2020

We thank Prashant Bharadwaj, Lucie Gadenne, Siddharth George, Aprajit Mahajan, Ted Miguel, and participants in various seminars for comments and suggestions. This paper would not have been possible without the continuous efforts and inputs of the J-PAL/UCSD project team including Avantika Prabhakar, Burak Eskici, Frances Lu, Jianan Yang, Kartik Srivastava, Krutika Ravishankar, Mayank Sharma, Sabareesh Ramachandran, Simoni Jain, Soala Ekine, Xinyi Liu, and Vaibhav Rathi. Finally, we thank the Bill and Melinda Gates Foundation (especially Dan Radcliffe and Seth Garz) for the financial support that made this study possible. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Identity Verification Standards in Welfare Programs: Experimental Evidence from India
Karthik Muralidharan, Paul Niehaus, and Sandip Sukhtankar
NBER Working Paper No. 26744
February 2020
JEL No. D73,H53,O30,Q18

ABSTRACT

How should recipients of publicly-provided goods and services prove their identity in order to access these benefits? The core design challenge is managing the tradeoff between Type-II errors of inclusion (including corruption) against Type-I errors of exclusion whereby legitimate beneficiaries are denied benefits. We use a large-scale experiment randomized across 15 million beneficiaries to evaluate the effects of more stringent ID requirements based on biometric authentication on the delivery of India's largest social protection program (subsidized food) in the state of Jharkhand. By itself, requiring biometric authentication to transact did not reduce leakage, slightly increased transaction costs for the average beneficiary, and reduced benefits received by the subset of beneficiaries who had not previously registered an ID by 10%. Subsequent reforms that made use of authenticated transaction data to determine allocations to the program coincided with large reductions in leakage, but also significant reductions in benefits received. Our results highlight that attempts to reduce corruption in welfare programs can also generate non-trivial costs in terms of exclusion and inconvenience to genuine beneficiaries.

Karthik Muralidharan
Department of Economics, 0508
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508
and NBER
kamurali@ucsd.edu

Paul Niehaus
Department of Economics
University of California, San Diego
9500 Gilman Drive #0508
La Jolla, CA 92093
and NBER
pniehaus@ucsd.edu

Sandip Sukhtankar
Department of Economics
University of Virginia
Charlottesville, VA 22904
srs8yk@virginia.edu

A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/1620

1 Introduction

How should recipients of publicly provided goods and services prove their identity in order to access these benefits? From accessing welfare benefits to obtaining a driver's license to casting a vote, how stringent ID requirements should be is a perennially controversial question around the world. The core design issue is how to manage the tradeoff between Type-II errors of *inclusion* whereby benefits are paid to non-eligible or "ghost" recipients against Type-I errors of *exclusion* whereby legitimate beneficiaries are denied benefits to which they are entitled. While there is a large literature on how to *target* people to be put on program beneficiary lists (Alatas et al., 2012, 2016; Niehaus et al., 2013), there is much less evidence on the distinct question of how onerous it should be for citizens to prove their *identity* at the point of receiving benefits.

This question is particularly salient in developing countries. Historically, states have invested in the ability to better identify their citizens as they develop (Scott, 1998). During the past two decades in particular "the number of national identification and similar programs has grown exponentially... to the point where almost all developing countries have at least one such program" (Gelb and Metz, 2018). Around two-thirds of these use biometric technology, reflecting the view that this provides more reliable authentication than alternatives, particularly in settings with low levels of literacy and numeracy.[1] A leading case is India, where the government has now issued unique identification ("Aadhaar") numbers linked to biometric records to over 1.24 billion people and is gradually integrating Aadhaar-based authentication into a range of applications. The extent to which authentication should be mandated to receive welfare benefits has been a highly controversial issue, contested all the way to the Supreme Court. Proponents argue that this is necessary to prevent fraud, while critics argue that the requirement denies people their legal entitlements and in doing so "undermines the right to life" (Khera, 2017). In a September 2018 ruling, India's Supreme Court allowed the government to mandate the use of Aadhaar for accessing social programs, making it all the more urgent to understand how doing so affects errors of inclusion and exclusion.

This paper reports results from the first (to our knowledge) experimental evaluation of introducing Aadhaar as a requirement to collect welfare benefits. Specifically, we examine how this introduction shifted the tradeoff between errors of inclusion and exclusion in the Public Distribution System (PDS), India's largest welfare program, accounting for roughly 1% of GDP. The PDS is the primary policy instrument for providing food security to the poor in India, which has the largest number of malnourished people in the world (FAO et al., 2019). In principle, PDS beneficiaries are entitled to purchase fixed monthly quantities of grain and other commodities at a highly-subsidized price from a government-run fair-price ration shop (FPS). In practice, the resulting dual-price system creates strong incentives for corrupt intermediaries to divert grains to the open market, with nation-wide estimated rates of leakage at 42% as of 2011-2012 (Dreze and Khera, 2015). The Government of India aimed to reduce this leakage by requiring beneficiaries

[1] National biometric ID systems have recently been rolled out in Malawi, Senegal, and Uganda, amongst others, while Ghana, Kenya, and Tanzania are currently enrolling citizens. The World Bank has a dedicated initiative - ID4D - to help countries "realize identification systems fit for the digital age." (`https://id4d.worldbank.org/`).

to obtain an Aadhaar number for at least one member of their household, link (or "seed") it to their PDS account, and then authenticate their identity by scanning the fingerprints of a seeded household member each time they transacted at a ration shop.

To evaluate the impact of introducing Aadhaar in the PDS, we worked with the government of the state of Jharkhand to randomize the order in which biometric authentication was introduced across 132 sub-districts in 10 districts in the state of Jharkhand. Our evaluation sample is representative by design of 15.1 million beneficiaries in 17 of Jharkhand's 24 districts, and representative on observables of the rest of the state. Further, the integration of Aadhaar into the PDS was implemented by the Government of Jharkhand (GoJH) as part of a full-scale deployment that was being rolled out across the country. Thus, our study design allows us to directly estimate the policy-relevant parameters of interest.

The GoJH implemented this reform in two phases. In the first phase, electronic Point-of-Sale (ePoS) machines were installed in PDS shops in treated areas (as well as the rest of the state in non-study districts) and beneficiaries were required to use Aadhaar-based Biometric Authentication (ABBA) to collect their rations. The control group continued with the default of authentication based on presenting a paper "ration card" to collect benefits, where it was much more difficult for the government to verify if beneficiaries had in fact collected their rations. In the second phase ("reconciliation"), which started 11 months later after stabilizing the implementation of the ABBA system, the GoJH also started using the electronic records of authenticated transactions from ePoS machines to adjust the amount of grain that was disbursed to PDS shops, simultaneously in treatment and control areas.[2] This was done after deploying the ePoS machines to control areas in the two months prior to the onset of reconciliation. We therefore present experimental estimates of the impact of requiring ABBA to collect benefits, and non-experimental estimates of the impact of reconciliation using a pre-specified event study framework (that also uses unreconciled commodities as a control group).

We conduct our analysis with a combination of administrative and survey data. The former include data on disbursals of commodities by month across treatment and control blocks. The latter include baseline and three follow up surveys of 3,840 PDS beneficiaries to measure receipts of benefits as well as transaction costs in accessing them. One follow up survey was conducted before the onset of reconciliation and before the rollout of ePoS machines to control areas, and we use this to study the effects of ABBA alone. Two further follow-up surveys were conducted to study the effects of reconciliation.

We first examine program implementation. Jharkhand is a relatively "low capacity" state (see Section 2 for details) which one might expect to struggle with the sheer logistics of a major reform such as Aadhaar integration. Yet we find the opposite: by our first follow-up survey, which measured beneficiary experiences 6 to 8 months after treatment onset, 97% of beneficiary households in

[2]Specifically, prior to reconciliation, the government would send each FPS the full amount of grain needed to satisfy the entitlements of all beneficiaries registered to the FPS each month. After reconciliation started, the government would only disburse the amount of grain that was shown to have been distributed by the FPS as per the ePoS records assuming that the FPS dealer still had the undistributed grains in stock.

treated areas had at least one member with an Aadhaar number seeded to the PDS account, and 90% reported that transactions at their FPS were being authenticated. Overall, implementation of the reform was thorough and quick.

We then examine how this affected the transfer of value from the government to beneficiaries: the value of goods disbursed by the former (from administrative data), received by the latter (from household surveys), and the difference between these (i.e. leakage). The reform was framed by the central government as part of a broader effort to achieve fiscal savings by reducing fraud, with the possibility but no commitment that such savings would be fed back into the PDS or other safety net programs in the future. Ex ante, the impact was ambiguous, with the extent of leakage reductions depending on how exactly the government would choose to use the data on authenticated transactions to determine grain disbursals. Further, ABBA could potentially help to mitigate "identity fraud", whereby benefits were collected by others. On the other hand, ABBA could potentially increase exclusion error or add to transaction costs due to authentication failures. Moreover, it was not designed to address the problem of "quantity fraud" (where PDS dealers give beneficiaries less than their entitlement). As one official put it: "Even if someone uses his thumb and gets 2 kg of rice instead of 5 kg, what can you do?"[3] The design of the reform contrasts in this sense with the related biometric "Smartcards" reform we previously studied in Andhra Pradesh (Muralidharan et al., 2016), where the government focused more on improving the end user experience rather than on fiscal savings (a contrast to which we return below).

In practice, we find that the impacts of ABBA on beneficiaries were small on average and, where significantly different from zero, negative. Consistent with not using the data on authenticated transactions for adjusting grain disbursal in the first phase of the reform, ABBA did not decrease (and if anything slightly increased) government spending. It also did not increase mean value received by beneficiaries substantially or significantly, and had no impact on leakage. We can reject changes in value received by beneficiaries outside of $[-4.3\%, 3.9\%]$ of value disbursed, and changes in leakage outside of $[-1.7\%, 6.5\%]$. We see no meaningful changes in measures of the quality of goods received, of their market prices, or of beneficiaries' food security. Beneficiaries, however, did incur a 17% higher transaction costs to collect their benefits (a Rs. 7 increase on a base of Rs. 41), driven mainly by an increase in the number of unsuccessful trips made to the ration shop, and the opportunity cost of doing so.

While average benefits received may not have decreased, more stringent ID requirements could have excluded a vulnerable minority who were unable to meet the new identification standards. Extensive margin effects are consistent with this possibility, as treatment increased the probability that a beneficiary received no commodities at all in any given month by 2.4 percentage points ($p = 0.099$). To examine this further, we focus on the 23% of households who did not have at least one member's Aadhaar number seeded to their PDS account at baseline, and were thus at greater risk of being excluded. Unseeded households tend to be poorer and less educated than their seeded

[3] "Death by Digital Exclusion," The Hindu, `https://www.thehindu.com/news/national/other-states/death-by-digital-exclusion/article28414768.ece`, accessed 13 July 2019.

peers. Among unseeded households, exclusion errors increased significantly: the mean value of rice and wheat received fell by Rs. 49, or 8.4% of value disbursed, and the probability of receiving none of these commodities increased by 10 percentage points (a 50% increase on a base of 20%).

Overall, these results are consistent with the critique that biometrically authenticating transactions per se caused at least some "pain without gain" (Dreze et al., 2017). A potential counterargument, however, is that authenticating transactions was a necessary first step towards the second phase of the reform which the government subsequently introduced. After rolling out Aadhaar and ePoS devices in both treatment and control subdistricts, the GoJH began in July 2017 to reconcile its monthly shipments of grain to each FPS against authenticated transaction data for previous months, sending less grain to dealers it believed - based on the data - should have had more stock remaining from their previous months' allotment.

The data suggest that reconciliation initially had a substantial impact. Focusing first on the control group, which began reconciliation shortly after beginning ABBA itself, we find that the introduction of reconciliation coincided with an 18% (Rs. 92 per ration card per month) fall in the value of reconciled commodities disbursed by the government. This drop is specific to the two commodities (rice and wheat) subject to reconciliation and does not appear for the three unreconciled commodities (sugar, salt, and kerosene). Of the drop, we estimate that 22% represents a reduction in value received by beneficiaries (based on household-survey data). Thus, of the total reduction in disbursals, 78% represents a genuine reduction in leakage.

In treatment areas, effects of reconciliation were more pronounced overall, and the tradeoff between errors of exclusion and inclusion somewhat less advantageous. Disbursements fell initially by 36% (Rs. 182 per ration card per month), of which 34% represents a drop in value received by beneficiaries and 66% a reduction in leakage. The former effect was largely felt on the extensive margin: the probability that a household received no reconciled commodities increased by 10 percentage points, enough to account fully for the drop in average receipts. The larger effect in treated areas reflects the fact that, dealers in the treated areas had been implementing ABBA for roughly nine months longer. Thus, based on the ABBA records maintained on the ePoS machines, treated areas held significantly greater undisbursed stocks of grain, and thus received significantly less grain from the government. In practice, dealers may not have had these stocks of grain (likely due to having diverted undisbursed stocks to the open market), and thus the larger reduction of grain disbursement in treated areas may have been passed on to beneficiaries since they may not had grains left to distribute.

Interestingly, the reduction in leakage is also consistent with the expectations of PDS dealers: those treated early reported a 72% lower expected future bribe price for FPS licenses, suggesting that they expected a substantial fall in the potential for extracting rents from an FPS. However, in cases where the dealers had already diverted or not properly stored the undisbursed grain in prior months, the reduction in disbursal from the government may have been passed on to beneficiaries, as seen in our results.

Overall the reconciliation policy was unpopular, drawing complaints from both dealers and beneficiaries and demands for waivers and exemptions. Effects on both leakage and beneficiaries were short-lived and largely attenuated within three months, after which the GoJH temporarily rescinded the policy. One interpretation of this outcome, and especially of the difference in the effects of reconciliation between experimental arms, is that holding dealers accountable for past diversion may have contributed to the increased exclusion we find after reconciliation. The government might have obtained a more favorable tradeoff between inclusion and exclusion errors had it instead given all dealers a "clean slate" when introducing reconciliation. We examine this by exploiting the experimentally induced variation in opening balances to extrapolate to a "clean slate" scenario, and estimate that this would have reduced leakage while if anything weakly increasing value received by beneficiaries.

A longer-run analogue to reconciliation is to delete ration cards that have not been seeded, thereby removing them from the eligibility list entirely. We find that the rate of card deletions increased after the onset of ABBA and reconciliation, and that deleted cards included both true "ghosts" as well as non-ghost recipients (based on data on the subset of households we sampled and surveyed). While purely descriptive, these results highlights another margin along which the rollout of Aadhaar likely cut leakage at the cost of some exclusion.

Overall, biometric authentication in Jharkhand's ePOS was not a free lunch: depending on how it was used, it either did not reduce errors of inclusion or leakage or did so at the cost of increased exclusion error. While the combination of ABBA and reconciliation did reduce leakage, around 22-34% of the reduced disbursals represented a reduction in beneficiary receipts. If we consider the conservative case of reconciliation in the control group, a planner would need to value marginal revenues at at least 28% (i.e. 22%/78%) of the value placed on transfers to marginal households in order to prefer such a policy to the status quo. In this specific case, the benefits of reduced disbursal may have been even lower as the savings were only notional, yielding an increased stock of grain in public warehouses as opposed to reduced spending.[4] In the conclusion, we discuss several practical ways to reduce the likelihood of exclusion errors while still achieving leakage reductions.

Our most important contribution is to provide the first experimental evidence on the trade-off between Type I and Type II errors from introducing stricter ID requirements for receiving welfare benefits, and to do so using an at-scale experiment across 15 million people in the context of the largest welfare program (PDS) in the country with the largest biometric ID program in the world (India). Our results showing that reductions in leakage came at the cost of increased exclusion errors and inconvenience to beneficiaries are directly relevant to policy discussions regarding the use of more stringent ID requirements to access public services in India and other countries.[5] More broadly, they add to the evidence base on how transaction costs affect the incidence of welfare benefits (e.g. Currie (2004) and more recently Alatas et al. (2016)). As predicted by Kleven and Kopczuk (2011), they illustrate how increasing the complexity of the process of obtaining benefits

[4] Over time, fiscal savings may be possible by reducing the amount of grain procured from farmers, but no such policy change has been announced during the period of introducing ABBA and reconciliation into the PDS.

[5] They also provide a counterpoint, for example, to recent panel-data evidence that voter ID requirements have had surprisingly little effect on voter participation in the United States (Cantoni and Pons, 2019).

can affect their overall incidence by decreasing takeup among eligible households. “Complexity” does not appear to have been an effective screening device in this case, however, as the households excluded generally appear less well off on socioeconomic measures.[6]

Second, our results illustrate the potential “shadow costs” of controlling corruption. Following early theoretical debate about the costs and benefits of corruption (Leff, 1964; Huntington, 1968; Shleifer and Vishny, 1993), the last decade or so of micro-empirical work has (as we read it) typically found that corruption is harmful and that there exist methods of reducing it that are cost-effective, at least in the sense that the benefits are large compared to the direct costs of intervention, such as the cost of hiring auditors (see for example Olken (2007) or Duflo et al. (2013), among others). However, in many settings the indirect costs may also matter. Rigid procurement procedures, for example, may limit the scope for graft but also slow down decision-making and make it hard to act on “soft information” (Wilson, 1989). In our setting, the cost of controlling corruption was not just the direct cost of issuing Aadhaar numbers and biometric readers, but included the (considerable) indirect cost of excluding marginalized households from their legally entitled benefits.[7]

Finally, our findings are also relevant to ongoing discussions about “external validity” in program evaluation – that is, how best to predict the effects of reforms from the results of past experiments. In prior work (Muralidharan et al., 2016), we found that introducing biometric payments in rural welfare programs in the state of Andhra Pradesh (AP) both reduced leakage and improved the payment experience. However, the impacts of ABBA and reconciliation in the PDS in Jharkhand were quite different. Our results, and in particular the fact that the rollout in Jharkhand was much faster than that in AP, suggest that this was less because of differences in state or implementation capacity, but more because of key differences in program design and emphasis. Leakage fell in *both* cases, but the rules in AP aimed to enhance the beneficiary experience (and thus passed on the benefits of reduced leakage to beneficiaries, while achieving no fiscal savings), while those in Jharkhand focused on achieving fiscal savings (resulting in reduced disbursals and leakage, but also passing on some pain to beneficiaries). Overall, the results highlight the importance of *construct* as well as *context* in extrapolating results from one case to another. They also caution against a simplistic attempt to characterize the effects of new technologies such as biometric authentication without paying careful attention to design details and to the beneficiary experience.[8]

The rest of the paper is organized as follows. Section 2 describes the context and intervention. Section 3 presents the research design including data collection and estimation strategy. Section 4 describes results of point-of-sale authentication, Section 5 describes results of reconciliation and ration card deletion, and Section 6 offers a concluding summary.

[6]In addition to the tradeoffs we discuss here, implementing large-scale biometric ID schemes typically involves tradeoffs between state capacity and privacy. See Gelb and Metz (2018) for further discussion.

[7]In related work, Lichand and Fernandes (2019) find that the threat of audits reduced corruption but also displaced spending on services such as public health care in Brazilian municipalities, and that this led to worsening of some local public health outcomes.

[8]The point that the impacts of technology-based interventions depend crucially on design details is also seen in the literature on education technology, where impacts on learning outcomes vary widely as a function of how effectively technology is (or is not) used to improve pedagogy (see Muralidharan et al. (2019a) for a review).

2 Context and intervention

Malnutrition remains a serious problem today in India, which ranked 102 of 117 countries in the most recent Global Hunger Index Rankings (Grebmer et al., 2019) and had an estimated 38% of children stunted and 36% underweight as of 2015-2016 (UNICEF et al., 2017). The Public Distribution System (PDS) is a central piece of the government's efforts to provide food security to the poor. Through a network of over 527,000 ration shops known as "Fair Price Shops" (FPS), it disburses subsidized wheat and rice to targeted households on a monthly basis, and other commodities such as sugar, salt, and kerosene on an occasional basis. Under the National Food Security Act of 2013, the government has a mandate to include 75% (50%) of the rural (urban) population as beneficiaries. Individual states administer targeting and distribution within their boundaries. Overall, the PDS costs roughly 1% of GDP to operate.[9]

Because it creates a dual-price system, distributing commodities at prices well below their market prices, the PDS has historically suffered from various forms of diversion. Commodities "leak" from the warehouses and trucking networks meant to deliver them to the Fair Price Shops, or from the shops themselves; dealers adulterate commodities or over-charge for them. Historically estimated leakage rates have been high; Dreze and Khera (2015) estimate that 42% of foodgrains nationwide and 44% in Jharkhand were diverted in 2011-2012, which is itself an improvement on the estimate of 73% by the Planning Commission in 2003 (The Programme Evaluation Organisation, 2005).

Various reforms meant to address these challenges are underway, including several grouped under the broad heading of "PDS computerization." We focus on one of the major components of computerization: the introduction of electronic point-of-sale (ePOS) devices to process and record transactions between dealers and beneficiaries. As we describe below, these devices enabled Aadhaar-based biometric authentication (ABBA) as well as the creation of a digital transaction ledger. Rollout of these devices was well underway elsewhere in India by the time the GoJH began its deployment; as of July 2016 an estimated 23% of India's FPSs had received devices, rising to 54% by December 2017[10] with the rollout ongoing.[11]

ePOS devices perform biometric authentication using Aadhaar, India's landmark unique ID system. The Government of India launched Aadhaar in 2009 with the goal of issuing an identification number linked to biometric information for every resident of the country. As of June 2019, it had issued Aadhaar numbers to 1.24B people, or 91% of the country's population.[12] Investments in ID

[9]The PDS is enabled in part by India's policy of a Minimum Support Price for essential commodities like rice and wheat combined with public procurement of these commodities from farmers. The resulting stocks of foodgrain with the government are then distributed to the poor through the PDS. In this way, Indian agriculture and food policy intervenes in both the production and distribution side of the market. For PDS expenditures, see `http://www.indiabudget.gov.in/ub2018-19/eb/stat7.pdf`. For GDP estimates, see `https://dbie.rbi.org.in/DBIE/dbie.rbi?site=statistics`. Both sources accessed on 5 March, 2018

[10]For July 2016 statistics, see `http://164.100.47.190/loksabhaquestions/annex/9/AS26.pdf`/. For December 2017 statistics, see `http://pib.nic.in/PressReleseDetail.aspx?PRID=1512902`. Both sources accessed 5 March 2018.

[11]Other PDS computerization initiatives included digitization of beneficiary databases, computerization of supply-chain management, and creation of grievance redressal mechanisms and online transparency portals.

[12]For statistic on number of Aadhaar UIDs generated, see `https://uidai.gov.in/aadhaar_dashboard/india.php`.

could be particularly important in India given its historically unusual situation as a country with a substantial welfare state at relatively low levels of per capita income, and indeed the government has touted Aadhaar as an enabling technology which will support reforms to the implementation of a wide range of government schemes – "a game changer for governance," as the Finance Minister at the time put it (Harris, 2013). Abraham et al. (2017) estimate that it was being applied to at least 558 use cases as of 2017. Government claims regarding the fiscal savings achieved by introducing Aadhaar have at times been met with skepticism (Khera, 2016), however, in part because they did not differentiate between real reductions in leakage and increased exclusion of legitimate beneficiaries. To our knowledge, however, there has been no experimental evidence to date on the impacts of an Aadhaar deployment in any welfare program.

Jharkhand is a relatively challenging environment in which to roll out an ambitious reform such as ABBA. In terms of state capacity, it ranked 17th of 19 major states on the most recent Governance Performance Index (Mundle et al., 2012), well below 3rd-ranked Andhra Pradesh in which our previous evaluation of biometric authentication was set. As one concrete example, it had the highest rate of teacher absence among *all* Indian states in both 2003 and 2010 (Muralidharan et al., 2017). Jharkhand also rated relatively low in terms of key pieces of enabling infrastructure such as rural teledensity (40 telephone or mobile phone connections per 100 people in rural Jharkhand as of 31 October 2017, ranked 19 out of 19 reported states) and at the middle of the pack for Aadhaar penetration (93% penetration as of 31 December 2017, ranked 17th of 36 states).[13]

2.1 The intervention

In August 2016, the GoJH introduced ePoS machines in FPSs to authenticate beneficiaries when they came to collect their rations (Figure 2 provides the rollout timeline). In August 2016, the GoJH introduced ePoS machines in FPSs to authenticate beneficiaries when they came to collect their rations (Figure 2 provides the rollout timeline). Prior to the intervention, authentication in the Jharkhand PDS was relatively informal. Each beneficiary was assigned to a unique FPS and issued a ration card listing members of the household and displaying a photograph of the household head. To collect benefits, any one of these listed household members was required to appear in person with the ration card at the assigned FPS. Anecdotally it was not uncommon for neighbors or friends to collect benefits on their behalf, or for dealers to hold on to beneficiaries' ration cards themselves. Dealers were expected to record transactions both on ration cards and in their own ledgers; ledgers were typically not audited, and anecdotally there was wide variation in record-keeping practices.[14]

The reform modified authentication and record-keeping processes. The state gave each dealer

For total population statistics, see `https://data.worldbank.org/indicator/SP.POP.TOTL`.

[13]For rural teledensity statistics, see `http://164.100.47.190/loksabhaquestions/annex/13/AU2751.pdf`, accessed March 5, 2018. For Aadhaar penetration statistics, see `https://uidai.gov.in/enrolment-update/ecosystem-partners/state-wise-aadhaar-saturation.html`, accessed January 31, 2018.

[14]One common practice is to keep separate "official" and "unofficial" ledgers, where the unofficial ledgers accounted for actual transactions including leakage while official clean ledgers would be produced in case of a government audit.

an ePOS device configured to authenticate beneficiaries in one of three modes: online, offline, and partially online.[15]

In *online mode*, the device required the operator to input a ration card number. It then displayed a list of all individuals who were both (i) listed as beneficiaries on the relevant ration card, and (ii) had an Aadhaar number linked ("seeded") to the card. The dealer selected the beneficiary present, and the device then prompted him/her to place a finger of choice on the device's scanner to be authenticated against the central Aadhaar database. If fingerprint authentication failed on three consecutive attempts, the beneficiary could opt to receive a one-time password texted to their mobile phone number as a fallback method of authentication.[16]

In *offline mode*, the device simply captured and stored fingerprint information for the person collecting benefits but performed no authentication checks. However, transaction logs were meant to be synchronized with a server periodically (as explained below).

In *partially online mode*, the device functioned as in online mode if it detected a network connection and in offline mode otherwise. Dealers did not have discretion to select modes (but could potentially have tried to force the device to operate in offline mode by disrupting connectivity).

The government varied the mode assigned to each FPS in an effort to balance the risks of inclusion and exclusion error: it sought to enforce relatively strict authentication requirements in areas where connectivity was strong enough to provide a reliable connection to the central Aadhaar database, but not deny benefits to legitimate beneficiaries in areas where connectivity was weaker.[17] In our experimental design assignment to receive a machine was random but assignment to machine mode was not, so that the effects we report represent an average of mode-specific effects given the assignment policy described here. We also report a non-experimental decomposition assuming that, had they been treated early, control Fair Price Shops would have been treated with the same machine mode to which they were subsequently assigned.

ePOS devices also enabled digital record-keeping. After authentication, the device would display any previously uncollected commodity balances to which the beneficiary was entitled, including the current months' entitlement and any uncollected balance from the previous month (but not balances from two or more months previous). After completing a transaction the dealer would record the amount of each commodity purchased in the device, which would print a paper receipt and also voice the transaction details in Hindi. Dealers were instructed to give the receipt to the recipient as well as recording the transaction in their ration card. In practice, recipients often reported not receiving receipts or that these faded quickly. In any case, the digital ledger maintained in the device became the source of truth for balance information from the government's perspective,

[15] Most FPS were assigned to the online mode (81% of shops), with 15% offline and only 4% partially online on average prior to August 2017. In August 2017 the government ended the use of partially online mode after which 88% of FPSs operated in online mode with the remaining 12% offline.

[16] Some officials claimed that at least initially if neither method of authentication succeeded there was an "override" option available allowing the dealer to authenticate a beneficiary without using Aadhaar, but officially no such option was meant to exist.

[17] In data collected by our survey team, the proportion of FPS at which no cellular signal could be detected was 5% for shops with online devices, 10% for shops with partially online devices, and 58% for shops with offline devices.

though dealers were of course free to maintain their own parallel paper records if they wished.

The government accessed transaction data by synchronizing ("syncing") regularly with each device. Online devices synced their records with a central government server automatically in real time. Dealers using partially online and offline devices were instructed to sync data within 48 hours of a transaction, but did not face any obvious repercussions if they did not. Instead their binding constraint appeared to be monthly: devices would not authorize new transactions in a given month until the previous month's transactions had been synced.

The process of seeding Aadhaar numbers to ration cards was ongoing during the period we study. To seed their ration card, a household first needed to have at least one of the members listed on the ration card obtain an Aadhaar number, either at camps organized specially for this purpose or subsequently by applying at the local block or district office. It then needed to link this Aadhaar number to its ration card, again either at camps organized for this purpose during NFSA enrollment or by applying at the block or district office. As of May 2016, 76.5% of ration cards in areas assigned to treatment and 79.9% of those in areas assigned to control had been seeded with at least one Aadhaar number. These figures had risen to 94.5% and 92.7%, respectively, by October of 2016. Finally, by May 2018, these figures had risen further to 99.8% and 99.5%, respectively. The seeding process could itself have affected errors of inclusion and exclusion, e.g. if the government choose to delete ration cards that had not been seeded after some interval in an effort to eliminate ghosts. The GoJH's stated policy was not to do so, but anecdotes circulated of cases in which this occurred. We examine this further in Section 5.3 below.

2.2 Reconciliation

Prior to the introduction of Aadhaar-based biometric authentication using ePOS devices, the GoJH rarely (if ever) reconciled balances with FPS dealers. For example, if the grain needed to serve all PDS beneficiaries assigned to a given FPS was 100kg of rice per month, it was GoJH policy to ship 100kg of rice to that FPS each month irrespective of how much rice it had distributed to beneficiaries in previous months. This reflected in part the simple fact that the government had no timely and reliable data on transactions at the shops.

By June of 2017, ePoS devices were actively in use for authentication in 93% of Fair Price Shops in our study area, including those in control blocks, where they were rolled out during April and May. Starting in July, therefore, the government introduced a second reform, reconciling its disbursements of rice and wheat, though not of sugar, salt or kerosene. The full formula the government used to determine disbursements under this regime is in Appendix C. To summarize, the government's new policy was to calculate (a) the amount each dealer would need to meet claims by beneficiaries against the current month's entitlements, as well as any outstanding claims on the preceding one month's entitlements, and (b) the amount the dealer should have in stock given the full history of deliveries and transactions (starting from the time the FPS first used an ePoS device), and then disburse the difference between these quantities.

From a dealer's perspective, this reform (if implemented by the book) had two effects. First, it had a retrospective effect, reducing the amount of rice and wheat received starting in July: dealers who had not distributed the full amounts disbursed to them in previous months (as recorded by the ePoS machines) received less. We would expect this effect to be larger for dealers in treatment blocks, since as of July they had been using devices for 11 months as opposed to 1-2 months for dealers in control blocks. As we discuss below, many dealers had "opening balances" at the onset of reconciliation equivalent to over a month of entitlement (based on the amount of grain they had received in previous months against which no authenticated transaction log existed). This implied that by rule they should have received no incremental grain at all, since they were supposed to be holding enough stocks of grains to make all program-required disbursals in July. Second, the reform prospectively affected dealers' *marginal* incentives to report via the ePOS devices that they had distributed grain to ration card holders, since reporting less than full distribution would reduce the amount they received the next month. The legitimate incentive to do so was the commission of 1 rupee per kilogram of grain they received for distributing commodities. In addition, receiving more grains would also make it easy to divert some while still providing beneficiaries with a given level of benefits.

From a beneficiary's perspective the consequences of reconciliation are unclear. On one hand, dealers might pass on some share of the reduction in grains disbursed to them, reducing in turn the amounts distributed to beneficiaries. On the other hand, dealers might distribute more grain to beneficiaries in order to increase future disbursements. Note, however, that strictly speaking the reform created incentives for dealers to *report* that they had distributed grain, not to actually distribute it. They needed beneficiaries to appear and scan their fingerprints to do so, but did not need to give beneficiaries the amount of physical grain that they recorded. Anecdotally, some dealers told beneficiaries that they would enter the full amounts into the devices even while distributing less or none, since otherwise there would be less grain to distribute in the subsequent month. To summarize, using the terminology in Niehaus and Sukhtankar (2013), ABBA and reconciliation likely made it more difficult for dealers to divert grains through "over-reporting" the number of beneficiaries (including making up fake or ghost beneficiaries), but it may not have altered their incentives for "under-payment" of benefits to genuine beneficiaries.

2.3 Summary

Overall, the reforms introduced by the GoJH were representative of the way in which the Government of India has envisioned using Aadhaar to reform program administration. In particular, they made possession of an Aadhaar number effectively mandatory for the receipt of PDS benefits (despite a 2013 Supreme Court ruling prohibiting this).[18] A priori one would thus expect it to have both strong potential to reduce errors of inclusion, and a high risk of generating additional errors of exclusion. Media criticism has argued that it has done exactly that, leading in some cases to

[18] `http://judis.nic.in/temp/494201232392013p.txt`

preventable starvation deaths –"death by digital exclusion," as one headline put it.[19]

Compared to the Smartcards reform we previously studied in Andhra Pradesh (AP) (Muralidharan et al., 2016), the reform in Jharkhand exhibits some similarities and some differences. Both interventions introduced biometric authentication, but the implementation was generally stricter in Jharkhand. Many devices in Jharkhand operated in online mode and required connectivity to function, while all devices in AP operated offline; devices in AP featured a manual override mechanism for use if biometric authentication failed, while we have no evidence to suggest such a mechanism was used in Jharkhand. Meanwhile, (food) balances in Jharkhand were not reconciled initially, while (cash) balances in AP were reconciled from the outset. Finally, in AP the location where program participants collected their benefits was moved (from post offices to customer service providers within each village), while in Jharkhand the location was held constant at the Fair Price Shop.

3 Research design

Our research design follows a pair of pre-specified and pre-registered analysis plans, one for the evaluation of Aadhaar-based authentication itself and another for the analysis of reconciliation.[20] Appendix D provides a comprehensive list of analysis reported in addition to what was pre-specified.

3.1 Randomization

To obtain policy-relevant estimates, we sought to design an evaluation that was "at scale" in each of the three senses identified by (Muralidharan and Niehaus, 2017). These include conducting our study in a sample that is representative of the (larger) population of interest, studying the effects of implementation at large scale, and having large units of randomization to capture general equilibrium or other spillover effects such as changes in the market prices of subsidized commodities.[21]

We first sampled study districts. Of Jharkhand's 24 districts, we excluded 1 in which the intervention rollout had already begun and 6 in which a related reform (of Direct Benefit Transfers for kerosene) was being rolled out. From the remaining 17 districts, home to 24M people and 15.1M PDS beneficiaries, we randomly sampled 10 within which to randomize the rollout of the intervention.[22] This design ensures representativeness of the 17 districts in our frame. In practice

[19] "Death by Digital Exclusion." The Hindi, 13 July 2019. `https://www.thehindu.com/news/national/other-states/death-by-digital-exclusion/article28414768.ece`, accessed 13 July 2019.

[20] `https://www.socialscienceregistry.org/versions/39275/docs/version/document` and `https://www.socialscienceregistry.org/versions/39274/docs/version/document` respectively.

[21] Each of these three design choices helps to improve external validity. Conducting experimental evaluations in near-representative samples helps by reducing the risk of site-selection bias (Allcott, 2015). Evaluating a large-scale implementation helps because effect sizes have been shown to decline with size of implementation (Vivalt, forthcoming), Finally, randomizing large units into treatment and control status helps produce estimates that are inclusive of spillovers, which have been shown to be salient for policy in several studies including Cunha et al. (2018), Muralidharan et al. (2016), and Egger et al. (2019).

[22] We used stratified random sampling, stratifying on three variables related to geography and socio-economic status. We used these 3 binary variables to classify the 17 available districts into 8 (2x2x2) distinct categories. We then sampled half of the districts in each category, rounding down to the nearest integer and us-

our 10 study districts appear comparable on major demographic and socio-economic indicators to the 14 remaining districts of Jharkhand (Table 1). Our frame is thus arguably representative of the full population of 5.6M PDS households and 26M PDS beneficiaries in the state.

The evaluation was conducted within the context of a full-scale rollout, as the GoJH deployed ePOS devices to 36,000 ration shops covering the entire population of 26M PDS beneficiaries in the state. This deployment involved a major effort by the government and was the stated top priority of the Department of Food and Civil Supplies for the year and (anecdotally) the single largest use to which they put staff time. We thus measure the effects of implementation at full scale by a bureaucratic machinery fully committed to the reform, which are the effects of interest for policy purposes.

Finally, we assigned treatment to large units. We randomized the rollout at the level of the sub-district ("block"), which on average covers 73 Fair Price Shops and 96,000 people. Figure 1 maps treated and control blocks and illustrates their geographic balance and coverage of the state. We allocated 132 blocks into a treatment arm of 87 blocks and a control arm of 45 blocks, reflecting the government's preference to delay treatment in as few blocks as possible.[23] Treatment and control blocks are similar in terms of demographic and program characteristics, as one would expect (Table 2, Panel A). Of 12 characteristics we examine, one is marginally significant at the 10% level.

The GoJH complied closely and quickly with the treatment assignment. By the time of our follow-up survey, households in treated blocks reported that 96% of dealers in treated blocks possessed an ePOS device and 91% were using it to process transactions (Table 2, Panel B).[24] ePOS utilization was stable at 90-91% in treated blocks during January-March 2017, which increases our confidence that we are estimating steady state impacts and not transitional dynamics. In control blocks, on the other hand, 5% of dealers possessed a device and 6% were using it process transactions, largely reflecting early rollout in one control block.[25] Overall these figures suggest that it is sensible to estimate intent-to-treat effects and to interpret them as fairly close approximations of the overall average treatment effect.

ing probability proportional to size (measured as number of Fair Price Shops) sampling, and lastly sampled additional districts without stratification to reach our target of 10. Full details in the Pre-Analysis Plan: `https://www.socialscienceregistry.org/versions/39275/docs/version/document`.

[23]Within each district, we assigned blocks to treatment status as follows: We first divided blocks into rural and urban samples, then stratified them into groups of three by ordering them on the first principal component of three variables related to household size and benefit category. Within each group of 3 blocks we randomly assign 2 to treatment and 1 to control. Full details, including how we dealt with districts with residual strata of fewer than 3 blocks, in the Pre-Analysis Plan: `https://www.socialscienceregistry.org/versions/39275/docs/version/document`.

[24]This rollout was substantially faster than the Smartcards rollout in Andhra Pradesh, for example, which took 2 years to cover 50% of transactions. The rollout of (non-Aadhaar) biometric authentication in Andhra Pradesh is analogous to the one we study here took six months to complete, according to government claims. See `http://dfpd.nic.in/1sGbO2W68mUlunCgKmpnLF5WHm/mergedoriginal.pdf`, accessed 27 July 2018. The difference in speed of execution reflects a combination of the priority placed on implementation by the national Department of Food and Public Distribution as well as the fact that Aadhaar was more prevalent across citizens prior to the integration with the PDS compared to Smartcards where enrollment had to be done from scratch.

[25]Of the 31 control households that report a dealer using an ePOS device, 24 are in one block. The remaining 7 are scattered across 6 other blocks and most likely reflect reporting errors.

3.2 Sampling and Data Collection

Our data collection focused on measuring three core concepts: the value of commodities disbursed by the government, the value of commodities received by beneficiaries (both net of price paid), and the real transaction costs incurred by dealers and beneficiaries to implement/obtain this transfer of value. Leakage in this framework is simply the difference between value disbursed and value received. Our concept of transaction costs includes a number of costs incurred by beneficiaries such as the amount of time required to collect rations.

To measure these quantities we begin with administrative records. These include information on monthly quantities of commodities disbursed to individual Fair Price Shops, which we obtained from the National Informatics Commission.[26] We obtained this data for *all* FPS shops in study blocks and we use this administrative data as our primary source of data on disbursals (as opposed to that reported by the sample of dealers we surveyed).

We next used the administrative database of eligible PDS beneficiaries and their assignment to Fair Price Shops to draw samples of dealers and households to survey, and attempted to survey them four times – once at baseline and then at three subsequent follow-ups. We sampled as follows: from administrative records we drew a sample of 3 Fair Price Shops (FPS) in each study block, for a total of 396 shops.[27] We successfully interviewed the dealers operating 367 (93%) of these shops at baseline, and 373 (94%) of them in the endline. Dealer surveys covered measures of the quantity of commodities received by the shop each month, their operating costs, the dealers perceived value of FPS licenses and interest in continuing to operate a ration shop, and stated preferences for the reform as opposed to the status quo system. Enumerators also measured by hand the strength of the four major cellular networks at the shop in order to capture connectivity.[28].

For each sampled ration shop we sampled 10 households from the government's list of PDS beneficiaries,[29] which had been created as part of a targeting exercise conducted in 2015 to comply with the National Food Security Act of 2013. This generated a target sample of 3,960 households.

[26]In some cases we were also able to obtain and digitize disbursement records directly from District Supply Officers, Market Supply Officers, Block Development Officers, and godowns run by the Food Corporation of India and the state of Jharkhand. These records generally correlated strongly (from 0.87 to 0.95 for various commodity $\times$ month pairs) but not perfectly with the NIC records. We use the NIC records to ensure representative coverage, but obtain qualitatively similar results if we use the hand-captured ones instead.

[27]We sampled a fixed number of shops per block as we had previously selected study districts using PPS sampling.

[28]In follow-up surveys, we expanded the number of dealers surveyed, as a few (7.9%) of our sampled households had been re-assigned to new dealers in the normal course of operations during the 10 months since baseline. We report results for both the original and augmented dealer samples, as the reassignment rate of households is balanced across treatment and control, and the incremental dealers are not statistically distinguishable from the original ones on measured characteristics (Table A.1). Note also that the reassignment of households to other shops does not affect our ITT estimates because we follow the originally sampled households. It also does not affect the first-stage or the interpretation of our results because the reassignment was to other FPS in the same block, with the same treatment status (which is another advantage of randomizing at the block level)

[29]We define household here as those individuals belonging to a single ration card. We first sampled one village from the catchment area of each FPS using PPS sampling, with "size" defined as the number of ration cards in the village assigned to that FPS. We sampled ration cards using stratified random sampling, with strata including the method by which the household became eligible for the PDS and the benefit category to which the cardholder is entitled. Full details in the Pre-Analysis Plan, `https://www.socialscienceregistry.org/versions/39275/docs/version/document`.

We attempted to interview these households for baseline and three follow-up surveys to create a household-level panel. [30]

Of the 3,960 ration cards we sampled, we identified and interviewed the corresponding household at least once in 97% of cases.[31] Overall, we estimate that at most 3% of beneficiaries were ghosts (see Figure A.1 for a more detailed categorization of households). This is noteworthy as it suggests that the scope for eliminating leakage by removing ghosts (or non-existent households) from the beneficiary list was relatively limited in this setting.

We timed follow-up surveys and their associated recall periods to obtain continuous monthly data on beneficiaries' experiences with PDS from January through November of 2017. Figure 2 illustrates the recall window covered by each survey. We use data from follow-up 1, covering January through March, to measures the impacts of ABBA, and use data from all three follow-ups to examine the impacts of reconciliation. Topical coverage varied across surveys; follow-up 1 was most comprehensive, while follow-ups 2 and 3 measured a subset of outcomes (e.g. for households, the quantities of each commodity received). In particular, we did not measure market prices in follow-ups 2 and 3 and so do not examine price effects of reconciliation.

3.3 Estimation strategy: Aadhaar-based biometric authentication

To examine the impacts of ABBA we estimate intent-to-treat specifications of the form

$$Y_{hfbs}^{t} = \alpha + \beta Treated_{bs} + \gamma Y_{hfbs}^{0} + \delta_s + \epsilon_{hfbs}^{t} \tag{1}$$

where Y is an outcome measured for household h assigned to Fair Price Shop f in block b of stratum s.[32] Regressors include an indicator T for whether that block was assigned to treatment, the baseline value Y_{hfbs}^{0} of the dependent variable, and a stratum fixed effect δ_s. Where we observe baseline values for multiple months we take their average. Where the baseline value is missing we set it equal to the overall mean value, and include an indicator for baseline missingness. When using survey data we weight specifications by (inverse) sampling probabilities to obtain results that are representative of the sample frame.[33] We use analogous specifications for outcomes measured at the level of the Fair Price Shop or block. We pool observations for January-March 2017, following our pre-specified plan for dealing with the possibility of non-stationary treatment effects.[34]

[30] Because our frame is the universe of households previously deemed eligible for the program, our sample is not suited to examine errors of inclusion and exclusion in the process of determining *who is eligible* for the PDS, as in the extensive literature on poverty targeting. Our focus here is rather on studying changes in inclusion and exclusion resulting from increased stringency in verifying the *identity* of those previously deemed eligible for benefits.

[31] We successfully interviewed 3,410 (86%) of these households at baseline and 3,583 (90%), 3,618 (91%), and 3,562 (90%) at follow-ups 1, 2 and 3, respectively.

[32] Because the randomization algorithm created 6 strata (3 urban and 3 rural) of size 1, we create a single fixed effect δ_s for each of these two groups.

[33] Variation in sampling probabilities was driven largely by field logistics constraints, e.g. the need to plan to interview a fixed number of households per village rather than a fixed proportion.

[34] Specifically, we pre-specified that we would (i) estimate models for each month individually, pooled models, and pooled models with a linear interaction between treatment and month, and then (ii) choose which specification to privilege based on the overall tendency of the trend terms to be significant predictors of primary outcomes. We

Each regression table below reports the percent of the original sample for which data were non-missing and included in the estimation. In Tables A.2 and A.3 we examine missingness by treatment status and generally do not find evidence of imbalance, with 9% of differences significant at the 10% level. We impute zeros when calculating quantities and value received for verified "ghost" ration cards (which account for 1.6% of sampled households and do not differ across treatment and control groups).

3.4 Estimation strategy: reconciliation

To examine the effects of reconciliation we examine time series variation in value disbursed and received using the following pre-specified model:

$$Y_{hfbst} = \alpha_{hfbs} + \gamma t + \beta_R R_t + \beta_{Rt} R_t (t - t^*) + P_t + \epsilon_{hfbst} \tag{2}$$

where R_t is an indicator equal to one if disbursements for month t were calculated using the reconciliation formula (i.e. for July through October), t^* is the first month of reconciliation (i.e. July), and P_t is an indicator for the one post-reconciliation month in our data (i.e. November). We estimate the model separately for treated and control blocks. To compare the two, we pool the data and interact (2) with an indicator for treatment. We report standard errors clustered by FPS. In most cases, we have a single well-defined summary measure of outcomes such as value disbursed or received. We adjust for multiple-hypothesis testing when reporting outcomes at the individual commodity level, reporting both standard p-values and q-values adjusted to control the false discovery rate.

This specification embodies several substantive assumptions. First, we assume the effect of reconciliation is identified once we control for a linear pre-trend. This is a strong assumption, but the best that is realistic with 6 months of pre-treatment data (and as it turns out yields an excellent fit). Second, because we include a distinct indicator for November we do not impose that outcomes immediately revert to what they would have been absent the intervention. While the latter assumption would significantly improve power if true, we find it implausible. Third, we model the potential for (linear) time variation in the treatment effect. This reduces power and increases the risk of overfitting if the treatment effect is in fact time-invariant, but seems appropriate given both that (a) theory suggests reconciliation should generate transitional dynamics, and (b) anecdotes suggest that the government granted many waivers to the reconciliation policy, and these may vary over time. Finally, we present results using time-series variation in value disbursed and received for both reconciled commodities (rice and wheat), and unreconciled ones (salt, sugar, and kerosene). The latter commodities provide a plausible contemporaneous control group to examine the effects of reconciliation.

generally do not observe any evidence of trends, and therefore privilege the pooled estimators. This is consistent with the fact discussed above that program implementation also appeared to have stabilized by the time of our follow-up. For completeness we report the other estimators in Appendix B.

4 Results: Biometric authentication

4.1 Value transfer

We measure value (V) as the sum across commodities of quantity (Q) multiplied by the difference between the local market price (p^m) of that commodity and the statutory ration shop price (p^s).[35] Formally,

$$V_{ht} = \sum_c Q_{cht}(p^m_{ht} - p^s_{ht}) \tag{3}$$

In total, ration card holders are entitled to a meaningful monthly amount. The quantity of commodities each household can purchase is capped at levels that depend on the category of ration card it holds and the size of the household; the mean value of these entitlements evaluated using Equation 3 is Rs. 595 per month, equivalent to 14% of the national rural poverty line for an average household in our sample.[36]

In practice, households receive less than their entitlement. The mean value received in the control group at follow-up was Rs. 463 per month, or 78% of the mean entitlement. This was largely not because the government failed to disburse commodities, at least according to its own records, as it disbursed commodities worth an average of Rs. 584 per month, or 98% of mean value entitled. Rather, it reflects the fact that roughly 21% by value of the commodities the government did disburse did not reach beneficiaries.

4.1.1 Value disbursed

Table 3 summarizes impacts on value transfer during January-March 2017, beginning in Panel A with value disbursed by the government. Note that we observe this outcome for the universe of FPS's in our study area and therefore use all of these data, with outcomes expressed per rationcard $\times$ month. We expect no meaningful changes to disbursements, as the government's policy during this period was to disburse to each FPS in each month the full amount to which households assigned to that shop were entitled. We find this is largely the case, though we do find some modest substitution away from wheat and towards rice which results in a small but significant increase in total value disbursed of Rs. 12 per ration-card month, or around 2%. This may reflect adjustments in treated areas to an informal policy the government maintained of accommodating regional differences in preferences for rice as opposed to wheat while keeping the total quantity of foodgrains fixed at their entitled value.[37] In any case, there is no evidence that ABBA saved the

[35] We find very little evidence of over-charging (below), and hence our results are essentially the same if we use actual as opposed to statutory ration shop prices. We obtained data on local market prices for equivalent commodities as those provided by the ration shops. Even if the prices reflected higher quality of market grains, that would not affect our leakage calculations because we use the same price to estimate both value disbursed and value received.

[36] An average household in our sample had 4.4 members, and the national rural expenditure poverty line was Rs. 972 / person / month (Commission, 2014). The poverty line had not been updated since 2014; if we adjust it upwards for changes in the rural consumer price index from 2014-2017, then the mean entitlement was 13% of the poverty line for an average household.

[37] Specifically, government policy was to provide rice to rural blocks while providing rice and wheat in 3:2 proportions to urban blocks and to exceptional rural blocks that expressed a desire for wheat. Given preferences for rice in

government money.

4.1.2 Value received

Panel B reports effects on value received by households using survey data. We see some directional evidence of the shift from wheat to rice noted above, but no significant change in overall value received. A 95% confidence interval for this effect is Rs. $[-25.2, 22.8]$, letting us rule out a decrease greater than 4.3% of value disbursed or an increase greater than 3.9% of value disbursed. Any effects on value received by the average household were thus small in economic terms.

We also examine whether our null effect on total value masks offsetting changes in the underlying prices and quantities. If the intervention reduces quantities flowing into rural markets which in turn raises market prices, we might see no overall effect even though recipient welfare had changed. As Panel B of Table A.4 reports, we see no significant changes in the mean quantity of any commodity received, though directionally there appears to be a shift from wheat to rice as noted above. The market prices households faced for these commodities also did not change significantly, with the possible exception of a fall in the price of sugar which is marginally significant after adjusting for multiple testing (Table A.5, Panel A).[38]

Our quantity-based measure of value received does not account for potential variation in the quality of commodities. Allegedly, PDS dealers sometimes adulterate the goods they sell (e.g. by adding sand or stones to wheat) or sell spoiled goods (e.g. rotten grains). We address this in two ways. First, we asked respondents who had completed purchases whether they received adulterated or low-quality goods. Generally speaking, few beneficiaries report experiencing these issues and rates are unaffected by treatment (Table 4). In the control group, reported adulteration rates range from 1% to 9% and none change significantly with treatment. Reported rates of quality issues are similarly low with the exception that 38% of control households report receiving low-quality salt; this rate is 6% lower among treated households, with the difference significant before but not after adjusting for multiple comparisons.[39]

Second, we asked respondents what amount of money they would have been willing to accept in lieu of the bundle of goods they purchased at the ration shop in each month. This metric has important limitations; it measures stated as opposed to revealed preferences, and requires asking subjects a series of questions which they often find confusing.[40] On the other hand, it has the

Jharkhand, it is possible that more exceptions were made in treated areas. However, these exceptions were made on an ad-hoc basis and were not recorded.

[38]Interestingly, dealers do report facing lower prices, notably for rice (Table A.5, Panel B). We view these data cautiously as (i) unlike the household reports they are not based on actual transactions, and may not reflect the pricing that is relevant to the beneficiaries whose welfare we wish to examine, and (ii) only the effect on the price of rice survives the adjustment for multiple testing.

[39]Response rates for these outcomes are relatively low as we observe them only for households that purchased a positive quantity.

[40]For households that purchased PDS commodities in a given month we elicited their stated willingness-to-accept (WTA) by asking for a series of values ranging from Rs. 100 to Rs. 2,000 whether they would have preferred to receive that amount of money to the opportunity to purchase the commodities they obtained. We define their WTA as the smallest amount of money for which they answered "yes." For households that did not purchase any PDS commodities in a given month we define WTA as zero. An unusually high proportion (48%) of respondents answered "yes" for a

advantage of capturing all aspects of both quantity and quality as perceived by the beneficiaries. When we replace our default measure of total value received with this alternative measure, we estimate an insignificant reduction in value received of Rs. 11 per month, equal to 1% of the control group mean. A 95% confidence interval for the treatment effect is $[-5.5\%, 3.6\%]$, again letting us reject substantial changes in value received in either direction.

4.1.3 Leakage

Given that value disbursed increased slightly while value received as unchanged, we do not expect to find reductions in leakage. Panel C of Table 3 tests this directly. We use a Seemingly Unrelated Regressions framework with the ration card $\times$ month as the unit of analysis and with (i) value received as reported by the household, and (ii) value per ration card disbursed to the corresponding block as the dependent variables, and then report the difference between the estimated treatment effects on these variables.[41]

We estimate that leakage increased insignificantly by Rs. 14 per ration card $\times$ month. We can reject large decreases in leakage: a 95% confidence interval is $[-10, 38]$ which lets us reject changes in the share of value lost outside of $[-1.7\%, 6.5\%]$.

Because the value figures in Table 3 are based on the difference between market and statutory ration shop prices, they pick up leakage on the quantity margin (e.g. the diversion of food grains) but may not pick up leakage due to overcharging by the FPS dealer. We examine this separately in Panel D of Table A.5. Overall, we estimate that the average control group household overpaid by Rs. 8 for the bundle of commodities it purchased, representing a small share (less than 2%) of total value received. Treatment reduced overcharging by a statistically insignificant Rs. 2.6. This makes sense given that the intervention did not directly change marginal (dis)incentives for over-charging.

4.2 Transaction costs

The transaction costs incurred in order to transfer value from the government to beneficiaries include the government's cost to store and ship commodities, the FPS dealer's cost to receive, store, and retail them, and the beneficiaries cost to travel and collect them.

Using budgetary records, we calculate that the Government of Jharkhand spends an average of Rs. 127 per ration card per month operating the PDS. The government paid around Rs. 1,600 per month per e-PoS machine to an IT provider inclusive of equipment rental, maintenance, and training. The average FPS in our data has 257 households, yielding an estimated cost of ePoS deployment of Rs. 6.2 per ration card per month. While, it is possible that some administrative costs associated with paper-based record keeping were reduced (including time taken to do so),

range of values but then subsequently answered "no" for at least one higher value, likely reflecting confusion about the nature of exercise. We believe that the lowest "yes" coding is the most reasonable way to interpret the data, but clearly they should be treated with caution.

[41]This approach lets us take advantage of the fact that we observe the universe of disbursements while also exploiting potential efficiency gains due to covariance in the error terms in the two equations.

these savings were not reported in any official spending records. Thus, we treat the costs of ePoS deployment as the change in administrative cost in treatment areas, which was a 5% increase.

Using dealer survey data, we estimate that PDS dealers spend Rs. 7 per ration card × month on the direct costs of transporting and storing PDS commodities. This is a lower bound on the total cost of distributing PDS commodities as it does not reflect costs that are shared between PDS and other activities – for example, rent paid by a shopkeeper who uses space both to distribute PDS commodities and to retail other commercial products. With that caveat in mind, we reject economically meaningful treatment effects on the portion of dealer costs we do observe (Table 5, Columns 2-3), which is what we would expect given the lack of an impact on quantities.

Using household survey data, we estimate that the average control group household spent the monetary equivalent of Rs. 41, or 9% of mean value received, in order to collect its benefits in March 2017. We calculate this using information on the individual trips they took to the ration shop, whether each trip succeeded, the time each trip took, and any money costs incurred (e.g. bus fare), as well as information on the opportunity cost of time of the household member who made the trip.[42] Treatment increased these transaction costs by a small but significant amount: Rs. 7, or around 1.5% of value received (Table 5, Column 1). In Table A.6 we examine impacts on the variables that feed into our total cost measure; the cost increase appears to be due to (i) a significant increase in the number of trips that were unsuccessful in the sense that they did not result in any purchases, which more than doubled from 0.1 per household per month to 0.23, and (ii) an increase in the opportunity cost of time of the household member who collected benefits, consistent with the idea that the reform reduced households' flexibility to send whoever could be spared from other work.

4.3 Food security

Given the results above, we would not expect to see impacts on food security outcomes.[43] Table 6 confirms this. We examine two measures of a household's food security: a food consumption score that follows standard World Food Program methodology to calculate a nutrient-weighted sum of the number of times a household consumed items from each of a set of food groups in the last week, and a simple food diversity score defined as the number of groups from which the household consumed any items in the past week.[44] We see a tightly estimated null effect of treatment, with

[42]In our survey we asked both for the number of unsuccessful trips made by each individual on the household roster and for the total number of unsuccessful trips taken. When the latter exceeds the sum of the former we attribute the stated total number of trips to household members in proportion to their stated individual number of trips. The results are not sensitive to alternative approaches, e.g. simply using individuals' trip counts.

[43]It is possible that there could have been indirect effects on food security if the incentive to register for Aadhaar created by treatment affected households' access to *other*, non-PDS benefits which also required Aadhaar. At the time of our follow-up Aadhaar was being used in some form to control access to wage payments under the National Rural Employment Guarantee Scheme, pension payments, and conditional cash transfers to mothers under the Pradhan Mantri Matru Vandana Yojana. Note however that any effects through access to these benefits would likely be small given that the difference in Aadhaar registration rates between treatment (96%) and control (92%) was only 4%.

[44]For more details on these methods including the weights for each food group, which are defined based on the group's nutrient density, see `http://documents.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp197216.pdf?_ga=1.115126021.300736218.1470519489`

95% confidence intervals expressed in control group standard deviations of $[-0.11\sigma, 0.12\sigma]$ and $[-0.11\sigma, 0.09\sigma]$ respectively.

4.4 Distributional and heterogeneous effects

The notion that stricter ID requirements should trade of reductions in errors of inclusion (including, broadly defined, leakage) against increases in errors of exclusion itself suggests that effects are likely to be heterogeneous: many may be unaffected, or affected only to the extent that transaction costs change, while the main risk is that some lose access to most or all of their benefits.

The distributional effects of treatment suggest this was the case. Figure 3 plots the CDFs of value received in the treatment and control groups separately; these track each other closely except for values close to zero, where there is more mass in the treatment group. The probability that a treated household received zero value is 2.4 percentage points higher than a control household (Table 7, Column 1), significant at the 10% level.

For a sharper test, we examine how impacts differed for the household we would expect to be most likely to lose access to benefits, namely the 23% of households that were "unseeded" at baseline in the sense that they did not have at least one member whose Aadhaar number had been seeded to their ration card. Figure A.2 plots the distributions of household income and mean years of schooling completed for the two most educated household members, separately by seeded status. Unseeded households tend to have lower values on both metrics, and in both cases a Kolmogorov-Smirnov test rejects equality of distributions. Unseeded households are also 5% less likely be upper caste ($p < 0.01$).

Losses in value received are concentrated among unseeded households. Table 8 reports estimated treatment effects split by this variable. The reform lowered value received by Rs. 49 per month for unseeded households, equivalent to 12.6% of the control group mean for this category. This is significantly different from zero as well as from the mean effect among seeded households. On the extensive margin, treatment lowered the probability that unseeded households received any benefit by 10 percentage points, also significantly higher than the (insignificant) impact on seeded ones. While we cannot of course identify specific households that counterfactually would not have been excluded, this decrease fully accounts in an arithmetic sense for the overall decrease in the fraction of households reporting receiving any benefits in a given month. Treatment effects on stated willingness to accept are also significantly lower for unseeded households, though not in this case significantly different from zero. Transaction costs, on the other hand, increase the most for seeded households, consistent with the idea that they are the ones able to continue transacting with the system, albeit at a higher cost (and that unseeded households may not have bothered making multiple trips). Overall, this suggests that the reform did cause a significant reduction in value received for the households least ready for the reform, likely driven by the total loss of benefits of a subset of these households.

We also examine heterogeneity along several additional pre-specified dimensions, including (i) characteristics likely to matter for understanding the distributional and political consequences of the reform such as caste, education level, and income level, and (ii) characteristics of the location likely to predict heterogeneity in the implementation of the reform such as rural status, cellular network signal strength, and the device mode (online, partially online, or offline).[45] Note that to examine heterogeneity by machine mode we need to infer the mode that a given control FPS would have been using had it been treated at the time. We do so by assuming that it would have used the same machine type it was ultimately assigned to use once the control blocks were treated.[46]

In general we find limited evidence of heterogeneity along these dimensions (Tables A.7, A.8, and A.9.). There is some evidence that wealthier and better-educated households receive differentially more value and that wealthier households incur larger increases in transaction costs.

4.5 Stakeholder preferences and perceptions

We next examine beneficiaries' and dealers' stated preferences for the reform relative to the status quo. These preferences provide both a cross-check on any inferences about well-being we draw from the experimental estimates, and insight into the longer-term political viability of the reform.

Overall, views on the reform were sharply divided (Table 9). Fifty-three percent of households and 51% of dealers preferred the reform to the status quo method of authenticating. Even among unseeded households, which were most hurt as a group, 50% of households prefer the reform to the status quo. Views are quite polarized, with 89% (87%) of households (dealers) holding a strong as opposed to a weak view one way or the other. One interpretation is that respondents view the question as being as much about their political allegiance as about the direct effect of the policy on their wallets. Indeed, respondents often framed their responses in broader political terms (for example, "I prefer the machine because finally the government is doing something about the PDS").

We also asked dealers their expectations about future business prospects (Table 10). Roughly the same share of treated dealers expect to continue running their FPS (Columns 1-2), but they predict that the going price to obtain a dealer's licenses in the first place will drop substantially, by 74% (Columns 3-4). (Expected payments to renew a license turn out to be negligible and unaffected by treatment (Columns 5-6)). We interpret this result cautiously given that it is a sensitive question and only a minority of dealers provided an answer. That said, it is intriguing that those who did answer expect the price of licenses to fall substantially, despite the fact that the reform requiring Aadhaar-based authentication to avail PDS by itself had not affected leakage. One possibility is that dealers in treated blocks anticipated that the government would soon begin using authenticated transaction data to reconcile commodity balances and that this would meaningfully reduce their ability to divert grain onto the open market. We turn to this next.

[45] We observe the mode in which machines were operated at the level of the individual transaction. To create a machine × month level series, we assign each observation the modal transaction type conducted by that machine and during that month. On average the modal transaction type accounts for 99% of the transactions in a given machine × month cell, suggesting this is a reasonable approximation.

[46] This assumption generally appears reasonable in the sense that the distribution of machine types 6-8 months after the reform was implemented appear similar in both treatment and control areas, with the one exception that the government ended the use of partially online mode in August 2017 and so we impute fewer partially online machines in control.

5 Reconciliation & ration card deletion

5.1 Policy implementation

Anecdotally, reconciliation was only partially implemented. FPS dealers pressed for and were often granted adjustments or exceptions to offset its effects. Had they not, the consequences for grain disbursement could have been dramatic: by our calculations, for example, 32% of FPS in treated blocks had reported opening balances worth one month or more by June of 2017 and should therefore have received no disbursement of wheat or rice in the next month had the policy been strictly implemented. Overall, we estimate that foodgrain distributions would have fallen by 43% in July had reconciliation been strictly implemented. In Appendix C we examine in more detail how strictly the government implemented its stated reconciliation formulae. Generally speaking the policy had bite, but the government clearly made numerous exceptions and overall tended to disburse grain somewhat more leniently than strict "by the book" implementation would have implied. The results that follow should thus be interpreted as the net effect of the rules as actually implemented, net of various adjustment and exceptions which were made.

5.2 Effects of reconciliation on value transfer

Overall, the onset of reconciliation coincided with a drop in both disbursements and receipts of reconciled commodities (rice and wheat), but not of unreconciled ones (sugar, salt and kerosene). Figure 4 illustrates this, plotting the evolution of value disbursed (Panel (a)) and received (Panel (b)) separately for for reconciled and unreconciled commodities. It also overlays the raw data with the fit we obtain from estimating Equation 2 and 95% confidence bands around this fit.[47] For reconciled commodities, both value disbursed and received show little change until the onset of reconciliation, after which both drop sharply. They then rebound gradually until, by October, they return to roughly the levels we would have expected based on pre-reconciliation trends. For both series our pre-specified functional form fits the temporal patterns quite well.

For unreconciled commodities – arguably a control or placebo group, though of course it is possible that policy changes affecting the reconciled commodities could have spillover effects onto the unreconciled ones – both value disbursed and received drift slightly downwards over time without any substantial change during the period of reconciliation. If anything, the value of unreconciled commodities received during this period was high relative to trend.

In Table 11 we quantify these effects, splitting the data by treatment arm.[48] We begin with control blocks, which had received ABBA shortly before the onset of reconciliation itself. We estimate that value disbursed to control blocks fell by Rs. 92, or 19%, with the onset of reconciliation. Of this, an estimated 22% represents a reduction in value received by legitimate beneficiaries (Panel B), while the remaining 78% represents a reduction in leakage (Panel C). As a proportion of January

[47]We value commodities using market prices obtained from follow-up 1, as follow-ups 2 and 3 did not elicit updated market price data. The evolution of value metrics thus reflects the evolution of quantities.

[48]Figure A.3 provides the corresponding plots split by treatment arm.

2017 means, these figures represent a 5% reduction in value received by legitimate beneficiaries and a 92% reduction in leakage, respectively.

Overall, these results highlight the tradeoff between errors of inclusion and of exclusion. This tradeoff in turn like reflects a mixture of two offsetting effects of reconciliation. Recall that from the point of view of dealers the reconciliation reform had both an incentive effect (increasing their incentive to record authenticated transactions, and thus potentially increasing beneficiaries' bargaining power) and a resource constraint effect (reducing available grains to distribute to the extent the dealer had accumulated arrears). The former effect should tend to help beneficiaries, while the latter seems more likely to hurt them. If this logic is correct, it suggests that reconciliation might have impacted beneficiaries less negatively had it been introduced on a "clean slate" basis, ignoring previously accumulated arrears and simply enforcing the rules on a go-forward basis.

The prior randomization of Aadhaar-based biometric authentication gives us one way to examine this hypothesis. Because dealers treated in that experiment had ePOS devices for longer than those in control blocks (11 months vs 2), they also accumulated greater "opening balances" (corresponding to undisbursed grain received earlier) by the beginning of July 2017. Figure 5 plots the distribution of total grain stocks at end of June for the treatment and control group separately, illustrating the difference. On average, the government held treated shops responsible for 7,815kg of grain as opposed to 2,481kg for control shops ($p < 0.0001$). As a result, we expect to see the resource constraint effect concentrated in treated blocks, while the incentive effect applies equally everywhere.

Consistent with this intuition, we estimate that reconciliation had larger effects overall in treated areas and generated a less advantageous tradeoff between errors of exclusion and inclusion. In treated areas, reconciliation coincided with a Rs. 182 drop in value disbursed, of which Rs. 62 represents reduced value received by beneficiaries and Rs. 121 represents reduced leakage. In each case these figures are significantly larger than the corresponding figures for control areas. The *proportion* of the drop in value disbursed that represents exclusion error is higher in treated than in control areas, by 34% to 22%, (though statistically we do not reject the null that these proportions are the same, $p = 0.18$).

Pushing this logic further, we can use our experimental variation to extrapolate to predict what might have happened had the government counterfactually introduced a fully clean-slate reconciliation policy. To do so we estimate Equation 2 fully interacted with the fps-level stock balance as of the beginning of July 2017, expressed per ration card (the "opening balance"). Since opening balances are likely endogenous, but vary systematically with assignment to early v.s. late treatment, we instrument for balance (and its interactions) in this specification with assignment to treatment (and its interactions). This instrument is valid if assignment to treatment altered the effects of reconciliation only through its effects on opening balances. We then interpret the point estimates on the main effect of reconciliation, with opening balances set to zero, as estimates of the initial effects of a clean slate reconciliation.

The estimates from this procedure suggest that a clean-slate reconciliation would have weakly *increased* value received by beneficiaries (Table A.10). Column 4 examines effects on value received; setting the opening balance to zero, reconciliation would have increased value received (insignificantly) by Rs. 7.3. The interaction of reconciliation with opening balance is significantly negative, implying that it was the opening balances for which dealers were held accountable rather than the reconciliation policy per se which led to the drop in benefits received.

It is less clear how a clean slate reconciliation would have affected value disbursed. If implemented strictly it should have no initial impact, precisely because at the onset of reconciliation dealers would not have been in arrears. In practice, the government implemented a somewhat loose interpretation of its reconciliation formulae (Appendix C), and our data suggest that doing so would initially have reduced value disbursed by Rs. 59 per ration card (Column 1), less than the drop we see in our data but still greater than zero. What is notable is that the data suggest the government could have achieved this reduction without adversely affecting beneficiaries.

This exercise should be interpreted with some degree of caution. It necessarily involves extrapolation: the mean opening balance in the control group, while only 32% of that in the treatment group, was not zero, so that we do not directly observe the consequences of reconciliation at a zero balance. It estimates how the initial effects of a clean slate reconciliation would have differed, but does not as readily forecast how outcomes would subsequently have evolved given that these would in turn depend dynamically on counterfactual decisions by the dealers that we do not observe. That said, the results are consistent with the basic intuition that the incentive effects of reconciliation benefit recipients, while the resource constraint effects (which are magnified when dealers are held responsible for cumulative undisbursed stocks of grain) may not.

In practice, the GoJH ultimately suspended reconciliation starting in November 2017 in the face of negative press coverage and complaints from both beneficiaries and FPS dealers. It reintroduced reconciliation starting in August 2018 with two modifications: its stated policy is now to adjust disbursements to reflect estimated stock as of four months previously (i.e. disbursements in August were based on estimated stock as of April) while also conducting physical verification of stock.

5.3 Ration card deletion

We close by examining the deletion of ration cards, another margin along which the government traded off errors of inclusion and exclusion during the rollout of Aadhaar. Some card deletion was normal, due for example to the death of beneficiaries, dissolution of beneficiary households, or discovery of fraud. From 2014-2016, for example, the GoJH reported deleting 453,958 ration cards, or around 2% of the total,[49] From October 2016 to May 2018 we observe an additional 6% of ration cards removed from the database for the 10 districts we study, suggesting that the overall rate of deletion was higher after the rollout of ABBA in the PDS.

We examine this more closely in Table 12. In the top panel we examine the universe of ration cards in our 10 study districts. We define a card as deleted if it appeared in the list we obtained

[49]Data from `https://data.gov.in/resources/stateut-and-year-wise-number-deleted-ration-cards-2014-2018-ministry-consumer-affairs-food`, accessed 12 July 2019.

in October 2016, but not the list we obtained in May 2018. We define a card as unseeded if no household member had seeded an Aadhaar number to that card as of October 2016. We observe that unseeded cards were substantially more likely to be deleted during this period, with 36% of unseeded cards ($80,085/213,089$) deleted compared to 2% ($64,076/2,236,521$) of seeded cards.

Two key questions are whether and how effectively this removal of unseeded cards differentially eliminated fraudulent accounts. In the bottom panel we restrict the analysis to our sampled ration cards, for which we can categorize beneficiaries as ghosts or not based on the results of our survey. We estimate that unseeded cards are disproportionately likely to belong to ghosts: they make up 9.5% (21/232) of unseeded cards, as opposed to just 0.8% (32/3616) of seeded cards ($p < 0.001$ from a two-way Pearson χ^2 test). Deleting unseeded cards does thus differentially target ghost accounts. However, because ghosts make up such a small share of the overall population, we estimate that the great majority (88%, or $(90 + 97)/213$) of the deleted cards, and even of the deleted unseeded cards (84%, or $15/(15 + 90)$), belonged to real households.

To be clear, our data do not allow us to estimate a causal effect of ABBA on card deletions. Further, deletions (even of real households) do not by themselves imply exclusion of deserving beneficiaries. For example, a card might remain unseeded because a (relatively) wealthy holder did not think it worth the trouble to seed it, and deleting this card would be desirable.[50] At the same time, the discussion above shows using data from our field surveys that a considerable number of deleted cards corresponded to real households, and thus illustrates another way in which attempts to reduce leakage may come at the cost of exclusion error of genuine households.

6 Conclusion

Navigating the tradeoffs between errors of inclusion and exclusion in social programming is a common challenge facing governments worldwide. This is especially so in developing countries such as India where state capacity is limited and lives may literally be at stake. In this paper we examine the effect on inclusion and exclusion error of stricter authentication requirements, and in particular authentication using Aadhaar, the world's largest ID system and largest biometric ID system. We examine the effects of an experimental at-scale rollout of Aadhaar authentication in the Public Distribution System, India's largest safety net program, as well as a subsequent reform to the reconciliation protocols that leveraged the authenticated transaction data this generated.

On its own, Aadhaar-based authentication of transactions had no measurable benefit; it slightly increased mean transaction costs for beneficiaries, excluded from their benefits a minority who did not have ID's "seeded" to their ration cards at baseline, and did not reduce leakage. When paired with the new reconciliation protocols, authentication appears to have substantially cut leakage but at the cost of concurrent reductions in mean value received by legitimate beneficiaries. These results illustrate how the costs of controlling corruption may include indirect "collateral damage" beyond the direct costs of intervening. Juxtaposed with prior evidence, they also demonstrate the

[50]In our data, cards held by the better-off among the unseeded are deleted somewhat more often, but this pattern is not statistically significant.

range of impacts that "biometric authentication" can have depending on the specific ways it is used. Together, our results suggest that "biometric authentication" is not per se a particularly stable or helpful construct for policy analysis: the design details of how the technology is used and how it alleviates (or aggravates) existing constraints matter a lot.

The comparison of these results with our own prior work on the impacts of biometric authentication in a public employment program and a pension program in the state of Andhra Pradesh (AP), is especially illustrative. A key point to note is that *both* programs reduced leakage. However, in the case of AP, the reduced leakage was passed on to the beneficiaries in terms of more money received (displaced from corrupt intermediaries), while there were no savings to government. In contrast, in the case of Jharkhand, the reduced leakage in the PDS led to reduced disbursals from the government, but did not improve the beneficiary experience in any way (and worsened it in some ways). In other words, the technology of biometric authentication "worked" in both settings in terms of reducing leakage. But the question of how the benefits of this leakage reduction should be shared between the government and beneficiaries is ultimately a design question and also a political one. Thus, the biggest reason for the difference in results (in our assessment) was not because of the technology (Smartcards vs. Aadhaar) or the context (AP vs. Jharkhand) but because of differences in program design. AP emphasized the beneficiary experience, whereas Jharkhand (implementing the policy decision of the Government of India) emphasized fiscal savings - and the results are consistent with this difference in emphasis.

The final lesson to note is the importance of data on beneficiary experiences for policy. In the AP Smartcards case, it would have been easy to think that there was no impact on leakage because there was no change in government expenditure on NREGS and pensions (Muralidharan et al., 2016). It was only with the matched data between administrative records and household surveys (and the existence of a control group) that we could see that leakage had fallen sharply and that more benefits were reaching people. Conversely, in the case of the PDS in Jharkhand, it would have been easy to interpret the reduction in disbursals as evidence of reduced leakage (and indeed, officials often made this claim). However, it was the matched data using household surveys that clarified that at least some of the reductions in disbursal were coming at the cost of exclusion errors.

Our results suggest that it is important to both build in procedures to guard against exclusion error at the program design phase, as well as monitor this during the implementation phase. Some design choices that may improve the terms of the trade-off between leakage reduction and exclusion errors include: (i) authenticating beneficiary lists (say annually) rather than transactions,[51] (ii) creating alternative methods of authentication or override mechanisms,[52] and (iii) conducting

[51]Gelb and Metz (2018) similarly argue that "point-of-service biometric authentication for every service can... be overkill" and that "programs also need to allow for people who cannot register or authenticate themselves in the regular way..." On the other hand, authenticating every interaction may be valuable in cases where physical attendance is an important margin. Bossuroy et al. (2019) find for example that requiring (non-Aadhaar) biometric authentication in health clinics increased adherence, and reduced over-reporting of adherence, to a tuberculosis treatment regimen. It may also be important for the portability of benefits, whereby beneficiaries (especially internal migrant workers) can access their benefits outside their home location.

[52]In a survey in Krishna district of Andhra Pradesh, for example, Aadil et al. (2019) find few reports of exclusion from PDS benefits due to Aadhaar authentication in Krishna district in AP, which they attribute to the availability

reconciliation on a clean-slate basis. Our estimates suggest that the last option might have been effective at reducing leakage without increasing exclusion errors.[53]

However, regardless of design safeguards, policy making can also benefit from improving real-time visibility into last-mile beneficiary experience with welfare programs to enable rapid course correction of policies that may be hurting vulnerable populations. One promising way of doing this may be to use outbound call centers to call representative samples of beneficiaries regularly and measure if they are receiving their benefits. Recent evidence suggests that such an approach may be a scalable way of measuring and improving last-mile service delivery (Muralidharan et al., 2019b).

of several override mechanisms.

[53]In conversations with senior government officials, we have been told that the Department of Food is now following such a protocol.

References

Aadil, Arshi, Alan Gelb, Anurodh Giri, Anit Mukherjee, Kyle Navis, and Mitul Thapliyal, "Digital Governance: Is Krishna a Glimpse of the Future?," Working Paper 512, Center for Global Development June 2019.

Abraham, Ronald, Elizabeth S. Bennett, Noopur Sen, and Neil Buddy Shah, "State of Aadhaar Report 2016-17," Technical Report, IDinsight 2017.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, and Julia Tobias, "Targeting the Poor: Evidence from a Field Experiment in Indonesia," *American Economic Review*, June 2012, *102* (4), 1206–40.

_ , _ , _ , _ , Ririn Purnamasari, and Matthew Wai-Poi, "Self-Targeting: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy*, April 2016, *124* (2), 371–427.

Allcott, Hunt, "Site Selection Bias in Program Evaluation," *The Quarterly Journal of Economics*, March 2015, *130* (3), 1117–1165.

Bossuroy, Thomas, Clara Delavallade, and Vincent Pons, "Biometric Tracking, Healthcare Provision, and Data Quality: Experimental Evidence from Tuberculosis Control," Working Paper 26388, National Bureau of Economic Research October 2019.

Cantoni, Enrico and Vincent Pons, "Strict ID Laws Don't Stop Voters: Evidence from a U.S. Nationwide Panel, 2008-2016," Working Paper 25522, National Bureau of Economic Research February 2019.

Commission, Planning, "Report of the expert group to review the methodology for measurement of poverty," Technical Report June 2014.

Cunha, Jesse M, Giacomo De Giorgi, and Seema Jayachandran, "The Price Effects of Cash Versus In-Kind Transfers," *The Review of Economic Studies*, 04 2018, *86* (1), 240–281.

Currie, Janet, "The Take Up of Social Benefits," Techreport, National Bureau of Economic Research May 2004.

Dreze, Jean and Reetika Khera, "Understanding Leakages in the Public Distribution System," *Economic and Political Weekly*, February 2015, *50* (7).

_ , Nazar Khalid, Reetika Khera, and Anmol Somanchi, "Pain without Gain? Aadhaar and Food Security in Jharkhand," *Economic and Political Weekly*, December 2017, *52* (50).

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan, "Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India," *The Quarterly Journal of Economics*, September 2013, *128* (4), 1499–1545.

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael W Walker, "General Equilibrium Effects of Cash Transfers: Experimental Evidence from Kenya," Working Paper 26600, National Bureau of Economic Research December 2019.

FAO, IFAD, UNICEF, WFP, and WHO, "The State of Food Security and Nutrition in the World 2019. Safeguarding against economic slowdowns and downturns," Technical Report, FAO 2019.

Gelb, Alan and Anna Diofasi Metz, *Identification Revolution: Can Digital ID Be Harnessed* *for Development?*, Center for Global Development, January 2018.

Grebmer, Klaus V., Jill Bernstein, Doris Wiesmann, and Hans Konrad Biesalski, "2019 Global Hunger Index: The Challenge of Hunger and Climate Change," 2019.

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

Huntington, Samuel P., *Political Order in Changing Societies*, Yale University Press, 1968.

Khera, Reetika, "Opinion: On Aadhaar Success, It's All Hype - That Includes The World Bank.," July 2016.

_ , "Impact of Aadhaar on Welfare Programmes," *Economic and Political Weekly*, Dec 2017, *52* (50).

Kleven, Henrik Jacobsen and Wojciech Kopczuk, "Transfer Program Complexity and the Take-Up of Social Benefits," *American Economic Journal: Economic Policy*, February 2011, *3* (1), 54–90.

Leff, Nathaniel H., "Economic Development Through Bureaucratic Corruption," *American Behavioral Scientist*, November 1964, *8* (3), 8–14.

Lichand, Guilherme and Gustavo Fernandes, "The Dark Side of the Contract: Do Government Audits Reduce Corruption in the Presence of Displacement by Vendors?," Working Paper, University of Zurich April 2019.

Mundle, Sudipto, Samik Chowdhury, and Satadru Sikdar, "The Quality of Governance: How Have Indian States Performed?," *Economic and Political Weekly*, December 2012, *47* (49).

Muralidharan, Karthik, Abhijeet Singh, and Alejandro J. Ganimian, "Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India," *American Economic Review*, April 2019, *109* (4), 1426–60.

_ **and Paul Niehaus**, "Experimentation at Scale," *Journal of Economic Perspectives*, 2017, *31* (4), 103–124.

_ **, Jishnu Das, Alaka Holla, and Aakash Mohpal**, "The Fiscal Cost of Weak Governance: Evidence From Teacher Absence in India," *Journal of Public Economics*, January 2017, *145*, 116–135.

_ **, Paul Niehaus, and Sandip Sukhtankar**, "Building State Capacity: Evidence from Biometric Smartcards in India," *American Economic Review*, 2016, *106* (10), 2895–2929.

_ **,** _ **,** _ **, and Jeffrey Weaver**, "Improving Last-Mile Service Delivery using Phone-Based Monitoring," Working Paper 25298, National Bureau of Economic Research July 2019.

Niehaus, Paul and Sandip Sukhtankar, "Corruption dynamics: The golden goose effect," *American Economic Journal: Economic Policy*, 2013, *5* (4), 230–69.

_ **, Antonia Atanassova, Marianne Bertrand, and Sendhil Mullainathan**, "Targeting with Agents," *American Economic Journal: Economic Policy*, February 2013, *5* (1), 206–38.

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

Organisation, Planning Commission The Programme Evaluation, "Performance Evaluation of Targeted Public Distribution System (TPDS)," Technical Report March 2005.

Scott, James C., *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*, Yale University Press, 1998.

Shleifer, Andrei and Robert W. Vishny, "Corruption," *Quarterly Journal of Economics*, 1993, *108* (3), 599–617.

UNICEF et al., "WHO, The World Bank. Joint child malnutrition estimates—Levels and trends (2017 edition). Geneva: World Health Organization," 2017.

Vivalt, Eva, "How Much Can We Generalize from Impact Evaluations?," *Journal of the European Economics Association.*, forthcoming.

Wilson, James Q., *Bureaucracy: What Government Agencies Do and Why They Do It*, Basic Books, 1989.

Table 1: Representativeness within Jharkhand

	Study district	Non-study district	Difference	p-value
	(1)	(2)	(3)	(4)
Panel A: 2011 Census				
Population in 2011	1,267,604	1,450,864	-183,260	0.50
Population growth, 2001-2011	0.23	0.24	-0.02	0.56
Population density	451	459	-8	0.94
% Literate	0.62	0.66	-0.04	0.22
Panel B: Beneficiary List				
Beneficiaries per FPS	308	293	15	0.45
% FPS rural	0.92	0.89	0.03	0.39
% AAY beneficiares	0.18	0.18	0	0.92
Number of blocks	13.20	12	1.20	0.58
Panel C: NSS 68				
% With salary income	0.11	0.16	-0.04	0.25
Monthly per capita consumption	1,097	1,298	-201	0.10
Consumption value food	4,050	3,518	532	0.34
Consumption value fuel/light	506	462	44	0.16
N	10	14		

This table compares the 10 districts studied with the remaining 14 districts in Jharkhand using data from the 2011 and 2001 Censuses (Panel A), the PDS beneficiary list prior to baseline (Panel B), and the 68th Round of the National Sample Survey (NSS 68) (Panel C). Column 3 reports the raw difference in means between columns 1 and 2. Column 4 reports the p-value from a test of equality of means. "Population density" is in population per square mile. "Beneficiaries per FPS" is the ratio of PDS beneficiaries to the number of Fair Price Shops. "% FPS rural" is the share of Fair Price Shops located in areas classified as rural. "% AAY beneficiaries" is the percentage of PDS beneficiaries covered by the more generous Antyodaya Anna Yojana (AAY) scheme. "% With salary income" is the share of the population that reports earning a salaried income. "Monthly per capita consumption" is household monthly per capita consumption in Rs.. "Consumption value food" is household monthly expenditure on food in Rs. "Consumption value fuel/light" is the household monthly expenditure on fuel and lighting in Rs. Statistical significance is denoted as: $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table 2: Baseline balance and program implementation

	Treatment	Control	Regression-adjusted difference	p-value
	(1)	(2)	(3)	(4)
Panel A: Baseline Characteristics				
Priority households	13080	12292	345	.82
AAY households	2922	2576	125	.68
Aadhaar numbers seeded per rationcard	2.4	2.4	.046	.58
Rice disbursed per priority household	23	20	7.5	.13
Rice disbursed per AAY household	35	32	11	.27
Number of FPS	73	71	-2.1	.8
Median household size	4.4	4.3	.069	.42
% of rationcard holders identified via SECC	.71	.68	.023	.27
% of rationcard holders identified by application	.16	.16	-.0014	.93
% of rationcard holders without eligibility info	.13	.16	-.022	.22
Whether at least one Aadhaar seeded	.77	.8	-.025*	.082
Missing whether any Aaadhaar seeded	.096	.16	-.024	.19
Panel B: Program implementation				
Dealer has an ePOS machine at endline survey	.96	.05	.91***	0.00
Dealer used ePOS in January 2017	.91	.06	.85***	0.00
Dealer used ePOS in February 2017	.91	.06	.85***	0.00
Dealer used ePOS in March 2017	.91	.06	.86***	0.00

This table compares treatment to control blocks within study districts on baseline characteristics (Panel A) which should be balanced due to randomization, and measures of program implementation (Panel B) which should not. Column 3 reports the regression-adjusted difference in means after conditioning on strata fixed effects, and column 4 reports the p-value from a test that this quantity equals zero. "Priority households" is the number of ration cards assigned to households under the priority households scheme; "AAY households" is the number of ration cards assigned to households under the Antyodaya Anna Yojana (AAY) scheme. "Aadhaar numbers seeded per ration card" is the average number of verified Aadhaar numbers seeded per ration card. "Rice disbursed per priority household" is kilograms of rice disbursed per PHH ration card. "Rice disbursed per AAY household" is kilograms of rice disbursed per AAY ration card. "Number of FPS" is the total number of Fair Price Shops. "Median household size" is the block median number of household members listed on ration cards. "% of ration card holders identified via SECC" is the share of ration card holders whose eligibility was established using data from the Socio Economic Caste Census. "% of ration card holders identified by application' is the share of ration card holders whose eligibility was determined by local authorities after submitting applications. "% of ration card holders without eligibility info" is the share of ration card holders for which we do not observe how they became eligible. "At least one Aadhaar number seeded" is an indicator equal to one if the household had at least one Aadhaar number seeded to its ration card at baseline. "Missing Aadhaar seeding status" is an indicator equal to one if we do not observe the count of Aadhaar numbers seeded to the ration card at baseline. Estimates in Panel B are weighted by inverse sampling probabilities. "Dealer has an ePOS machine at endline" is an indicator equal to one for endline survey respondents who reported that their FPS dealer had an ePOS machine. "Dealer used an ePOS machine in Month X 2017" is an indicator equal to one for endline survey respondents who reported that their FPS dealer used or attempted to use an ePOS machine in the corresponding month. Statistical significance is denoted as: $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table 3: Effects on value disbursed, value received, and leakage

	Total	Rice	Wheat	Sugar	Salt	Kerosene
	(1)	(2)	(3)	(4)	(5)	(6)
Mean entitlement	595	493	17	19	9	57
Panel A: Value disbursed						
Treatment	12**	35***	-27**	.093	.028	-.25
	(4.9)	(12)	(12)	(.15)	(.045)	(.18)
		[0.05]	[0.16]	[1.00]	[1.00]	[0.68]
Control mean	584	421	75	25	9.2	54
Observations	26,611	26,611	26,611	26,611	26,611	26,611
% of frame	99	99	99	99	99	99
Panel B: Value received						
Treatment	-1.2	17	-15	.55	.51	-.56
	(12)	(10)	(11)	(1.6)	(.58)	(1.1)
		[0.96]	[0.96]	[1.00]	[1.00]	[1.00]
Control mean	463	348	54	14	7.2	40
Observations	10,396	10,557	10,654	10,670	10,726	10,618
% of sample	88	89	90	90	90	89
Panel C: Leakage						
Treatment	14	18	-11	-.46	-.48	.56
	(12)	(12)	(7.2)	(1.6)	(.56)	(1.1)
		[0.72]	[0.72]	[1.00]	[1.00]	[1.00]
Control mean	109	63	16	11	2	17
Observations	10,396	10,557	10,654	10,670	10,726	10,618
% of sample	88	89	90	90	90	89

This table reports estimated treatment effects on the value of commodities disbursed by the government (Panel A), received by recipients (Panel B), and the difference (Panel C) in endline one (January - March). The unit of measurement is rupees per ration card-month throughout. In Panel A the unit of observation is Fair Price Shop × month and we use the universe of Fair Price Shops; in Panels B and C the observation is the ration card × month and a representative sample of ration card holders in Panels B and C. The dependent variable in columns 2-6 is the relevant quantity of the commodity multiplied by the difference between the median market price of that commodity in control blocks in the same district, and the statutory PDS price for that commodity. The dependent variables in column 1 is the sum of the values in columns 2-6. In Panel C, estimated effects are the difference between estimated effects on block-level mean value disbursed per ration card and value received per ration card, estimated within a Seemingly Unrelated Regression framework. All specifications include strata fixed effects and the baseline value of the dependent variable. Standard errors clustered at the block level are reported in parentheses, with statistical significance denoted as: $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$. q values adjusted to control the false discovery rate across the five commodities are reported in brackets.

Table 4: Effects on quality of ration received

	Rice	Wheat	Sugar	Salt	Kerosene
	(1)	(2)	(3)	(4)	(5)
Panel A: Adulteration					
Treatment	.0097	.0074	.01	-.0036	-.0019
	(.012)	(.027)	(.0079)	(.0076)	(.003)
	[0.88]	[1.00]	[0.88]	[1.00]	[1.00]
Control mean	.087	.068	.028	.033	.0095
Observations	3,308	971	2,305	2,381	3,205
% of sample	84	25	58	60	81
Panel B: Low quality					
Treatment	-.0014	.045	.014	-.058**	-.0008
	(.0068)	(.031)	(.0088)	(.029)	(.0026)
	[1.00]	[0.58]	[0.58]	[0.52]	[1.00]
Control mean	.036	.069	.056	.38	.0076
Observations	3,329	975	2,319	2,402	3,228
% of sample	84	25	59	61	82

This table reports estimated treatment effects on the quality of commodities received by beneficiaries in endline one (January - March). The unit of analysis is the ration card in both panels. The dependent variable in Panel A is an indicator equal to one if the respondent reported receiving adulterated commodities at least once in the past three months for each of the five commodities. The dependent variable in Panel B is an indicator equal to one if the respondent reported that the overall quality of commodities received over the past three months was "very bad" or "bad" (as opposed to "OK" or "good') for each of the five commodities. In both panels, observation of the outcome is conditional on the ration card holder purchasing a positive quantity of the commodity during January-March 2017. All regressions include strata fixed effects and the baseline value of the dependent variable. Standard errors clustered at the block level are reported in parentheses, with statistical significance denoted as: $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$. q values adjusted to control the false discovery rate across the five commodities are reported in brackets.

Table 5: Effects on transaction costs

	Beneficiary costs	Dealer costs		Government costs
	(1)	(2)	(3)	(4)
Treatment	6.9*	.51	.65	–
	(3.8)	(.95)	(.63)	–
Adjusted R^2	.09	.10	.28	–
Reference group mean	41	6.8	5.9	127
Observations	3,538	441	367	–
% of sample	89	–	93	–
Sample	–	Full	Restricted	–

This table reports estimated treatment effects on measures of transaction costs incurred transferring PDS commodities. In column 1 the unit of analysis is the ration card and the dependent variable is the total cost incurred in March by the household holding that ration card in purchasing or attempting to purchase PDS commodities, including time and money costs (see text for details). In columns 2 and 3 the unit of analysis is the Fair Price Shop and the dependent variable is the total cost incurred by the dealer to transport and store PDS commodities in an average month in January - March divided by the number of ration cards assigned to that dealer. In column 2 the sample includes all dealers surveyed, including those to whom sampled households switched between baseline and endline; in column 3 it includes only dealers drawn in the original sample. Column 4 reports the mean administrative cost per ration card $\times$ month incurred by the state government to administer the PDS, but does not report an estimated treatment effect as cost data disaggregated by block are not available. All specifications include strata fixed effects, and regressions in columns 1-3 include the baseline value of the outcome variable. Standard errors clustered at the block level are reported in parentheses with statistical significance denoted as: $^{*}p < .10,^{**}p < .05,^{***}p < .01$.

Table 6: Effects on food security

	Dietary diversity score	Food consumption score
	(1)	(2)
Treatment	-.011	.08
	(.061)	(1)
Adjusted R^2	.05	.10
Control mean	5.7	43
Observations	3,578	3,578
% of sample	90	90

This table reports estimated treatment effects on measures of food security in March. The unit of observation is the ration card. The dependent variable in column 1 is the sum of a series of indicators each equal to one if the household has consumed any items from within a major food group during the previous week. The dependent variable in column 2 is a weighted sum of the number of times the household consumed items from each major food group in the past week, with weights based on the group's nutrient density.The major food groups are: main staples, pulses, vegetables, fruit, meat and fish, milk, sugar, oil, and condiments. The definition of food groups and their weights can be found from the World Food Programme. All regressions include strata fixed effects. Standard errors clustered at the block level are reported in parentheses. Statistical significance is denoted as: $^{*}p < .10,^{**}p < .05,^{***}p < .01$.

Table 7: Effects on the extensive margin of value received

	Any Commodity	Rice	Wheat	Sugar	Salt	Kerosene
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-.024*	-.025	-.064	-.011	.0049	-.013
	(.014)	(.016)	(.047)	(.019)	(.018)	(.018)
		[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
Adjusted R^2	0.10	0.10	0.32	0.05	0.04	0.10
Control mean	.85	.83	.28	.28	.29	.75
Observations	10,396	10,557	10,654	10,670	10,726	10,618
% of sample	88	89	90	90	90	89

This table reports estimated treatment effect on the extensive margin of the values received by beneficiaries per month in endline one (January - March). The unit of analysis is the ration card-month. The dependent variable in columns 2-6 is an indicator equal to one if the ration card holder received a positive quantity of the commodity in a given month. The dependent variable in column 1 is an indicator of whether the household received a positive quantity of any commodity in a given month. All regressions include strata fixed effects and the baseline value of the dependent variable. Standard errors clustered at the block level are reported in parentheses, with statistical significance denoted as: $^{*}p < .10,^{**}p < .05,^{***}p < .01$. q values adjusted to control the false discovery rate across the five commodities are reported in brackets.

Table 8: Heterogeneous effects by Aadhaar seeding

	At least one Aadhaar seeded?		
	(1) No *(N=797)*	(2) Yes *(N=2645)*	(3) Δ
Value received (market prices)	-49***	.054	49***
	(18)	(14)	(15)
Value received>0	-.1***	-.023	.079***
	(.024)	(.015)	(.022)
Value received (WTA)	-31	37*	68*
	(36)	(22)	(35)
Transaction costs	6.8	8.9**	2.1
	(6.6)	(4.3)	(6.9)

This table reports estimated differential treatment effects by Aadhaar seeding status in endline one (January - March). Column 1 (2) reports estimated treatment effects for households that did not (did) have at least one member with an Aadhaar number seeded to their ration cards at baseline. Column 3 reports the difference between these effects. Each row represents a different primary outcome; all estimates are derived from a single underlying regression that interacts treatment with an indicator equal to one for households with one or more Aadhaar numbers seeded. All specifications include strata fixed effects and the baseline value of the dependent variable when available. Standard errors clustered at the block level are reported in parentheses with statistical significance denoted as: $^{*}p < .10,^{**}p < .05,^{***}p < .01$.

Table 9: Perceptions of the ePOS intervention

	Households	Dealers
Overall, do you prefer ePOS to the old system of manual transactions?	*(N=2182)*	*(N=288)*
Strongly disagree	.44	.45
Weakly disagree	.03	.04
Weakly agree	.08	.09
Strongly agree	.45	.42
Did not know/answer	.01	–
Reasons for preference of manual transactions:	*(N=1023)*	*(N=150)*
It is cheaper to run FPS operations	–	1
Manual transactions faster	.66	.75
Manual transactions easier to understand	.44	.62
There are no problems with network or software	.28	.74
Anyone can collect rations on my behalf	.43	.06
Could give ration to those who did not have ration cards	–	.03
Dealer to verify amounts purchased	.01	–
It is more profitable	–	.01
Other	.01	.02
Reasons for preference of ePOS enabled transactions:	*(N=1165)*	*(N=137)*
ePOS transactions are faster	.54	.60
There is a lower chance of fraud by the FPS dealer	.56	–
The official transaction is equal to what I receive	.14	–
I know my exact ration entitlement and payment amounts	.15	–
Nobody else can collect ration in my name	.38	–
Ration balance carry forward if I don't collect	.06	.13
I receive physical receipts after ePOS transactions	.18	–
I receive text messages after ePOS transactions	.02	–
The dealer calls me to buy ration as he cannot hide supply	.07	–
Better relationship with beneficiaries	–	.64
Beneficiaries are more informed	–	.35
Nobody can steal ration from beneficiary	–	.48
Other	–	.02

This table reports summary statistics of households' and FPS dealers' stated preferences for and perceptions of the ePOS intervention in March. The sample is restricted to households and dealers in treated blocks. In Panel B the sample is further restricted to respondents who said they strongly or weakly disagreed in Panel A, while in Panel C it is restricted to those who strongly or weakly agreed. Estimates are weighted by inverse sampling probabilities. Some values are missing because the list of options provided to households and dealers differed.

Table 10: Dealer expectations

	Intends to continue running FPS?		Expected bribes to obtain license?		Expected bribes to renew license?	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	.054	.033	-56,816**	-58,393*	-111	-83
	(.055)	(.059)	(28,561)	(33,370)	(123)	(147)
Adjusted R^2	.093	.13	.31	.27	.053	.035
Control mean	.73	.71	76,590	81,188	565	555
Observations	437	366	150	127	370	307
% of sample		92		32		78
Sample	Full	Restricted	Full	Restricted	Full	Restricted

This table reports estimated treatment effects on measures of Fair Price Shop dealers' expectations in March about the future. The unit of analysis is the Fair Price Shop. The dependent variable in columns 1-2 is an indicator equal to 1 if the dealer responded "yes" when asked whether they intended to continue running an FPS for the next two years and to 0 if they responded "maybe" or "no." The dependent variable in columns 3-4 is the dealer's estimate of the additional money (excluding official fees) someone would have to pay to obtain a new license to operate a FPS. The dependent variable in columns 5-6 is the dealer's estimate of the additional money (excluding official fees) an existing FPS dealer would have to pay to renew his or her license. In columns 1, 3 and 5 the sample includes all dealers surveyed, including those to whom sampled households switched between baseline and endline; in columns 2, 4 and 6 it includes only dealers drawn in the original sample. All specifications include strata fixed effects. Standard errors clustered at the block level are reported in parentheses with statistical significance denoted as: $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table 11: Effects of reconciliation

	Reconciled			Unreconciled		
	Treatment	Control	Difference	Treatment	Control	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
Mean entitlement	512	504	8	71	68	3
Panel A: Value disbursed						
Reconciliation	-182***	-92***	-90***	-2.1***	.24	-2.3***
	(2.3)	(2.9)	(3.8)	(.24)	(.39)	(.46)
Reconciliation * Month	64***	31***	33***	8.2***	7.9***	.3
	(1)	(1.2)	(1.6)	(.11)	(.15)	(.19)
January 2017 mean	501	494		85	82	
Observations	66,404	31,350		66,404	31,350	
% of frame	96	96		96	96	
Panel B: Value received						
Reconciliation	-62***	-20**	-43***	19	-4.1**	23
	(7.4)	(7.9)	(11)	(32)	(1.7)	(32)
Reconciliation * Month	13***	2.9	10*	-8.3	2.2**	-10
	(3.9)	(4.2)	(5.7)	(8.8)	(.84)	(8.9)
January 2017 mean	399	403		63	61	
Observations	25,469	13,447		25,349	13,334	
% of sample	89	91		88	90	
Panel C: Leakage						
Reconciliation	-121***	-72***	-49***	-21	4.7***	-25
	(9.1)	(10)	(14)	(32)	(1.8)	(32)
Reconciliation * Month	51***	28***	23***	17*	5.5***	11
	(4.7)	(5.4)	(7.1)	(9)	(.93)	(9)
January 2017 mean	97	78		22	24	
Observations	25,469	13,447		25,349	13,334	
% of sample	89	91		88	90	

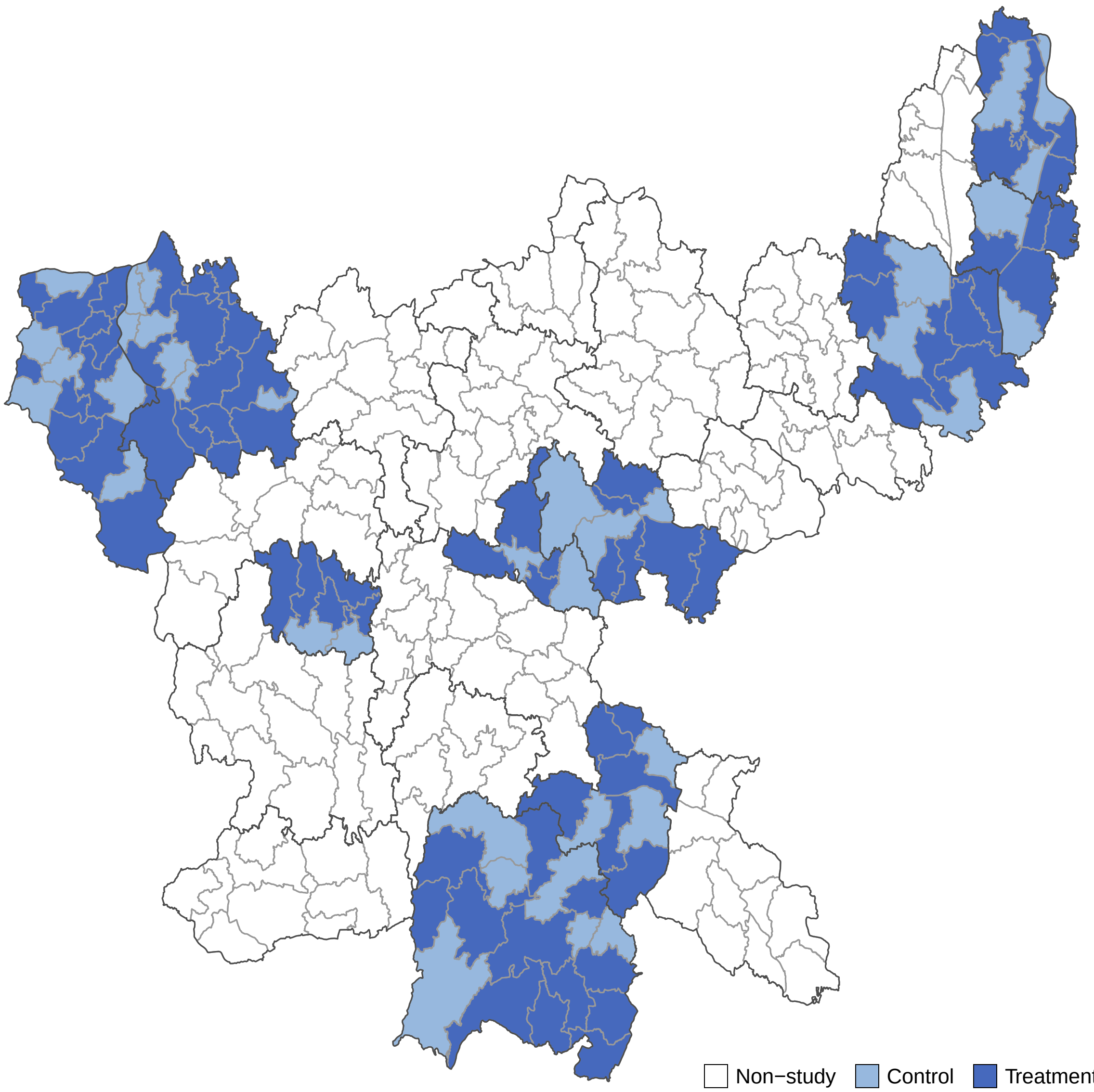
This table reports estimates of the effect of reconciliation on measures of the value disbursed by the government (Panel A), received by recipients (Panel B), and the difference (Panel C) separately for treatment and control areas using data from all three endlines. The unit of analysis is the Fair Price Shop-month in Panel A and the ration card-month in Panels B and C, but all figures are per ration card-month. Observation counts vary by panel because we use the universe of Fair Price Shops to estimate effects on disbursements in Panel A, and a representative sample of ration card holders in Panels B and C, but all samples are representative. The dependent variable in columns 1 and 2 is the sum of values for rice and wheat, and the dependent variable in columns 4 and 5 is the sum of values for sugar, salt, and kerosene. Per-commodity values are defined in the notes to Table 3 above. Columns 3 and 6 test the difference between columns 1 ,2 and 4,5, respectively. Standard errors clustered at the FPS level are reported in parentheses, with statistical significance denoted as: $^{*}p < .10,^{**}p < .05,^{***}p < .01$.

Table 12: Decomposition of ration card deletions

	Deleted	Non-deleted	Total	%
	(1)	(2)	(3)	(4)
Admin data				
Unseeded	80,085	133,004	213,089	8.7%
Seeded	64,076	2,172,445	2,236,521	91.3%
Total	144,161	2,305,449	2,449,610	100%
% of overall total	5.89%	94.11%	100%	
Survey data				
Unseeded and ghost	15	6	21	.5%
Unseeded and not ghost	90	142	232	5.9%
Seeded and ghost	11	21	32	.8%
Seeded and not ghost	97	3519	3616	92.7%
Total	213	3688	3901	100%
% of overall total	5.5%	94.5%	100%	

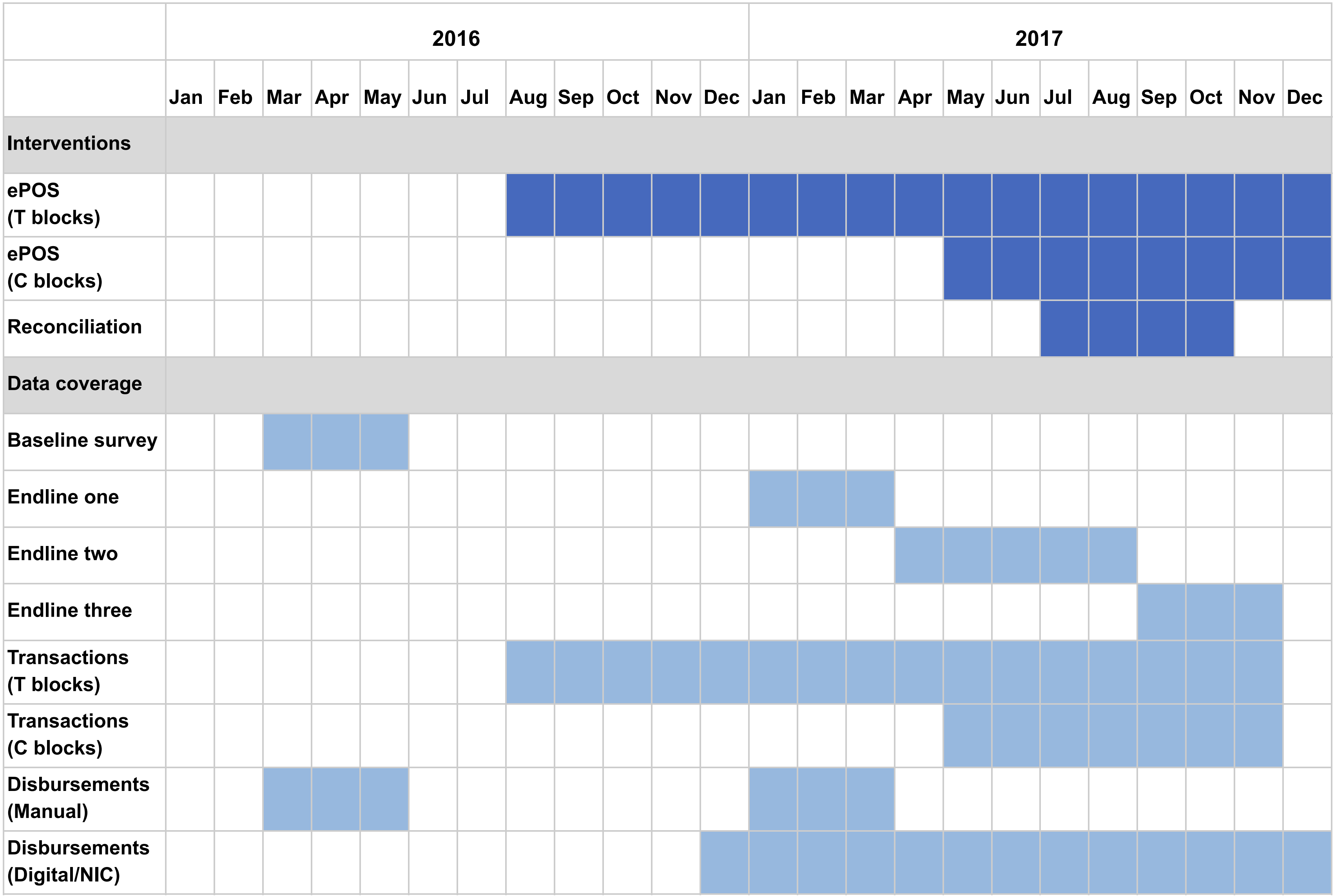
This table reports the decomposition of ration card deletions by Aadhaar seeding status. The top panel "Admin data" shows the results from the universe of ration cards in our 10 study districts, and the bottom panel "Suryey data" shows results from our sampled ration cards, for which we show counts adjusted by sampling probability and categorized beneficiaries as "ghosts" based on survey results. A ration card is "Deleted" if it was present in the beneficiary list in October 2016 but absent in May 2018, and is "Non-deleted" if still present in May 2018. A ration card is "Unseeded" if it did not have any Aadhaar number seeded to the card in October 2016, and "Seeded" if it did.

Figure 1: Blockwise treatment assignment



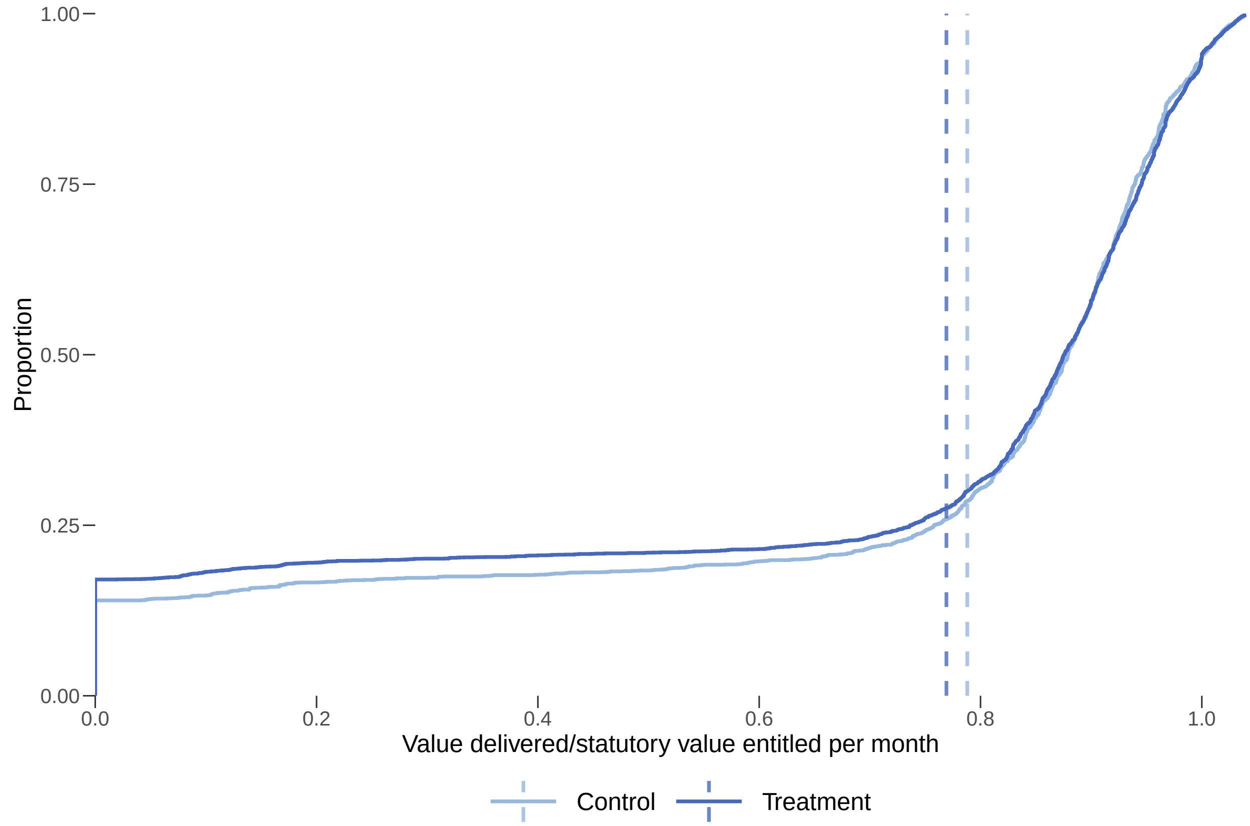
This figure shows the assignment of districts within Jharkhand to study (10) and non-study (14) status, and the assignment of blocks within these districts to treatment and control. Note that four of the census blocks depicted here are further sub-divided for the purposes of PDS administration into an urban and a rural “PDS block;” in these cases we give the entire census block the color corresponding to the treatment status of its larger, rural PDS block.

Figure 2: Intervention and data collection timeline



This figure plots the evolution of the interventions delivered by the Government of Jharkhand (top panel) and the coverage of the various data sources we use for analysis (bottom panel). Transaction data coverage in control areas in May 2017 is partial, as the rollout of ePOS devices in control areas began but did not finish in that month.

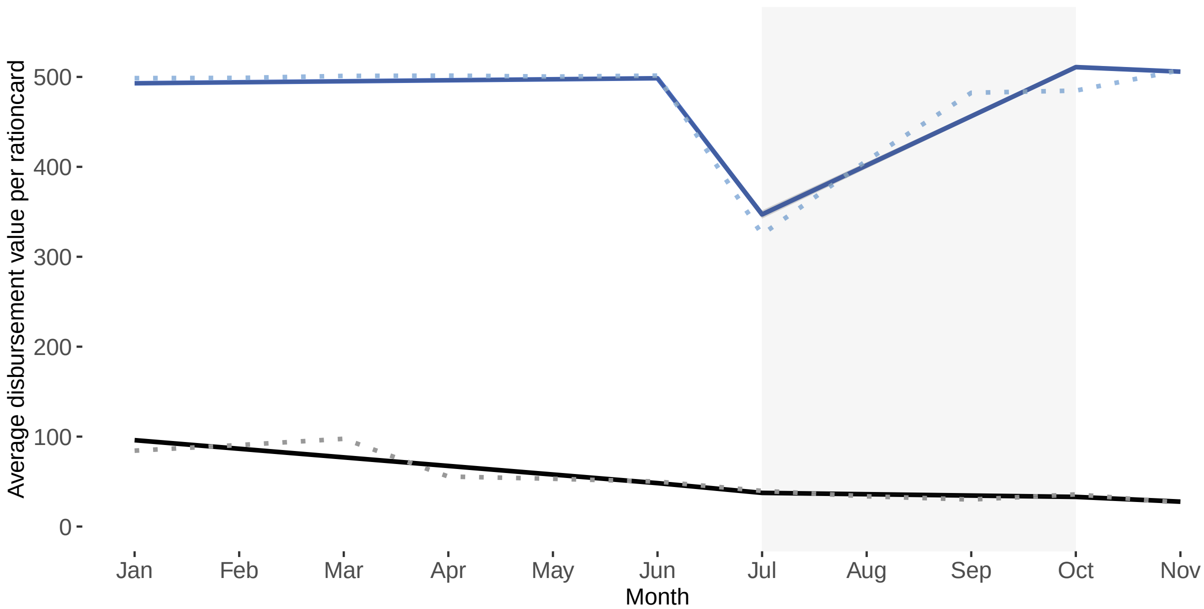
Figure 3: Value received as a proportion of entitlement



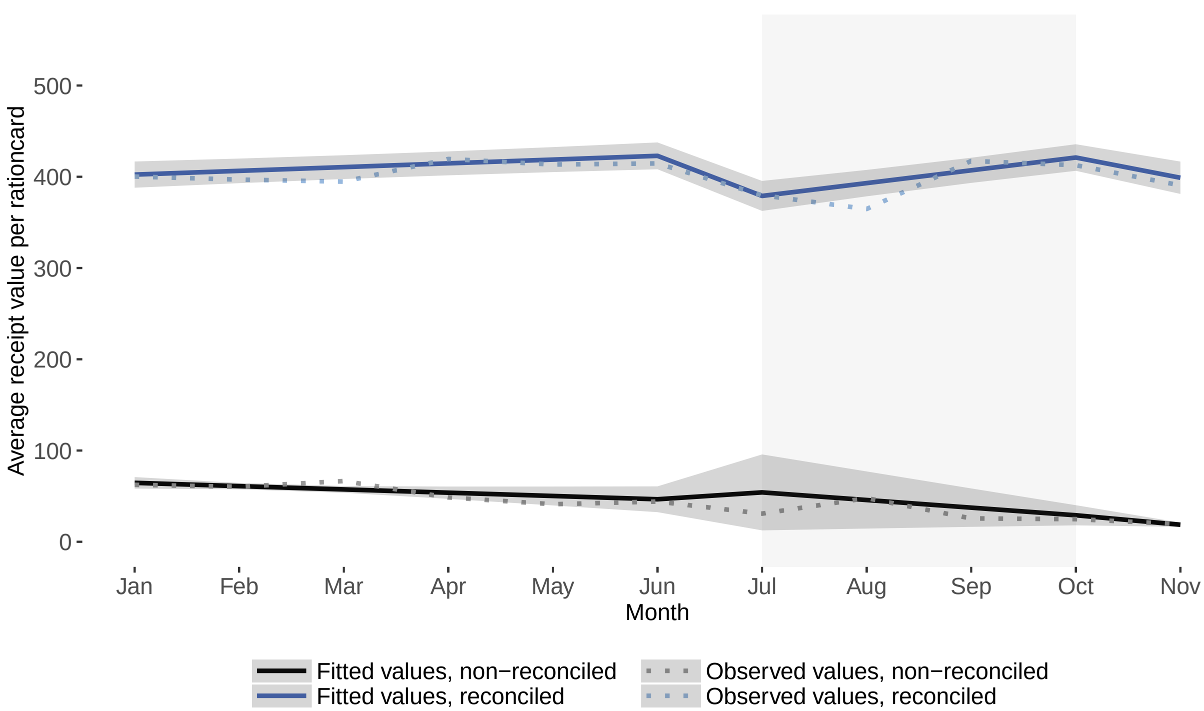
This figure plots the empirical cumulative distribution, separately for households in treatment and control blocks, of value received divided by value entitled per month, pooling the months of January, February, and March 2017. To improve legibility we right-censor the distributions at the 90th percentile.

Figure 4: Effects of reconciliation on value disbursed and received

(a) Value disbursed

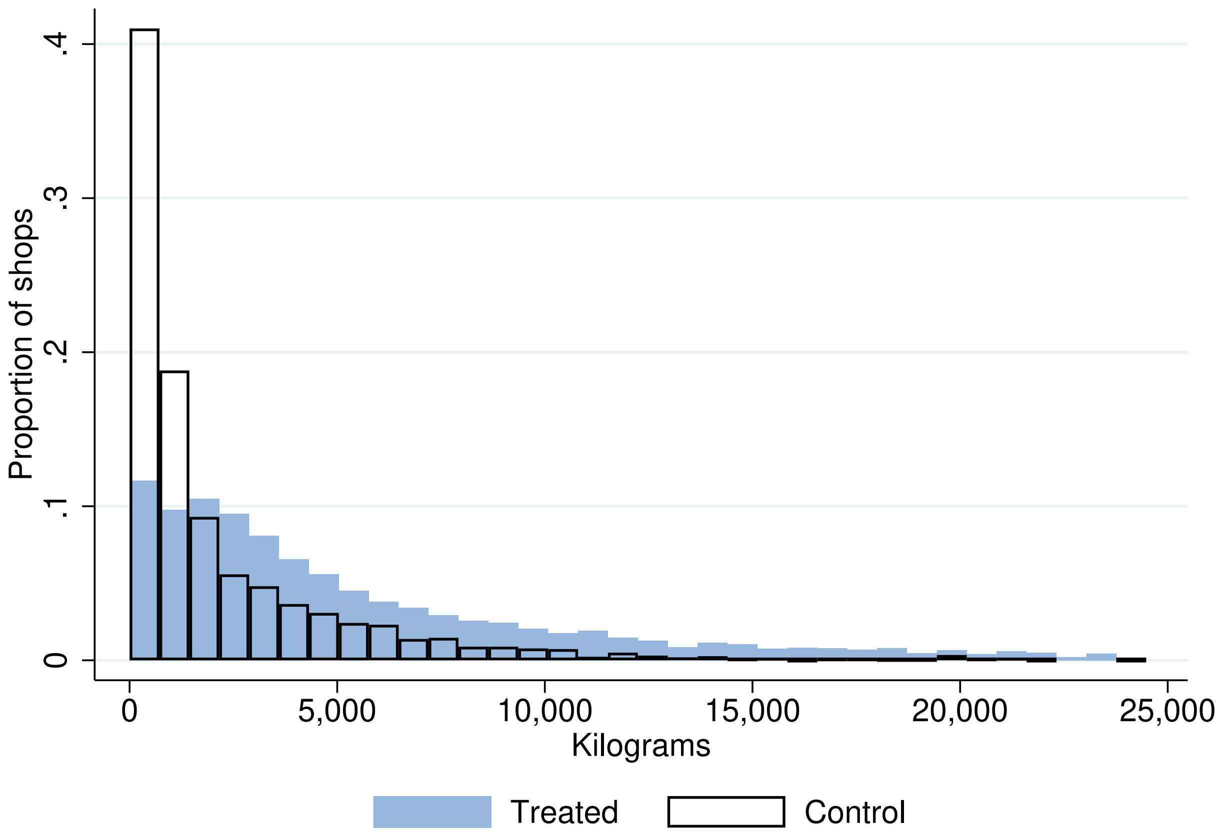


(b) Value received



This figures plots the evolution of the average value of commodities disbursed (Panel A) and received (Panel B) from January to November of 2017. The unit is the ration card-month. Value disbursed is calculated from administration data and value received from our series of endline surveys, using market price data as described in the notes to Table 3. Dashed lines represent the raw data, while solid lines represent fitted values obtained by estimating Equation 2. The shaded bands around the latter represent 95% confidence intervals for the fitted values. Values are shown separately for commodities that were (blue) and were not (black) separately subject to reconciliation. The shaded region from July to November indicates the period during which reconciliation was in effect.

Figure 5: Recorded grain stock as of June 2017, by treatment status



This figure shows the distribution of grain (i.e. rice and wheat) in kilograms held by Fair Price Shops at the end of June 2017 according to government records, separately for shops in treated and control blocks. The unit of observation is the Fair Price Shop. To increase legibility the distributions are right-censored at the 95th percentile.

A Supplemental exhibits

Table A.1: Comparison of dealer samples

	Original sample	Additional sample	Regression-adjusted difference	p-value
Treatment	.67	.77	-.081*	.07
Age	44	43	1.5	.33
Years of education	9.8	10	-.22	.74
Has an FPS dealer in family	.13	.27	-.12	.18
Years as FPS dealer	14	14	-.094	.95
Has other income sources	.79	.73	.066	.32
Runs FPS out of own home	.61	.73	-.16*	.06
Days open per month	19	20	-1.4*	.10
Hours open per day	6.7	6.7	-.073	.80
Days mandated to be open per month	23	24	-.58	.56
Hours selling PDS commodities per day	6.5	7.3	-.78	.50
Hours mandated to be open per day	6.9	7	.01	.97
Number of total ration cards	269	291	-20	.51
Number of PH ration cards	225	249	-21	.46
Number of AAY ration cards	44	42	.7	.89
Number of villages	2	2.5	-.38	.12

This table compares the PDS dealers originally sampled at baseline ("original sample") with those added at the first endline as a result of ration card re-assignment across ration shops ("additional sample"). Columns 1 and 2 report the means of each variable for the respective groups. Column 3 reports the coefficient from a regression of the given variable on an indicator for being in the original sample, controlling for strata fixed effects, and column 4 reports the p-value for a test that this coefficient is zero. Estimates are weighted by inverse sampling probabilities. Statistical significance is denoted as: $^{*}p < .10,^{**} p < .05,^{***} p < .01$

Table A.2: Outcome missingness by treatment status: all households

	Treatment	Control	difference	p-value
HH classified as ghost	.013	.025	-.0059	.15
Quantity rice purchased in January	.032	.039	-.0039	.66
Quantity rice purchased in February	.035	.037	-.00042	.96
Quantity rice purchased in March	.034	.041	-.0071	.41
Quantity wheat purchased in January	.026	.025	.0031	.69
Quantity wheat purchased in February	.028	.025	.0034	.67
Quantity wheat purchased in March	.025	.024	.0016	.82
Quantity sugar purchased in January	.021	.029	-.0078	.18
Quantity sugar purchased in February	.02	.028	-.0076	.19
Quantity sugar purchased in March	.024	.026	-.0017	.75
Quantity salt purchased in January	.016	.024	-.0073	.17
Quantity salt purchased in February	.015	.023	-.0081	.11
Quantity salt purchased in March	.017	.02	-.0033	.51
Quantity kerosene purchased in January	.025	.038	-.0084	.35
Quantity kerosene purchased in February	.025	.036	-.0081	.38
Quantity kerosene purchased in March	.026	.038	-.0081	.33
Value rice purchased in January	.032	.039	-.0039	.66
Value rice purchased in February	.035	.037	-.00042	.96
Value rice purchased in March	.034	.041	-.0071	.41
Value wheat purchased in January	.026	.025	.0031	.69
Value wheat purchased in February	.028	.025	.0034	.67
Value wheat purchased in March	.025	.024	.0016	.82
Value sugar purchased in January	.021	.029	-.0078	.18
Value sugar purchased in February	.02	.028	-.0076	.19
Value sugar purchased in March	.024	.026	-.0017	.75
Value salt purchased in January	.016	.024	-.0073	.17
Value salt purchased in February	.015	.023	-.0081	.11
Value salt purchased in March	.017	.02	-.0033	.51
Value kerosene purchased in January	.025	.038	-.0084	.35
Value kerosene purchased in February	.025	.036	-.0081	.38
Value kerosene purchased in March	.026	.038	-.0081	.33
Total value purchased in January	.049	.054	-.0019	.84
Total value purchased in February	.045	.054	-.0069	.47
Total value purchased in March	.049	.054	-.0047	.62

This table reports the rate at which various household outcomes measured in endline one (January - March) are not observed, by treatment status. We include all surveyed households and all households categorized as "ghosts." Columns 1 and 2 report the mean of each outcome among treatment and control households, respectively. Column 3 reports the simple differnece between these, and Column 4 reports the p-value on a test of the null that this difference is equal to zero. Estimates are weighted by inverse sampling probabilities.

Table A.3: Outcome missingness by treatment status: surveyed households

	Treatment	Control	difference	p-value
Sampled rationcard is inactive	.019	.013	.0089	.097*
Household does not know of sampled rationcard	.054	.073	-.015	.18
Willingness to accept in January	.12	.12	.0036	.78
Willingness to accept in February	.14	.12	.025	.11
Willingness to accept in March	.14	.12	.022	.2
Rice was low quality	.071	.06	.021	.043**
Wheat was low quality	.77	.66	.08	.15
Sugar was low quality	.36	.3	.038	.22
Salt was low quality	.33	.34	.0021	.95
Kerosene was low quality	.1	.094	.023	.097*
Rice was adulterated	.077	.067	.019	.095*
Wheat was adulterated	.77	.66	.081	.15
Sugar was adulterated	.36	.3	.039	.21
Salt was adulterated	.34	.35	.00093	.98
Kerosene was adulterated	.1	.11	.016	.27
Access cost in January	.11	.13	-.015	.27
Access cost in February	.11	.13	-.015	.31
Access cost in March	.0085	.016	-.0057	.13
Total access cost in March	.0085	.016	-.0057	.13
Overcharge on rice in January	.045	.062	-.0099	.37
Overcharge on rice in February	.047	.063	-.01	.35
Overcharge on rice in March	.045	.069	-.02	.074*
Overcharge on wheat in January	.035	.045	-.0041	.69
Overcharge on wheat in February	.037	.05	-.0071	.51
Overcharge on wheat in March	.034	.048	-.0081	.41
Overcharge on sugar in January	.041	.052	-.0074	.34
Overcharge on sugar in February	.04	.056	-.017	.099*
Overcharge on sugar in March	.042	.045	-.0024	.74
Overcharge on salt in January	.032	.037	-.0037	.58
Overcharge on salt in February	.028	.036	-.0092	.2
Overcharge on salt in March	.031	.036	-.0053	.35
Overcharge on kerosene in January	.044	.058	-.0085	.48
Overcharge on kerosene in February	.044	.058	-.0092	.47
Overcharge on kerosene in March	.051	.059	-.0047	.7
Total overcharge in January	.094	.11	-.004	.79

Total overcharge in February	.089	.12	-.024	.12
Total overcharge in March	.097	.12	-.021	.17
FPS-level market price of rice in March	.027	.039	-.011	.62
FPS-level market price of wheat in March	.38	.47	-.098	.064*
FPS-level market price of sugar in March	.027	.045	-.019	.38
FPS-level market price of salt in March	.0051	0	.0056	.18
FPS-level market price of kerosene in March	.36	.42	-.078	.16
HH-level market price of rice in March	.53	.53	-.025	.22
HH-level market price of wheat in March	.87	.87	-.021	.16
HH-level market price of sugar in March	.45	.49	-.037	.2
HH-level market price of salt in March	.3	.3	-.005	.81
HH-level market price of kerosene in March	.87	.89	-.018	.19

This table reports the rate at which various household outcomes measured in endline one (January - March) are not observed, by treatment status. We include only surveyed households. Columns 1 and 2 report the mean of each outcome among treatment and control households, respectively. Column 3 reports the simple differnece between these, and Column 4 reports the p-value on a test of the null that this difference is equal to zero. Estimates are weighted by inverse sampling probabilities.

Table A.4: Effects on quantity disbursed, quantity received, and leakage

	Rice	Wheat	Sugar	Salt	Kerosene
	(1)	(2)	(3)	(4)	(5)
Mean entitlement	24	1	1	1	2
Panel A: Quantity disbursed					
Treatment	1.457**	-1.072*	0.007	0.003	0.000
	(0.568)	(0.559)	(0.008)	(0.005)	(0.000)
	[0.13]	[0.33]	[1.00]	[1.00]	[0.38]
Control mean	20	3.5	1.3	1	2.4
Observations	26,611	26,611	26,611	26,611	26,611
Panel B: Quantity received					
Treatment	.76	-.58	.026	.056	-.034
	(.5)	(.48)	(.08)	(.065)	(.048)
	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
Control mean	17	2.6	.72	.81	1.8
Observations	10,557	10,654	10,670	10,726	10,618
Panel C: Leakage					
Treatment	.68	-.5	-.019	-.053	.05*
	(.57)	(.32)	(.081)	(.063)	(.028)
	[0.88]	[0.68]	[1.00]	[1.00]	[0.68]
Control mean	3.6	.88	.59	.22	.65
Observations	10,557	10,654	10,670	10,726	10,618

This table reports estimated treatment effects on the quantity of commodities disbursed by the government (Panel A), received by recipients (Panel B), and the difference (Panel C) in endline one (January - March). The unit of analysis is the Fair Price Shop-month in Panel A and the ration card-month in Panels B and C. Observation counts vary by panel because we use the universe of Fair Price Shops to estimate effects on disbursements in Panel A, and a representative sample of ration card holders in Panels B and C, but all samples are representative. In Panel C, estimated effects are the difference between estimated effects on quantity disbursed per ration card and quantity received per ration card with block-level mean imputation in a Seemingly Unrelated Regression framework. All specifications include strata fixed effects and the baseline value of the dependent variable. Standard errors clustered at the block level are reported in parentheses with statistical significance denoted as: $^{*}p < .10,^{**} p < .05,^{***} p < .01$. q values adjusted to control the false discovery rate across the five commodities are reported in brackets.

Table A.5: Effects on market prices and overcharges

	Total	Rice	Wheat	Sugar	Salt	Kerosene
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Local market prices—reported by households						
Treatment	–	.043	5.6	-1.1***	.17	1.4
	–	(.29)	(5.3)	(.39)	(.16)	(1.8)
	–	[1.00]	[1.00]	[0.09]	[1.00]	[1.00]
Control mean	.	23	22	43	10	43
Observations	.	383	248	382	392	229
Panel B: Local market prices—reported by dealers						
Treatment	–	-1.3***	.021	-1.3*	-.31	-1.7*
	–	(.46)	(.92)	(.69)	(.54)	(.94)
	–	[0.06]	[1.00]	[0.25]	[1.00]	[0.25]
Control mean	.	19	17	38	8.5	38
Observations	.	344	109	283	282	251
Panel C: Statutory prices						
Treatment	–	–	–	–	–	.027
						(.03)
	–					
Control mean		–	–	–	–	18
Observations	.	.	.	.	.	396
Panel D: Overcharges						
Treatment	-2.6	.069	-.13**	-2.1	.016	-.66
	(1.9)	(.24)	(.056)	(1.7)	(.035)	(.51)
		[1.00]	[0.21]	[0.84]	[1.00]	[0.84]
Control mean	8.2	1.1	.22	.91	.17	6
Observations	9,623	10,183	10,317	10,260	10,375	10,185

This table reports estimated treatment effects on the market prices reported by beneficiaries (Panel A), market prices reported by FPS dealers (Panel B), statutory prices (Panel C), and total overcharges (Panel D) in endline one (January - March). The unit of analysis is the Fair Price Shop for Panels A and B, the block-month for Panel C, and the ration card-month for Panel D. Prices are in rupees per kilogram except for kerosene, which is priced in rupees per liter. Observation counts vary in panels A and B as we observe outcomes only when at least one household purchased the commodity and when the dealer reported the commodity is sold in the private market, respectively. In Panels A-C the dependent variables are the median market price reported by beneficiaries assigned at baseline to the given FPS, the local market price reported by FPS dealers (Panel B), and the statutory PDS price, respectively. We do not report effects on statutory prices for goods other than kerosene as these did not vary. In Panel D the dependent variable is the amount beneficiaries report paying above what they should have paid for the quantity they received, by commodity in columns 2-6 and in total in column 1. All regressions include strata fixed effects; those in Panels C and D also include the baseline value of the dependent variable. Standard errors clustered at the block level are reported in parentheses with statistical significance denoted as: $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$. q values adjusted to control the false discovery rate across the five commodities are reported in brackets.

Table A.6: Effects on determinants of beneficiary transaction costs

	Total Cost		Opportunity cost	Unsuccessful trip count	Unsuccessful trip length	Successful trip count	Successful trip length	Transport cost
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	7.2*	6.9*	.87*	.13***	.14	.016	.15	-.81
	(3.8)	(3.8)	(.5)	(.024)	(.32)	(.054)	(.092)	(.52)
Baseline lag		1.6**						
		(.79)						
Adjusted R^2	.06	.09	.08	.10	.01	.06	.06	.00
Control mean	41	41	11	.13	1.2	1.5	2.3	1.6
Observations	3,538	3,538	3,066	3,565	449	3,565	3,062	3,538
% of sample	90	90	78	91	11	91	77	90

This table reports estimated treatment effects on the costs incurred by beneficiaries to access PDS rations in March 2017. The unit of analysis is the ration card. The dependent variable in columns 1-2 is the total estimated cost as reported in Table 5, and the remaining columns show impacts on its components. The dependent variable in column 3 is the weighted mean opportunity cost in rupees per hour of household members, weighted by the number of trips each household member made to their FPS in March. The dependent variables in columns 4 and 5 are the number of unsuccessful trips made to the ration shop (defined as trips that did not result in the purchase of positive quantities of any rationed commodity) and the average time in hours spent on these trips; the dependent variables in columns 6 and 7 are the analogous quantities for successful trips. Finally, the dependent variable in column 8 is the average monetary cost in rupees of any transport fees paid to make these trips (e.g. bus fare). Thus, the total cost in column1 equals to unit opportunity cost * (total time spent on unsuccessful trips + total time spent on successful trips) + transporation cost * (number of unsuccessful trips + number of successful trips) = column 3 * (column 4 * column 5 + column 6 * column 7) + column 8 * (column 4 + column 6). All regressions include strata fixed effects. Standard errors clustered at the block level are reported in parentheses, with statistical significance denoted as: $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table A.7: Heterogeneous effects by household characteristics

	HH is upper caste?			HH above median education level?			HH above median annual income?		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	No *(N=1875)*	Yes *(N=1705)*	Δ	No *(N=2083)*	Yes *(N=1500)*	Δ	No *(N=1634)*	Yes *(N=1608)*	Δ
Value received (market prices)	-1.4	-5	-3.6	-21	23*	44***	-15	11	26*
	(15)	(12)	(13)	(13)	(14)	(13)	(14)	(15)	(14)
Value received (WTA)	55**	-14	-69***	38	-.027	-38	26	30	4.1
	(26)	(24)	(25)	(25)	(27)	(29)	(26)	(27)	(27)
Transaction costs	4.5	9.6**	5.1	7.9**	5.4	-2.5	2	14***	12***
	(4.4)	(4.7)	(5)	(3.9)	(4.5)	(3.7)	(3.9)	(4.6)	(3.8)

This table reports differential estimated treatment effects along dimensions of household characteristics for ration card-level outcomes in endline one (January - March). Each row represents a different primary outcome, and each column grouping represents a different dimension of heterogeneity. Within each column group, the first column reports the average treatment effect on households that do not satisfy the stated condition, the second column reports the average effect on those that do, and the third column reports the difference in these two effects, all estimated from a single underlying regression that interacts treatment with an indicator for the stated condition. The indicator "HH is upper caste" is equal to one if the household does not belong to Scheduled Caste or Scheduled Tribe. The indicator "HH above median education level" is equal to one if the average number of years of schooling of the two highest-educated household members is above the sample median. The indicator "HH above median annual income" is equal to one if the household's annual income (as predicted from assets and household characteristics at baseline) is above the sample median. All regressions include strata fixed effects and baseline value of the outcome variable when available. Standard errors clustered at the block level are reported in parentheses, with statistical significance denoted as: $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table A.8: Heterogeneous effects by location characteristics

	FPS in urban area?			Network strength above median?		
	(1)	(2)	(3)	(4)	(5)	(6)
	No *(N=3129)*	Yes *(N=513)*	Δ	No *(N=1565)*	Yes *(N=1444)*	Δ
Value received (market prices)	-.74	-5.9	-5.2	-3.1	-4.2	-1.1
	(13)	(19)	(23)	(13)	(16)	(18)
Value received (WTA)	30	-79	-109**	9.2	4.4	-4.8
	(23)	(48)	(53)	(29)	(29)	(37)
Transaction costs	7.1*	4.3	-2.8	5.9	1.1	-4.8
	(3.9)	(14)	(14)	(5)	(5.7)	(6.5)

This table reports differential estimated treatment effects along dimensions of location characteristics for ration card-level outcomes in endline one (January - March). Each row represents a different primary outcome, and each column grouping represents a different dimension of heterogeneity. Within each column group, the first column reports the average treatment effect on households that do not satisfy the stated condition, the second column reports the average effect on those that do, and the third column reports the difference in these two effects, all estimated from a single underlying regression that interacts treatment with an indicator for the stated condition. The indicator "FPS in urban area?" is equal to one if the household's FPS belongs to block that is administratively classified as urban. The indicator "Network strength above median?" is equal to one if the measured signal strength of the Airtel network (which was the most common SIM card type installed in ePOS machines) at the household's assigned FPS is above the sample median. All regressions include strata fixed effects and baseline value of the outcome variable when available. Standard errors clustered at the block level are reported in parentheses, with statistical significance denoted as: $^{*}p < .10,^{**}p < .05,^{***}p < .01$.

Table A.9: Heterogeneous effect by machine mode

	Value received (mkt prices)	Value received (WTA)	Transaction costs
	(1)	(2)	(3)
Treatment*Online	1.8	1	3.1
	(15)	(30)	(5.2)
Treatment*Offline	-1.6	-57	9.2
	(22)	(52)	(9.5)
Treatment*Partial	-37	-52	19*
	(29)	(104)	(10)
Adjusted R^2	.20	.06	.08
Control mean	463	1028	41
Observations	9733	9787	3337
% of sample	82	84	86

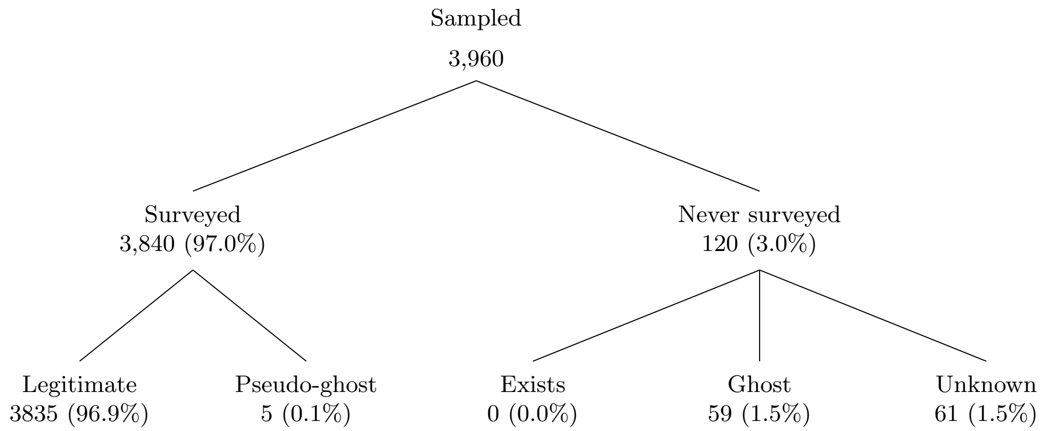
This table reports differential estimated treatment effects by machine mode assigned to dealers per month for ration card-level outcomes in endline one (January - March). The unit of analysis for columns 1 and 2 is the ration card-month and for column 3 the ration card. The dependent variable in column 1 is the sum of the values for each commodity, defined as the quantity multiplied by the difference between the median market price of that commodity in control blocks in the same district, and the statutory PDS price for that commodity. The dependent variable in column 2 is the household reported willingness to accept (WTA), constructed as the smallest value X for which the respondent reported that they would have preferred in cash to the commodities received. The WTA for ration cards that did not receive any ration is set to zero. The dependent variable in column 3 is the total cost incurred in March by the household holding that ration card in purchasing or attempting to purchase PDS commodities, including time and money costs. All regressions include strata fixed effects, and columns 1 and 3 include the baseline value of the outcome variable. Standard errors clustered at the block level are reported in parentheses, with statistical significance denoted as: $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table A.10: Heterogeneous effects of reconciliation on value disbursed and received by FPS balance

	Value disbursed			Value received		
	(1) Total	(2) Rice	(3) Wheat	(4) Total	(5) Rice	(6) Wheat
Reconciliation	-59***	-35***	-19***	7.3	3.7	-2.1
	(3.9)	(3.3)	(2.6)	(13)	(11)	(18)
Reconciliation*Balance	-.21***	-.24***	.03	-.1***	-.1***	-.014
	(.0089)	(.0092)	(.056)	(.028)	(.026)	(.61)
Reconciliation*Month	18***	14***	3.8***	-4	-3.2	3
	(2.1)	(1.9)	(1)	(8.6)	(6.6)	(16)
Reconciliation*Month*Balance	.078***	.083***	.022	.027	.027*	-.072
	(.0045)	(.0045)	(.021)	(.017)	(.015)	(.6)
January 2017 mean	499	442	57	400	355	45
Observations	92489	95581	93258	35518	36269	36046
% of frame/sample	91	94	92	82	83	83

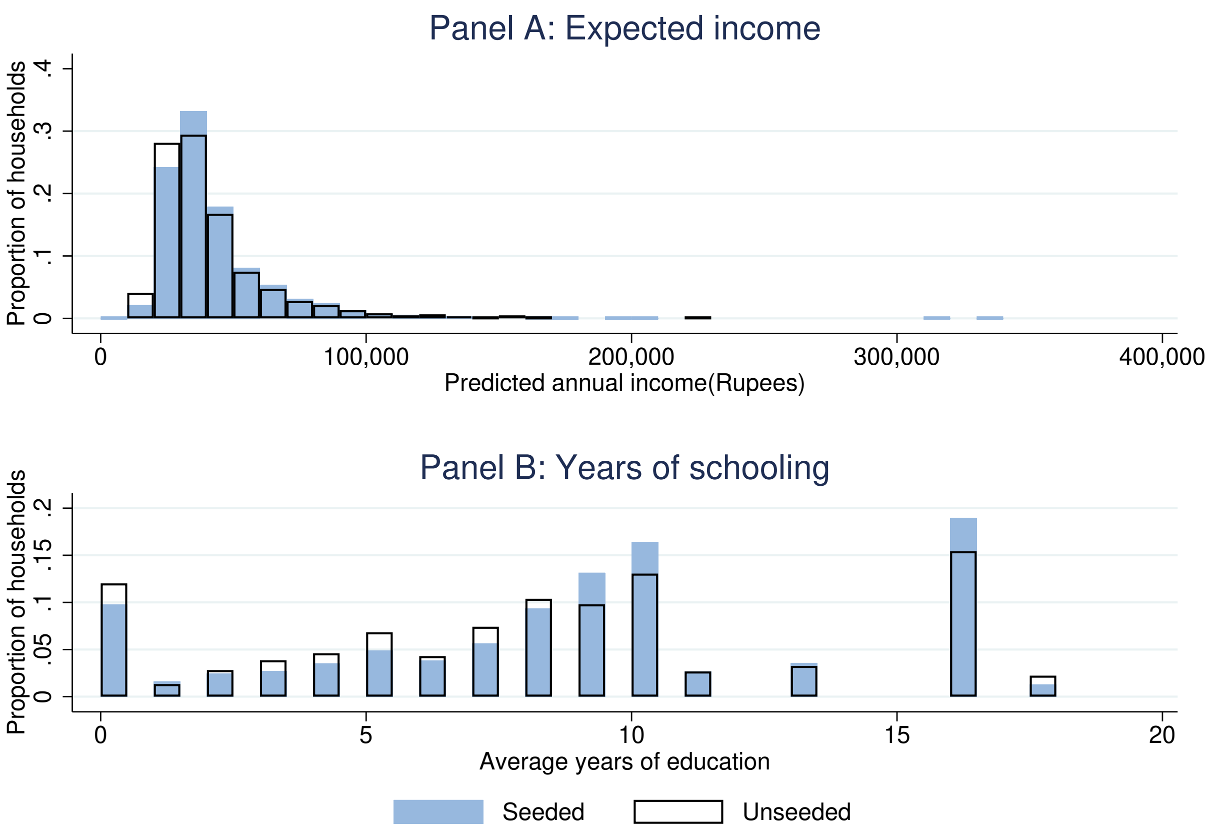
This table reports differential effects of accumulated stock balance at the time of reconciliation onset (July 2017) on value disbursed by dealers (columns 1 - 3) and received by beneficiaries (columns 4 -6). The unit of analysis is the Fair Price shop-month for columns 1-3 and ration card - month for columns 4-6. Observation counts vary because we use the universe of Fair Price Shops to estimate effects on disbursements in columns 1-3, and a representative sample of ration card holders in columns 4-6, but both samples are representative. The dependent variable in columns 2, 3, 5, and 6 is the per-commodity value disbursed and received as defined in the notes to Table 3 above. The dependent variable in column 1 is the sum of the values from columns 2 and 3, and the dependent variable in column 4 is the sum of values in columns 5 and 6. We calculate average balance per ration card as the balance per FPS at the beginning of July 2017, provided by NIC, divided by ration card counts per FPS, and we instrument for this FPS-level average balance per ration card using the block's initial ePOS treatment assignment. Standard errors clustered at the block level are reported in parentheses, with statistical significance denoted as: $^{*}p < .10,^{**}p < .05,^{***}p < .01$. All regressions include strata fixed effects and their interactions with stock balance.

Figure A.1: Household classification results



This figure classifies the households listed in the ration cards we originally sampled and attempted to survey. "Surveyed" households are those we were able to locate and survey at least once across baseline and the three endline surveys, as opposed to "never surveyed" households. Among the former "legitimate" are those that knew of the ration card based on which we sampled them, and "quasi-ghosts" are those that were unaware of the existence of this ration card across all surveys. Among the latter, households "exist" if we were able to locate the household but not survey it; are a "ghost" if we could not locate it after an exhaustive search and confirmed with multiple neighbors that it did not exist; and as "unknown" otherwise.

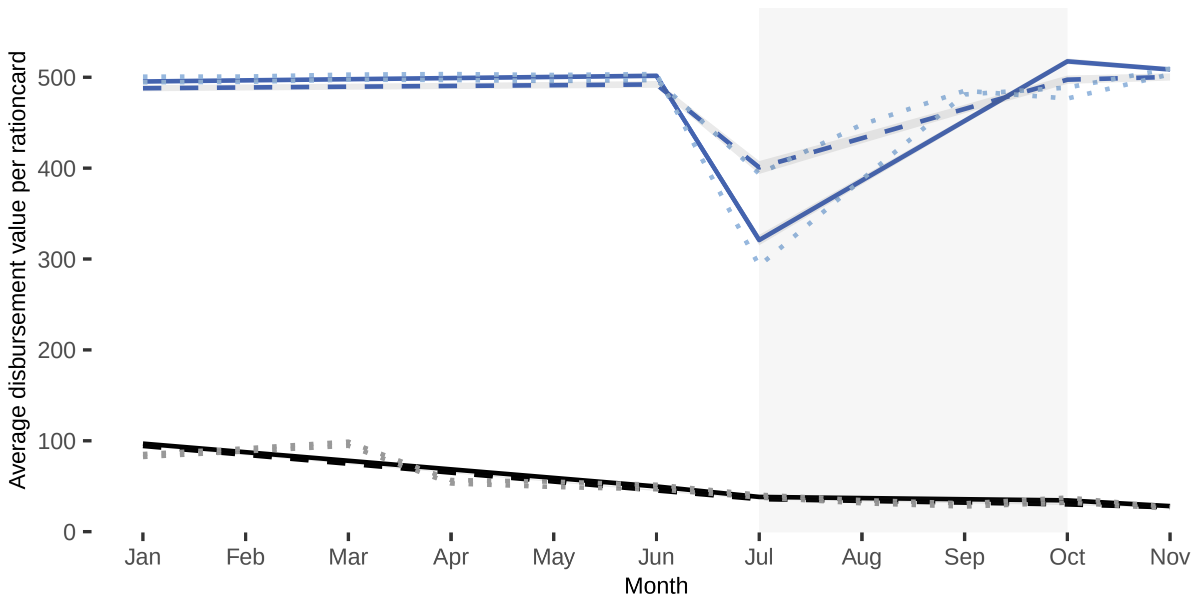
Figure A.2: Household expected income and years of schooling by seeding status



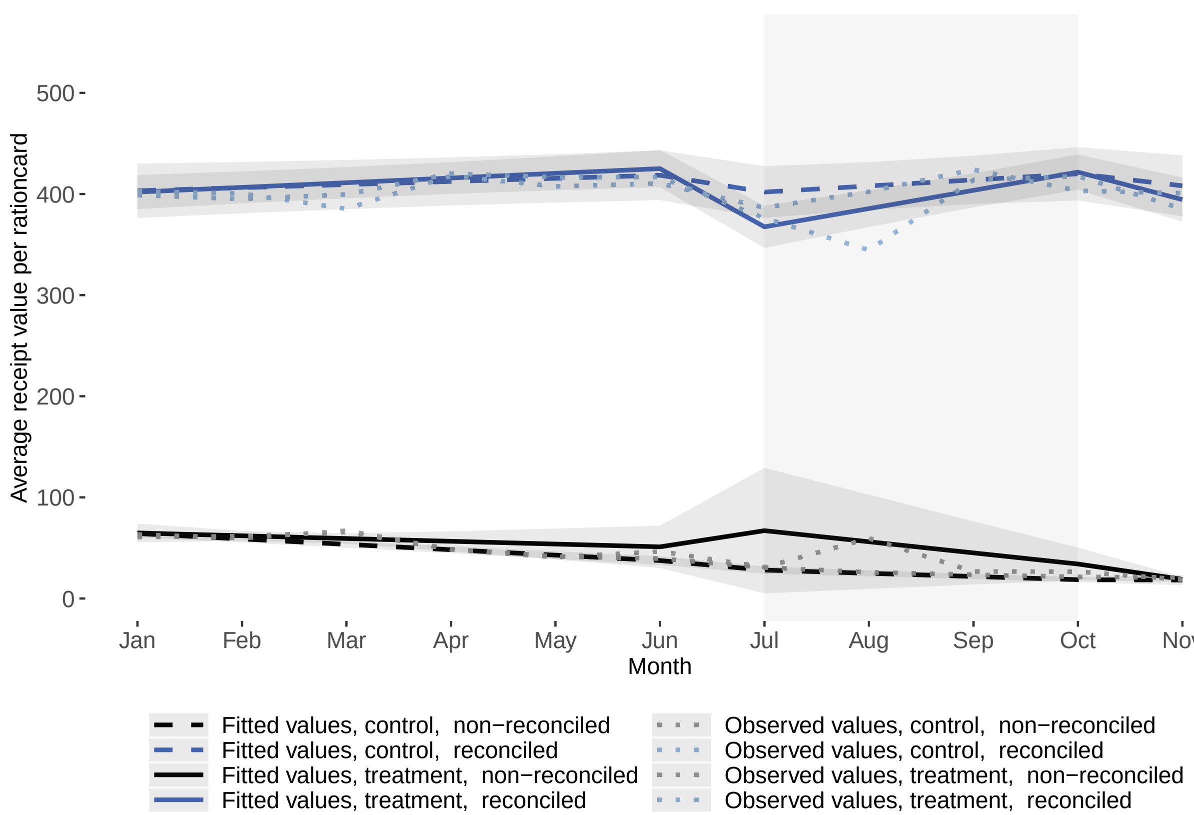
This figure shows the distribution of measures of household income (Panel A) and education (Panel B) by whether the ration card that household was attached to had at least one Aadhaar number seeded at baseline. In Panel A, the outcome is the linear prediction of annual income based on assets and household characteristics at baseline. In Panel B, the outcome is the average years in education of the two most educated members in the household.

Figure A.3: Effects of reconciliation on value disbursed and received, by treatment

(a) Value disbursed



(b) Value received



This figure plots the evoluton of the average value of commodities disbursed (Panel A) and received (Panel B) by treatment from January to November of 2017. The unit is the ration card-month. Value disbursed and value received are as described in Figure 4. Dotted lines represent the raw data, while solid lines and dashed lines represent fitted values obtained by estimating Equation2 for treatment and control, respectively. The shaded bands represent 95% confidence intervals for the fitted values. Values are shown separately for commodities that were (blue) and were not (black) separately subject to reconciliation. The shaded region from July to November indicates the period of reconciliation.

B Temporal heterogeneity

Table B.1: Effects on value received using alternative specifications

	Total	Rice	Wheat	Sugar	Salt	Kerosene
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: March only						
Treatment	9.7	31***	-19*	.049	1.1	.94
	(13)	(11)	(10)	(2.1)	(1.2)	(1.8)
		[0.09]	[0.45]	[1.00]	[1.00]	[1.00]
Control mean	456.71	337.81	55.57	15.80	8.56	37.83
Observations	3,460	3,517	3,553	3,551	3,575	3,533
Panel B: Pooled data with linear trend						
Treatment	-12	-1.1	-9	3	-.88	-1.9
	(18)	(16)	(12)	(5.5)	(1.8)	(1.8)
		[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
Month	-5.7	-8.3*	1.5	1.9	1.2*	-1.9***
	(5)	(4.5)	(1.6)	(1.4)	(.68)	(.47)
Treatment X Month	5.2	9	-3*	-1.2	.69	.66
	(6.7)	(5.6)	(1.8)	(2.3)	(.89)	(.86)
Control mean	463.30	348.18	53.73	13.80	7.25	39.64
Observations	10,396	10,557	10,654	10,670	10,726	10,618
Panel C: Pooled data with no baseline lag						
Treatment	.85	14	-13	.38	.62	-.33
	(14)	(12)	(11)	(1.5)	(.58)	(1.1)
		[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
Control mean	463.30	348.18	53.73	13.80	7.25	39.64
Observations	10,396	10,557	10,654	10,670	10,726	10,618

This table reports alternative specifications for Panel B in Table 3 by reporting models with March data only (Panel A), pooled data with a linear trend (Panel B), and pooled data without baseline lag (Panel C). All regressions include strata fixed effects, and regressions in panels A and B include the baseline value of outcome variable. Standard errors clustered at the block level are reported in parentheses, with statistical significance is denoted as: $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$. q values adjusted to control the false discovery rate across the five commodities are reported in brackets.

Table B.2: Effects on quantities received using alternative specifications

	Rice	Wheat	Sugar	Salt	Kerosene
	(1)	(2)	(3)	(4)	(5)
Panel A: March only					
Treatment	1.5***	-.75	.0084	.13	.034
	(.55)	(.47)	(.11)	(.13)	(.083)
	[0.10]	[0.64]	[1.00]	[1.00]	[1.00]
Control mean	16.28	2.68	0.82	0.95	1.71
Observations	3,517	3,553	3,551	3,575	3,533
Panel B: Pooled data with linear trend					
Treatment	-.15	-.29	.14	-.097	-.1
	(.77)	(.53)	(.27)	(.2)	(.08)
	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
Month	-.42*	.071	.1	.14*	-.04*
	(.22)	(.08)	(.071)	(.075)	(.021)
Treatment X Month	.45	-.14	-.059	.077	.035
	(.28)	(.089)	(.12)	(.099)	(.039)
Control mean	16.81	2.59	0.72	0.81	1.76
Observations	10,557	10,654	10,670	10,726	10,618
Panel C: Pooled data with no baseline lag					
Treatment	.53	-.5	.018	.069	-.019
	(.6)	(.52)	(.076)	(.064)	(.051)
	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
Control mean	16.81	2.59	0.72	0.81	1.76
Observations	10,557	10,654	10,670	10,726	10,618

This table reports alternative specifications for Panel B of Table A.4 by reporting models with March data only (Panel A), pooled data with a linear trend (Panel B), and pooled data without baseline lag (Panel C). All regressions include strata fixed effects, and regressions in panels A and B include the baseline value of outcome variable. Standard errors clustered at the block level are reported in parentheses, with statistical significance is denoted as: $^{*}p < .10,^{**}p < .05,^{***}p < .01$. q values adjusted to control the false discovery rate across the five commodities are reported in brackets.

Table B.3: Effects on overcharges using alternative specifications

	Total	Rice	Wheat	Sugar	Salt	Kerosene
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: March only						
Treatment	.7	.033	-.11*	.88***	-.12	-1.4
	(.61)	(.25)	(.063)	(.31)	(.078)	(1.4)
		[1.00]	[0.53]	[0.06]	[0.53]	[0.91]
Control mean	7.74	1.12	0.22	0.94	0.32	5.18
Observations	3,184	3,391	3,438	3,418	3,447	3,377
Panel B: Pooled data with linear trend						
Treatment	-7.3	.14	-.15***	-6.6	.25**	.11
	(4.8)	(.28)	(.056)	(4.5)	(.11)	(.95)
		[1.00]	[0.11]	[0.55]	[0.18]	[1.00]
Month	-.63**	.011	.0071	.16	.13**	-1***
	(.27)	(.032)	(.02)	(.16)	(.055)	(.19)
Treatment X Month	2.4	-.038	.0067	2.3	-.12*	-.38
	(1.7)	(.062)	(.022)	(1.6)	(.062)	(.59)
Control mean	8.22	1.08	0.22	0.91	0.17	5.95
Observations	9,623	10,183	10,317	10,260	10,375	10,185
Panel C: Pooled data with no baseline lag						
Treatment	-2.2	.11	-.13**	-1.8	.019	-.59
	(1.8)	(.29)	(.056)	(1.5)	(.035)	(.49)
		[1.00]	[0.23]	[0.92]	[1.00]	[0.92]
Control mean	8.22	1.08	0.22	0.91	0.17	5.95
Observations	9,623	10,183	10,317	10,260	10,375	10,185

This table reports alternative specifications for Panel D of Table A.5 by reporting models with March data only (Panel A), pooled data with a linear trend (Panel B), and pooled data without baseline lag (Panel C). All regressions include strata fixed effects, and regressions in panels A and B include the baseline value of outcome variable. Standard errors clustered at the block level are reported in parentheses, with statistical significance is denoted as: $^{*}p < .10,^{**}p < .05,^{***}p < .01$. q values adjusted to control the false discovery rate across the five commodities are reported in brackets.

Table B.4: Effects on willingness to accept using alternative specifications

	January	February	March	Pooled	
	(1)	(2)	(3)	(4)	(5)
Panel A: All households					
Treatment	-7.6	-23	-1.8	-11	-19
	(27)	(28)	(29)	(26)	(29)
Month					-22***
					(5.9)
Treatment X Month					4
					(7.2)
Control mean	1,045	1,041	1,000	1,028	1,028
Observations	3,395	3,522	3,520	10,437	10,437
Panel B: Excludes HHs who did not purchase ration in a given month					
Treatment	23	11	32	22	9.6
	(21)	(24)	(22)	(21)	(26)
Month					-26***
					(7.5)
Treatment X Month					6.3
					(8.9)
Control mean	1,163	1,157	1,111	1,143	1,143
Observations	3,165	3,122	3,102	9,389	9,389

This table reports a robustness check to Panel B of Table 3 by measuring value as the amount a household is willingness to accept in lieu of ration. We report estimates from both the sample that includes (Panel A) and excludes (Panel B) households that did not purchase any ration in a given month. The unit of analysis is the ration card - month. The dependent variable is the household reported willingness to accept (WTA), constructed as the smallest value X for which the respondent reported that they would have preferred in cash to the commodities received. The WTA for ration cards that did not receive any ration is set to zero. All regressions include strata fixed effects. Standard errors clustered at the block level are reported in parentheses, with statistical significance is denoted as: ${}^{*}p < .10, {}^{**}p < .05, {}^{***}p < .01$.

C Reconciliation protocol and implementation

As provided to us, the protocol governing the disbursements of reconciled commodities (i.e. wheat and rice) which the government of Jharkhand introduced was given by

$$
\begin{aligned}
D_t &= max(0, E_t + C_t - S_{t-1}) && (1)\\
S_t &= S_{t-1} + \hat{D}_t - O_t && (2)\\
C_{i,t} &= \begin{cases} E_{i,t-1} - (O_{i,t-1} - C_{i,t-1}) & O_{i,t-1} > C_{i,t-1} \\ E_{i,t-1} & O_{i,t-1} \leq C_{i,t-1} \end{cases} && (3)
\end{aligned}
$$

Equation 1 defines the amount D_t to disburse at the beginning of period t as a function of the amount $E_t = \sum_i E_{i,t}$ to which recipients assigned to the FPS in period t were entitled, the amount $C_t = \sum_i C_{i,t}$ is the total carryover commitment owed to recipients in period t because they did not collect their entitlement in period $t-1$, and the amount of stock S_{t-1} the government believes the FPS should have been holding at the end of the preceding month. Equation 2 defines the law of motion for stock, which increases with disbursements D_t and decreases with offtake $O_t = \sum_i O_{i,t}$ by beneficiaries. Offtake is interpreted as first accruing against carryover commitments from the previous period until these have been exhausted, and then accruing against current period commitments.

C.1 Adherence

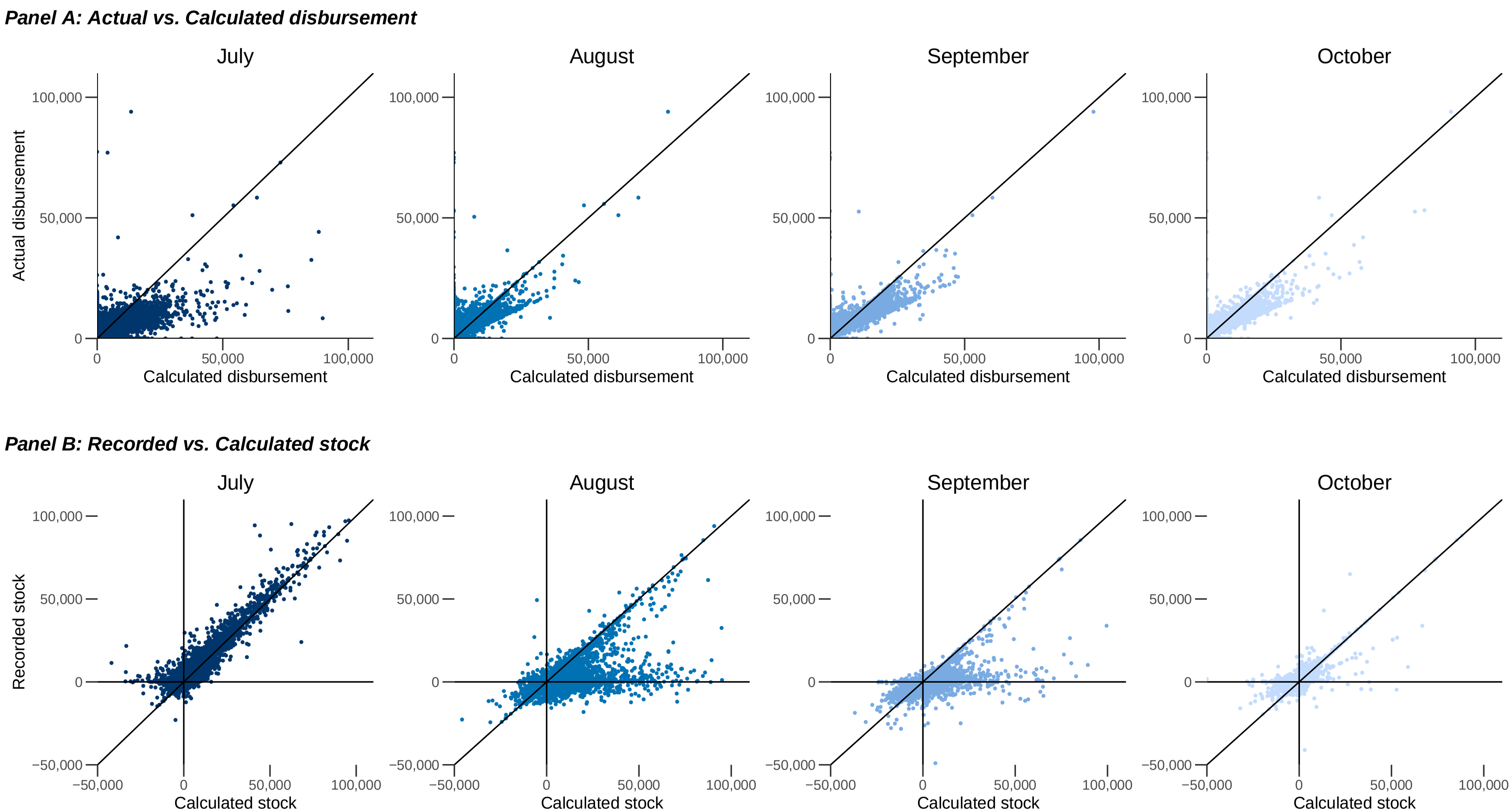
We next examine how closely the government adhered to this protocol during the four-month period of reconciliation. To do so we examine two sources of data: measures of the aggregates S_t, E_t, C_t, D_t and O_t which the National Informatics Commission maintained for each Fair Price Shop and month, and transaction-level data directly from the ePOS devices themselves which record $O_{i,t}$ for each household and month.

According to its own records, the government was fairly flexible in its implementation of reconciliation. Figure C.1 illustrates this, showing scatterplots of the left- and right-hand sides of Equations 1 and 2 for the months July-October 2017 and for wheat and rice pooled. For both relationships, the actual quantities on record are positively associated with those we obtain by mechanically recalculating them using the reconciliation formulae, but with substantial noise. It is particularly striking the extent to which the government held dealers accountable for less stock than implied by formula in August and September, and that in many cases they recorded a *negative* closing stock, something that should not have been possible if dealers accurately reported transactions.

Indeed, there is some evidence that the government did not take transaction data on offtake by beneficiaries at face value in all cases. Panel A of Figure C.2 plots the relationship between offtake as recorded in the official aggregates and our own independent measure of offtake calculated from the raw transaction data. In the majority of cases the two coincide exactly (85%) or are within 1% of each other (92%), but in other cases the transaction data imply higher offtake than the government aggregates acknowledge. This suggests the government may have suspected that the transaction data were over-stating actual offtake, as for example would be the case if FPS dealers asked beneficiaries to "sign" for more grain than they actually received. Indeed, Panel B – which restricts focus to our sampled ration cards – illustrates that beneficiaries generally report receiving less grain than the transaction data say they did.

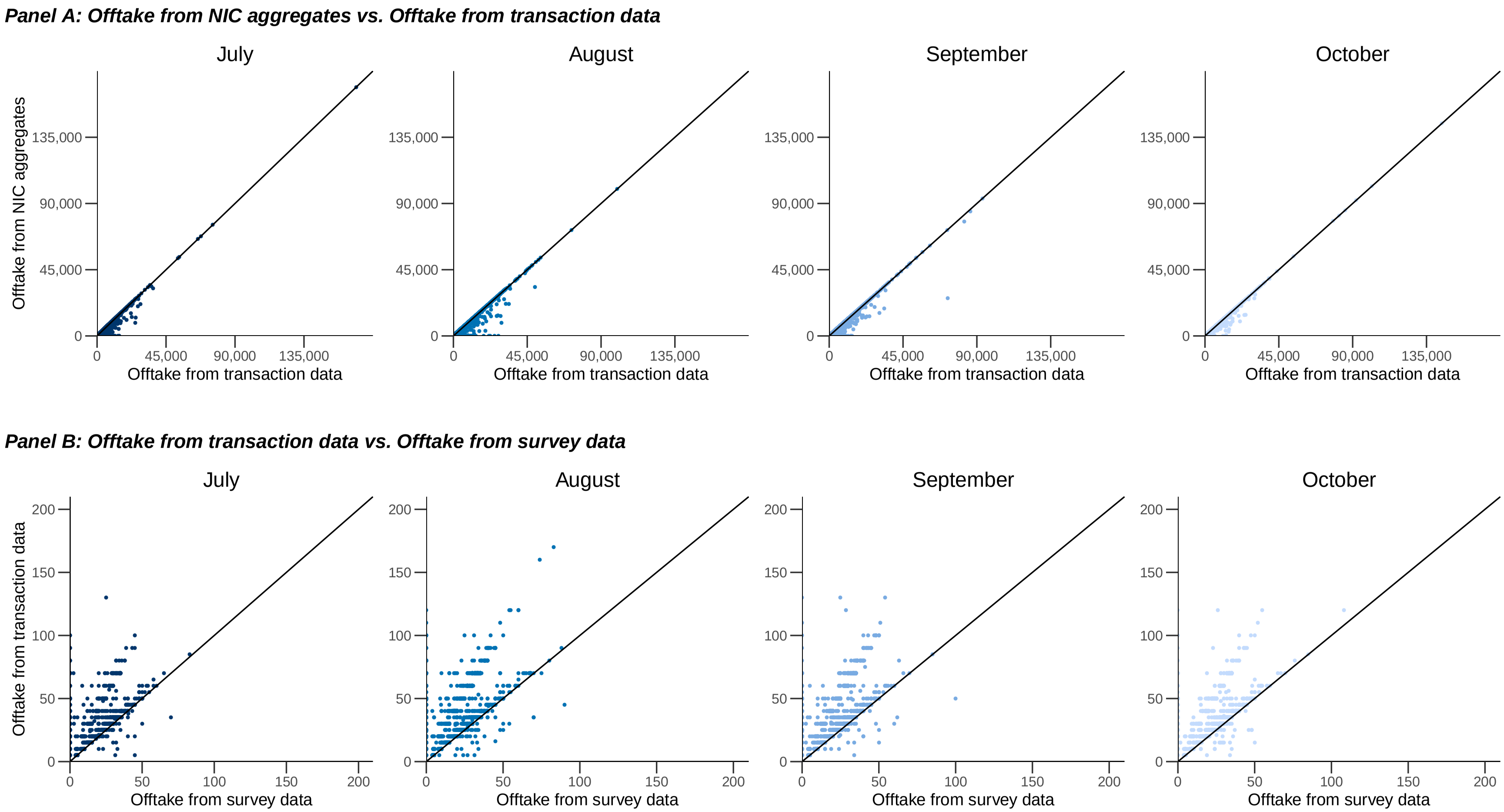
Overall, reconciliation as actually implemented was somewhat less punitive than reconciliation strictly by the book would have been. We illustrate this in Figure C.3, where we plot the evolution of actual grain disbursements per FPS against the counterfactual time series obtained by (i) recalculating disbursements assuming that the government implemented reconciliation strictly, while (ii) leaving unchanged the reported offtake of grain. If the government had implemented Equations 1 and 2 and disallowed negative values for stock, we estimate that grain disbursements would have been lower than actually observed in all four months. If it had instead allowed stock to take negative values, disbursements would initially have fallen more than they did but eventually rebounded to be even higher than entitlements, reflecting the implausibly large disbursements some ration shops were reporting and the resulting negative stock values they accumulated.

Figure C.1: Adherence of disbursement and stock to reconciliation policy, by month



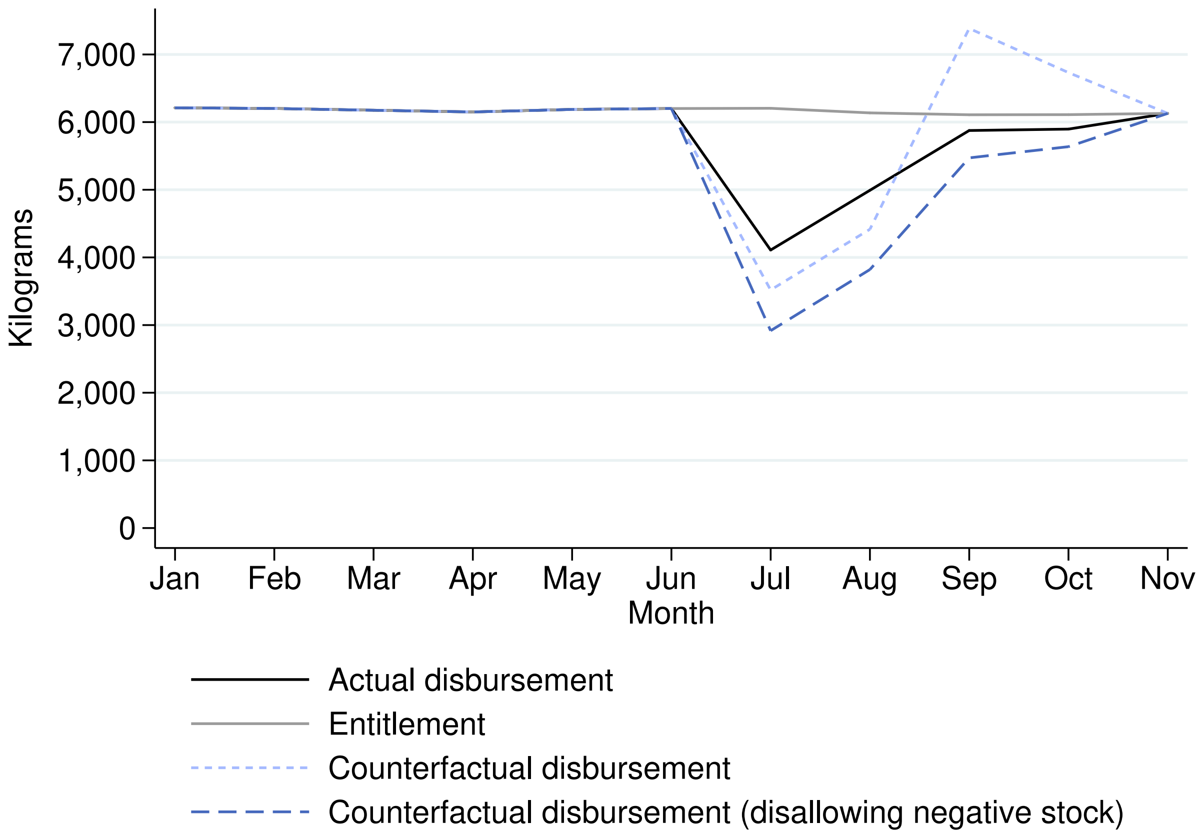
This figure shows the scatterplots of the left- and right-hand sides of Equations 1 and 2for the months July-October 2017 using aggregate measures obtained from NIC data in kilograms. Panel A has actual disbursements on the y-axis and uncensored calculated disbursements on the x-axis, and the Panel B recorded stock on y-axis and calculated stock on x-axis.

Figure C.2: Comparison of offtake from three data sources, by month



This figure compares offtake from three data sources: NIC aggregates provided by the government, transaction data from ePOS records, and survey data from beneficiaries in kilograms. Panel A plots the relationship between FPS-level offtake as recorded in the official aggregates and our own independent measure of offtake calculated from the raw transaction data. Panel B plots the relationship between ration card-level offtake as recorded in transaction data and reported by beneficiaries.

Figure C.3: Adherence to reconciliation protocol



This figure plots the evolution of actual grain disbursements per FPS against the counterfactual time series obtained by recalculating disbursements assuming that the government implemented reconciliation strictly and while leaving unchanged the reported offtake of grain. "Counterfactual disbursement (disallowing negative stock)" censors negative stock in the calculation.

D List of additional analysis conducted

The following pieces of analysis reported in the paper are additional to those pre-specified:

- Table 1, which reports the representativeness of our study area with Jharkhand;
- Panel B of Table 2, which compares measures of program implementation in treated and control areas;
- Table 7, which examines impacts on the extensive margin of value received;
- Table 8, which reports heterogeneous effects by Aadhaar seeding;
- Figure 3, which plots the distribution of value received in treated and control areas;
- Figure A.2, which compares household income and education levels between seeded and unseeded households;
- Table A.1, which compares the dealers we originally sampled to those we added to the sample as households were re-assigned to them;
- Figure A.3, which plots the evolution of value disbursed and received in treated and control areas;
- Figure 5, which plots the distribution of grain stocks as of June 2017 by treatment status.