Identity verification standards in welfare programs: experimental evidence from India*

Karthik Muralidharan[†]
UC San Diego

Paul Niehaus[‡]
UC San Diego

Sandip Sukhtankar[§]
University of Virginia

August 8, 2020

Abstract

How should recipients of government transfers prove their identity? ID requirements must trade off reductions in fraud and corruption against the risk of denying benefits to legitimate beneficiaries. We evaluate reforms that integrated more stringent, biometric ID requirements into India's largest social protection program, using large-scale randomized and natural experiments. Theoretically, the effects depend on the underlying mechanics of corruption. Empirically, corruption fell but with substantial costs to legitimate beneficiaries: approximately 2 million of whom lost access to benefits at some point during the reforms. 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.

JEL codes: D73, H53, O30, Q18

Keywords: identity verification, corruption, service delivery, biometric authentication, PDS, India, Aadhaar

*We thank Prashant Bharadwaj, Michael Callen, Gordon Dahl, Lucie Gadenne, Siddharth George, Roger Gordon, Ashok Kotwal, Lee Lockwood, 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, Thomas Brailey, Vaibhav Rathi, and Xinyi Liu. Finally, we thank the Bill and Melinda Gates Foundation (especially Dan Radcliffe and Seth Garz) for the financial support that made this study possible.

[†]UC San Diego, JPAL, NBER, and BREAD. kamurali@ucsd.edu.
[‡]UC San Diego, JPAL, NBER, and BREAD. pniehaus@ucsd.edu.
[§]University of Virginia, JPAL, and BREAD. sandip.sukhtankar@virginia.edu.

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 trade-off between errors of *inclusion* whereby benefits are granted to non-eligible recipients against errors of *exclusion* whereby legitimate beneficiaries are (fully or partly) denied benefits to which they are legally entitled. Managing this trade-off is especially challenging in settings with corrupt intermediaries who can exacerbate both types of errors: adding fake beneficiaries to rosters, for example, or taking a cut from transfers to legitimate ones.

This challenge 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 programs 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 issued unique identification ("Aadhaar") numbers linked to biometric records to over 1.24 billion people and is integrating Aadhaar-based biometric authentication into a range of applications. Whether Aadhaar should be mandated to receive welfare benefits has been controversial, contested all the way to India's Supreme Court. Proponents argue that this is necessary to prevent fraud, while critics argue that it denies people their legal entitlements and thus "undermines the right to life" (Khera, 2017). In a September 2018 ruling the Court allowed the government to mandate 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 experimental evaluation of introducing Aadhaar as a requirement to collect welfare benefits. Specifically, we examine how this introduction affected government expenditure, leakage, and beneficiary receipts 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). PDS "ration card" holders are entitled to purchase fixed monthly quantities of grain and other

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

commodities at a highly-subsidized price from a government-run Fair Price Shop (FPS). In practice, the resulting dual-price system creates strong incentives for corrupt intermediaries to divert grains to the open market; the most recent nation-wide estimate is that 42% of grain was diverted as of 2011-12 (Dreze and Khera, 2015).

Aiming to reduce leakage, the government reformed PDS implementation in two phases. In the first phase it installed electronic Point-of-Sale (ePOS) machines in FPSs and required beneficiaries to obtain an Aadhaar number, link (or "seed") it to their PDS account, and authenticate by scanning their fingerprints each time they transacted. We refer to this reform as "Aadhaar-based biometric authentication", or ABBA. ABBA generated a digital record of transactions for which beneficiaries had "signed" biometrically, in contrast with the status-quo approach of authentication based on physical ration cards and record-keeping on paper. In the second phase, the government began using these records to adjust downwards (or "reconcile") the amounts of grain it disbursed to each FPS each month to reflect the amount the FPS should still have in stock. This contrasted with the status quo approach, which was to disburse the full amount of grain needed to satisfy the entitlements of all beneficiaries.

Theoretically the effects of these reforms are ambiguous and, as we formalize in Section 3, depend among other things on the form of corruption in the status quo. They should be relatively effective at reducing disbursals and leakage on the accounts of beneficiaries that do not demand transfers, either because they cannot use them (e.g. in the case of out-migrants) or because they are in fact "ghost beneficiaries" created by agents for the sole purpose of extracting rents. However, effects on beneficiaries who *do* demand transfers but receive less than their entitlement under the status quo are ambiguous since their leverage to demand those entitlements may either increase or decrease. In addition, reduction of disbursals against the accounts of those who do not collect their benefits could spill over to those who do try to collect them (similar to the corruption displacement shown by Yang (2008)). Finally, ABBA could mechanically increase exclusion errors and transaction costs if some beneficiaries are unable to seed their ration cards or authenticate at the point of sale. It is thus a priori unclear whether authentication and reconciliation would allow the government to squeeze rents from the system without also harming beneficiaries.

To evaluate this empirically we worked with the government of the state of Jharkhand (GoJH) to randomize the order in which it introduced ABBA across 132 sub-districts in 10 districts. Our evaluation sample is representative by design of 15.1 million individuals in 3.3 million beneficiary households in 17 of Jharkhand's 24 districts, and representative on observables of the rest of the state. GoJH then launched reconciliation simultaneously in all areas two (eleven) months after deploying ePOS devices in control (treatment) areas. 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 (and also using as a placebo other PDS commodities not subject to reconciliation). The integration of Aadhaar into the PDS was implemented by GoJH as part of a full-scale deployment that was being rolled out across the country. Thus, our design allows us to directly estimate policy-relevant parameters of broad interest.

Our focal outcomes are the value of goods disbursed by the government, value received by beneficiaries, and the difference between these (i.e. leakage). We measure these using comprehensive administrative data on disbursals of commodities to all ration shops matched to original survey data on commodity receipts and transaction costs collected in four rounds (one baseline and three follow-up rounds) from a panel of 3,840 PDS beneficiaries. The first follow-up was used to study the effects of ABBA, and the remaining to study reconciliation.

At the time of the first follow-up, we estimate that the average leakage rate in the control group was 20% of value disbursed, driven primarily by 15% of beneficiaries getting zero benefits in any given month. Of these, at most 3% were outright "ghosts"; the majority were real beneficiaries who reported that they were unable to collect their benefits.

GoJH implemented ABBA quickly and thoroughly, and largely complied with the experimental design. Six months after treatment onset, 95% of beneficiary households in treated areas had at least one member with an Aadhaar number seeded to the PDS account, and 91% reported that transactions at their FPS were being authenticated, while only 6% of control households reported the same. ITT estimates can thus be reasonably interpreted as those of ABBA.

ABBA by itself had small effects on average. It did not decrease (and if anything slightly increased) government spending, and did not substantially change mean value received by beneficiaries or mean leakage. We also find no meaningful changes in measures of the quality of goods received, market prices, or beneficiaries' food security. However, beneficiaries did incur 17% higher flow transaction costs to collect their benefits (a Rs. 7 increase on a base of Rs. 41), due in part to a doubling in the number of unsuccessful trips to the FPS.

Yet for a minority of beneficiaries ABBA had substantial negative impacts. The probability that a beneficiary received no commodities at all in a given month increased by an estimated 2.4 percentage points ($p < 0.1$), implying that nearly 300,000 people lost access to benefits. Exclusion increased significantly among the 23% of households who had not "seeded" their ration cards at baseline: the mean value of rice and wheat received fell by Rs. 49, or 10.6%, and the probability of receiving none of these commodities increased by 10 percentage points. This pattern of incidence is regressive, as unseeded households tend to be poorer and less educated than their seeded peers. These results are consistent with the critique that ABBA per se caused at least some "pain without gain" (Dreze et al., 2017).

A potential counterargument is that authentication was a necessary first step towards reconciliation. Consistent with this, treated dealers reported 18% lower profits and a 72% lower expected future bribe price for FPS licenses, suggesting that they expected a relatively large fall in future rents reflecting the impending introduction of reconciliation. In line with this, we find that reconciliation coincided with a sharp 37% drop in value disbursed by the government in treated areas. However, only 66% of this was a reduction in leakage, and the remaining 34% represented a drop in value received by beneficiaries. On the extensive margin we estimate that an additional 1.6 million people did not receive PDS benefits in July 2017 (the first month of reconciliation). Thus, while reconciliation did meaningfully reduce dealer rents, dealers were able to pass on a considerable amount of pain to beneficiaries.

This large pass-through likely reflects to some extent the fact that the government reduced disbursals when it introduced reconciliation to reflect stock that, according to ABBA records, dealers should previously have accumulated (but in practice had likely diverted). If we focus on the control group, which had accumulated only two months of these balances, we find more modest effects: a 19% (Rs. 92 per ration card) initial reduction in value disbursed, of which 22% represents a reduction in value received and the other 78% a reduction in leakage. Going further and using our experimental variation to predict the impacts we would have observed had dealers held *no* opening balances, we estimate a reduction in leakage combined with a small, insignificant *increase* in value received by beneficiaries. This suggests that reconciliation per se has the potential to reduce corruption without harming beneficiaries, though as actually introduced it generated large-scale exclusion.[2]

The trade-off between leakage reduction and exclusion errors may change over time.[3] Thus, our findings should not be interpreted as representing an immutable set of trade-offs between leakage reduction and exclusion errors. Rather, they highlight the possibility that even well-intentioned reforms supported by international development agencies and carried out by a democratically-elected government can generate substantial social costs with over two million legitimate (and disproportionately vulnerable) beneficiaries losing access to essential benefits at various points during the roll-out. These issues are relevant to several countries around the world considering similar reforms, and we discuss specific ideas for improving the trade-off between leakage reduction and exclusion errors in the conclusion.

Our main contribution is to provide the first experimental evidence on the trade-off be-

[2]Indeed, reconciliation proved very unpopular, forcing GoJH to pause it after four months in the face of complaints from both dealers and beneficiaries before reintroducing it the following year. We discuss this and present estimates of impact over the entire 4-month period of reconciliation in Section 6.

[3]For example, exclusion errors and transaction costs may attenuate in the longer run as Aadhaar seeding becomes more widespread and connectivity improves (though treatment effects appear stable during the window we study). Similarly, reconciliation may not generate as much "pass through" of the pain of reduced stocks from dealers to beneficiaries in a longer-run steady state.

tween leakage reduction and exclusion errors from introducing stricter ID requirements for receiving welfare benefits, and to do so in the context of the largest welfare program (PDS) in the country with the largest biometric ID system in the world (India).[4] Further, the structure of the principal-agent problem we study is common around the world, and our conceptual framework highlights ways in which the effects of similar reforms may vary with context. More broadly, we 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), the complexity of the process of obtaining benefits affects their incidence; here, "complexity" does not appear to have been an effective screening device as the households excluded generally appear less well off socioeconomically.

Second, our results highlight the importance of accounting for the economic incidence of anti-corruption policies in addition to their fiscal consequences. Unlike reforms in which governments attempted to crack down on corruption while holding program spending fixed (as for example those studied by Di Tella and Schargrodsky (2003), Olken (2007)), the reforms in Jharkhand aimed to squeeze out intermediary rents and thereby *reduce* program spending, while keeping net benefits unchanged. However, because dealers had some ability to pass these spending cuts on, the intervention had not only direct costs (e.g. procuring machines) but also substantial indirect ones through exclusion. In this sense the results also illustrate how and when controlling corruption may incur substantial "shadow costs."[5]

Finally, our findings inform research and policy discussions on using technology to improve governance and state capacity in developing countries. In prior work we found that introducing biometric authentication of payments in rural welfare programs in the state of Andhra Pradesh (AP) both reduced leakage *and* improved the beneficiary experience (Muralidharan et al., 2016). The impacts of ABBA and reconciliation in the PDS in Jharkhand were quite different, likely reflecting differences in both the nature of the underlying leakage problem and in policy priorities as reflected in the details of intervention design (as discussed further in the conclusion). The contrasting results caution against simplistic characterizations of the effects of new technologies such as biometric authentication, without reference to details of program and intervention design. In other words, discussions of external validity need to pay attention to differences in program "construct" as well as context.[6]

[4]While there is a large literature on how to *target* people to be put on program beneficiary lists (Alatas et al., 2012; Niehaus et al., 2013; Alatas et al., 2016), there is much less evidence on the distinct question of how to *identify* them at the point of receiving benefits.

[5]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 worsened some local public health outcomes. More broadly, a long tradition of work has argued that efforts to control corruption may generate indirect costs by (for example) slowing down decision-making, precluding the use of "soft information," or deflecting attention from other important matters (Klitgaard, 1988; Wilson, 1989).

[6]There are close parallels with the literature on education technology, where impacts on learning outcomes

The rest of the paper is organized as follows. Section 2 describes the context and reforms; Section 3 characterizes potential effects of these reforms theoretically; Section 4 presents the research design, data collection, and estimation strategy; Section 5 reports impacts of ABBA; Section 6 reports impacts of reconciliation; and Section 7 offers a concluding summary.

2 Context and intervention

India was ranked 102 out 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 the main program by which the Government of India aims to provide food security to the poor. Through a network of over 527,000 ration shops known as "Fair Price Shops" (FPS), it delivers highly-subsidized wheat and rice to targeted households on a monthly basis, and other commodities such as sugar, salt, and kerosene on an occasional basis.[7]

Under the National Food Security Act (NFSA) 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. The NFSA entitles eligible households across India to 5kg/month of highly-subsidized grain for each member of the household. Rice and wheat are provided at a price of Rs. 3/kg and Rs. 2/kg through the FPSs, which is a 80-90% subsidy relative to market prices. Some states (though not Jharkhand) use their own budgets to augment benefits further through a combination of expanding eligibility, increasing quantities, and lowering prices even further. Overall, the PDS costs roughly 1% of GDP to operate.[8]

In part because it creates a dual-price system, distributing commodities at prices well below their market prices, the PDS has historically suffered from high levels of corruption. Commodities "leak" from the warehouses and trucking networks meant to deliver them to the FPS, or from the shops themselves. At the retail level, dealers have been reported to adulterate commodities, over-charge for them, or provide beneficiaries with less grains than their legal entitlement. 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).

have been found to vary widely as a function of design details (Muralidharan et al., 2019a).

[7]Throughout the paper, we use the term "disbursal" to refer to commodities sent by the government to FPS dealers, and the term "distribution" to refer to commodities provided by FPS dealers to beneficiaries.

[8]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`. Accessed on 5 March, 2018.

To help address these issues, the Government of India has introduced several reforms including the use of electronic point-of-sale (ePOS) devices to process and record transactions between dealers and beneficiaries and enable reconciliation of shipments to dealers. Rollout of these devices was well underway throughout India by the time GoJH began its deployment; as of July 2016 an estimated 23% of India's FPSs had received devices, rising to 54% by December 2017.[9]

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. As of June 2019, it had issued Aadhaar numbers to 1.24B people, or 91% of the country's population.[10] Investments in ID 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 the government has touted Aadhaar as an enabling technology for reforms to a wide range of 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), in part because they did not differentiate between real reductions in leakage and increased exclusion of legitimate beneficiaries. In an important study, Barnwal (2019) uses difference-in-differences to show that delivering petroleum subsidies to Aadhaar-linked bank accounts led to a substantial reduction in government expenditure. The absence of matched household survey data, however, makes it difficult to rule out the possibility that there was also some increase in exclusion. To our knowledge there has been no experimental evidence to date on the impacts of an Aadhaar deployment in a welfare program, and also no well-identified study on the impacts of Aadhaar with *matched* administrative and household data on disbursals and receipts.

Jharkhand is a relatively challenging environment in which to roll out an ambitious reform such as ABBA. On state capacity, it ranked 17th out of 19 major states on the most recent Governance Performance Index (Mundle et al., 2012). 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 ranked low among states on rural teledensity (40 telephone or mobile phone connections per 100 people in rural Jharkhand as of 31 October 2017, 19th

[9] 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.

[10] For statistic on number of Aadhaar UIDs generated, see https://uidai.gov.in/aadhaar_dashboard/india.php. For total population statistics, see https://data.worldbank.org/indicator/SP.POP.TOTL.

out of 19 major states) and in the middle for Aadhaar penetration (93% penetration as of 31 December 2017, 17th out of all 36 states).[11]

2.1 Aadhaar-based biometric authentication

Prior to the reforms, 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, but these records were typically not audited.

In August 2016, GoJH began introducing ePOS devices in FPSs to authenticate beneficiaries when they came to collect their rations. The state gave each dealer a device configured to operate in one of three modes, depending on connectivity: online (81% of shops), offline (15%), and partially online (4%). 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 individual 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 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. In *offline mode*, the device captured and stored fingerprint and transaction information, to upload later from a location with connectivity, but did not perform real-time authentication checks. In *partially online mode*, the device functioned in online mode if it detected a network connection and in offline mode otherwise. Dealers did not have discretion to select modes. Our experimental design (below) randomized receipt of a device but not device mode, so that reported treatment effects are an average of mode-specific effects given the assignment policy.

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 month's entitlement and any uncollected balance from the previous month. Unclaimed entitlements from two or more prior months would lapse. After

[11] 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.

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 hearing transaction details (volumes could be turned down). In any case, the digital ledger maintained in the device became the source of truth for balance information from the government's perspective. The government accessed this data in real time in the case of online devices, and dealers were expected to upload and synchronize data within 48 hours of a transaction in the case of partially online and offline devices.

GoJH launched ABBA after Aadhaar-seeding rates exceeded 75% and aimed to complete the process of seeding Aadhaar numbers to ration cards shortly thereafter.[12] As of May 2016 (three months prior to ABBA launch), 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 92.6% by October of 2016 and to 99.8% and 99.5% by May of 2018 (roughly one year after the period we examine experimentally).

2.2 Reconciliation

Prior to the introduction of ABBA using ePOS devices, 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 regardless 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 in active use in 93% of FPSs in our study area, including those in control blocks, where they were rolled out during April and May. Starting in July, GoJH began reconciling its disbursements of rice and wheat, though not of sugar, salt or kerosene. The full formula used to determine disbursements under this regime is in Appendix D. 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, and then disburse the difference between these quantities.

[12]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 or by applying at the block or district office.

While GoJH introduced reconciliation at the same time in both treatment and control blocks, the fact that it held dealers responsible for their full digital transaction history – starting from the time the FPS first used an ePOS device – implies that reconciliation may have had differential effects in these blocks. Specifically, treated blocks had been using ePOS devices for substantially longer than control blocks, and therefore should according to the digital transaction records have accumulated larger balances of grain for which GoJH could hold them accountable. Moreover, to the extent dealers anticipated this, we should interpret the experimentally estimated effect of early receipt of an ePOS device as potentially including the effect of the anticipation that transactions recorded using that device would be subject to future reconciliation. In other words, while reconciliation itself was not randomized, the *reconcilability* of transactions was. We return to these issues in interpreting the results below.

3 Conceptual framework

In administering the PDS, GoJH faces a principal-agent problem: it wishes to transfer value to beneficiaries while minimizing diversion by intermediate agents. We present a simple framework to describe this generic "transfer problem" and examine the conditions under which the authentication and reconciliation reforms that GoJH introduced are most likely to achieve a favorable trade-off between reducing corruption and protecting legitimate beneficiaries.

A government uses an agent to make transfers to a group of N beneficiaries indexed by i. In each period t the government sends a quantity $q_i^t \in [0,1]$ to the agent with instructions to deliver q_i^t to beneficiary i. Aggregated across beneficiaries, the government sends $Q^t = \sum_{n=1}^{N} q_i^t$. A proportion γ of beneficiaries do not demand the transfer; these may be "ghost beneficiaries" created by the agent solely for the purpose of diversion, but also beneficiaries who have migrated and no longer live in the area, or who do not find it worth the time and cost to collect inferior rationed goods. The other $1-\gamma$ of beneficiaries who do demand the transfer negotiate with the agent over how to divide q_i^t. The government may then receive a signal that is informative about this division. Either way she then chooses the next period's disbursements q_i^{t+1} and potentially additional actions affecting the players' payoffs.

To illustrate the key conceptual issues we make two simplifications: we focus on the finite-horizon case $t = 1, 2$, and for the most part on the case of a single beneficiary ($N = 1$). We consider the possibility of spillovers across beneficiaries later.

3.1 The status quo

In this case the government sends $q_i^t = 1$ every period, receives no informative signals about the outcome, and takes no other actions.

The agent pockets entirely those transfers directed to the share γ of beneficiaries who do not demand them. Consider next the case of one of the $1-\gamma$ beneficiaries who does demand the transfer. In the second period, this beneficiary will negotiate with the agent over how to share the value $q_i^2 = 1$ they have received. Let $b_i(q) \in [0, q]$ be the amount he receives out of an arbitrary transfer q; in equilibrium $b_i(1)$ will often determine payoffs in which case we will abuse notation by calling $b_i = b_i(1)$. $b_i > 0$ may reflect for example the beneficiary's ability to complain to local leaders if treated too badly, the agent's concern for the beneficiary, etc. On the other hand $b_i < q_i^2$ may reflect for example the dealer's ability to plausibly claim that he received less than full allotment from the government in the first place, or power asymmetry between the dealer and beneficiary. In the first period, the beneficiary will similarly negotiate with the agent over how to share the value $q_i^1 = 1$ they have received. Since the future division of rents is determined, the problem is the same and the beneficiary will again receive b_i.

Overall, the per-period and per-beneficiary average amounts transferred to the agent and to beneficiaries respectively are

$$v_A = \gamma + (1-\gamma)(1-\bar{b}) \tag{1}$$

$$v_B = (1-\gamma)\bar{b} \tag{2}$$

where $\bar{b}$ is the bargaining power of the average beneficiary. From the government's point of view there are two distinct issues. First, some transfers "leak" because the beneficiary for whom they are nominally intended does not demand them – either because he is a ghost or for other reasons. Second, even transfers intended for a beneficiary that demands them leak to the extent that the beneficiary's influence over the agent is weak (with leakage as high as 100% if $b_i = 0$).

One standard approach to reducing such leakage is to keep program expenditure fixed but conduct selective audits and punish the agent directly if he has diverted value (Becker and Stigler, 1974). In introducing authentication and reconciliation, GoJH took an alternative approach: it aimed to measure leakage and then *reduce* program expenditure by the amount missing. The key issue, then, is what the incidence of these reductions will be.

3.2 Authentication with reconciliation

Suppose the government introduces stricter authentication requirements for beneficiaries, as a precursor to reconciling transactions. To the extent agents enforce these requirements, one immediate impact may to exclude some beneficiaries who are unable to meet them. By itself this would be equivalent to an increase in the proportion γ of beneficiaries for whom transfers are undelivered and thus diverted by the dealer, which decreases value received and increases leakage (since disbursals are unchanged).

Next suppose the government asks the agent and the beneficiaries to send reports $\hat{q}_i^t \in [0, 1]$ quantifying the value transferred to each beneficiary i in period t, and that these reports are *authenticated* meaning that both parties must agree to sending one. This captures the fact that after the introduction of ABBA, transaction data could be logged in ePOS devices and remitted to the government only if both parties scanned their fingerprints. Further, suppose the government follows a policy of *reconciliation* meaning that it sets $q_i^{t+1} = \hat{q}_i^t$ (with "no message" interpreted as $\hat{q}_i^t = 0$.)

The agent can still divert in period 1 the full amount of transfers sent to the share γ of beneficiaries who do not demand them, but in these cases cannot generate a message $\hat{q}_i^1$ and thus receives no transfer from the government in period 2. Per-period average leakage thus falls from 1 to 1/2 (and in the unmodelled longer run to 0).[13] Authentication and reconciliation thus unambiguously help to reduce disbursals and leakage due to ghosts, and more generally to beneficiaries who simply do not demand transfers.

Now consider a beneficiary i who does demand the transfer. The effects of authentication and reconciliation on him are ambiguous for (at least) three reasons. First, he now negotiates with the dealer over both (i) a division of the first-period transfer, and (ii) what message to send to the government (which we may simply think of as negotiating directly over q_i^2). The efficient outcome is clearly any division of the current transfer and $q_i^2 = 1$, while if they fail to agree and thus send no message then $q_i^2 = 0$. The beneficiary's influence over this joint negotiation may be either greater or less than over the simpler problem of dividing the current period's transfer, which we would interpret as an increase or decrease in b_i.[14]

Second, the loss of rents from transfers to the share γ of beneficiaries who do *not* demand transfers may generate spillovers in how the agent approaches the negotiation with those who

[13] In the case where the beneficiary is a real, local resident, the agent might be able to induce him to join in submitting a report $\hat{q}_i^1 > 0$. If the beneficiary is truly a "ghost" or migrant, however, biometric authentication precludes this.

[14] Notice that this is not simply because the agent needs to concur with the message: even if the beneficiary could unilaterally choose the message $\hat{q}_i^t$, he could not credibly threaten to report $\hat{q}_i^t = 0$ as this would hurt him in the subsequent period. The key issue is that the government's sole instrument for punishing the agent for reported diversion is to reduce future disbursements, which also hurts the beneficiary, and thus creates a disincentive for him to report diversion in the first place.

do. If for example the dealer derives concave utility from his rents (Equation 1) relative to the benefits of avoiding conflict with beneficiaries, then the loss of these rents will make him more willing to extract compensating rents from others.[15] The result is that cracking down on leakage on the accounts of beneficiaries who do not demand transfers "displaces" some of it into other channels (Yang, 2008). We could represent this formally by extending our notation for the amount beneficiary i receives to $b_i(Q, q_i)$, capturing the fact that it depends not only on q_i, but also on the amounts disbursed for *other* beneficiaries.

Third, the way in which GoJH *introduced* reconciliation could also affect agent and beneficiary outcomes. Because it held dealers responsible for balances they should have accumulated for several months before reconciliation, GoJH sharply reduced disbursals in the first month of reconciliation. If dealers had in fact diverted rather than accumulating these balances, then this fall in Q^t could further reduce the amount $b_i(Q^t, q_i^t)$ that beneficiary i receives at time t (the onset of reconciliation).

The first two channels represent permanent effects of ABBA and reconciliation and we refer to these jointly as the "bargaining" channel. The third one is a transitory channel and we refer to it as the "opening balances" channel. Empirically we observe only $b_i(Q^t, q_i^t)$, which limits our ability to distinguish these channels, but we discuss in Section 6 a strategy for doing so exploiting the initial randomization of ABBA itself.

Taking stock, the framework implies that authentication and reconciliation may help to reduce disbursals and leakage, but that the size of these effects and their cost in terms of exclusion error depend on three factors: (i) the direct risks of exclusion that authentication introduces, (ii) the share of beneficiaries on the official rolls who are ghosts or otherwise do not demand transfers, and (iii) the impacts of reconciliation on the (relative) bargaining power of beneficiaries who *do* demand their transfers. We return to these issues in discussing and interpreting the results.

4 Research design

The design follows a pair of pre-specified and pre-registered analysis plans, one for ABBA and one for reconciliation.[16] Appendix B contains results from any analysis pre-specified in these plans that we do not discuss in the paper, and also lists all analyses in the main text that are additional to those pre-specified. Below we summarize randomization, sampling,

[15] One plausible reason for concave utility in rents is that dealers often pay bribes to obtain PDS licenses (which we present evidence of later) and may have planned to generate a certain amount of rents each month to cover these costs, especially if they have borrowed to make those bribe payments.

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

and data collection methods, with further details in Appendix C and full details in the pre-analysis plans.

4.1 Randomization

To obtain policy-relevant estimates of impact, 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.[17]

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.[18] From the remaining 17 districts, home to 24 million people and 15.1 million PDS beneficiaries, we randomly sampled 10 districts within which to randomize the rollout of the intervention. This design ensures representativeness of the 17 districts in our frame. In practice our 10 study districts appear fairly comparable on major demographic and socio-economic indicators to all the 14 remaining districts of Jharkhand (Table A.1). Our frame is thus arguably representative of the full population of 5.6 million PDS households and 26 million PDS beneficiaries in the state.

Finally, we assigned treatment to large units. We randomized the rollout at the level of the sub-district ("block"), which on average covers 73 FPSs and 96,000 people. Figure A.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. Treatment and control blocks are similar in terms of demographic and program characteristics, as one would expect (Table 1, Panel A). Of 12 characteristics we examine, one is marginally significant at the 10% level.

The evaluation was conducted within the context of a full-scale rollout, as GoJH deployed ePOS devices to 36,000 ration shops covering the entire population of 26 million PDS beneficiaries in the state. This deployment involved a major effort by the government and was

[17] 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), Egger et al. (2019), and Muralidharan et al. (2020a).

[18] The intervention had already started in the capital Ranchi district at least three months prior to our experiment, allowing for any initial glitches in implementation to be resolved.

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.

Consistent with this commitment, we find that GoJH complied closely and quickly with the treatment assignment (Figure 1 provides the timeline of the intervention rollout in treated and control blocks). 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 1, Panel B). 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 to process transactions, largely reflecting early rollout in one control block.[19] 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.

4.2 Sampling and Data Collection

Our data collection focused on measuring the value of commodities disbursed by the government and the value received by beneficiaries (net of price paid), as well as the real transaction costs incurred by dealers and beneficiaries to transfer this value. Leakage is the difference between value disbursed and value received.

To measure these quantities we begin with administrative records. We obtained information on monthly quantities of commodities disbursed to all FPSs from the National Informatics Centre (NIC),[20] and the administrative database of eligible PDS beneficiaries and their assignment to FPSs from GoJH. We used the latter 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 selected 3 FPSs via PPS sampling in each study block, for a total of 396 shops. 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

[19] 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.

[20] 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.

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 using our own equipment the strength of the four major cellular networks at the shop in order to capture connectivity.[21]

For each sampled ration shop we sampled 10 households from the government's list of PDS beneficiaries, 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. We attempted to interview these households for baseline and three follow-up surveys to create a household-level panel. We ultimately identified and interviewed the corresponding household at least once in 97% of cases.

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 1 illustrates the recall window covered by each survey. We use data from follow-up 1, covering January through March, to measure 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.

4.3 Estimation strategy: Aadhaar-based biometric authentication

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

$$Y^t_{hfbs} = \alpha + \beta Treated_{bs} + \gamma Y^0_{hfbs} + \delta_s + \epsilon^t_{hfbs} \tag{3}$$

where Y is an outcome measured for household h assigned to FPS f in block b of stratum s.[22] Regressors include an indicator T for whether that block was assigned to treatment,

[21] 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.2). Note that the reassignment of households to other shops does not affect our ITT estimates because we track the originally sampled households and because their reassignment was to other FPS in the same block, with the same treatment status. Note also that dealers cannot move across FPSs, as they are licensed to operate a specific shop. We also confirm that there is no treatment effect on enrollment of household into the program (coefficient 0.2% of control mean, p-value 0.86).

[22] 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.

the baseline value Y^0_{hfbs} 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.[23] When using survey data we weight specifications by (inverse) sampling probabilities to obtain results that are representative of the sample frame. We use analogous specifications for outcomes measured at the level of the FPS or block. We pool observations for January-March 2017, following our pre-specified plan for dealing with the possibility of non-stationary treatment effects.[24]

Each regression table below reports the percent of the original sample for which outcome data were non-missing. Tables A.3 and A.4 examine missingness by treatment status and generally do not find 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). We report standard errors clustered by block. We report p-values for well-defined summary measures of outcomes such as value disbursed, and for outcomes at the individual commodity level also report q-values adjusted to control the false discovery rate.

4.4 Estimation strategy: reconciliation

GoJH introduced reconciliation in July 2017 across both treatment and control groups simultaneously, and suspended it in November 2017. We examine the effects of reconciliation using this time series variation by estimating the following pre-specified model:

$$Y_{hfbst} = \alpha_{hfbs} + \gamma t + \beta_R R_t + \beta_{Rt} R_t (t - t^*) + P_t + \epsilon_{hfbst} \quad (4)$$

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 all regressors with an indicator for treatment. We

[23] This approach should yield consistent estimates given that $Treated_{bs}$ is experimentally assigned. Abrevaya and Donald (2017) show that if the regressor of interest is correlated with the partially observed covariates it may be necessary for consistency to include interactions between the missingness indicator and all the other regressors, but this is not relevant in our case.

[24] 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 generally do not observe evidence of trends, consistent with the fact that program implementation also appeared to have stabilized, and therefore privilege the pooled estimators. Time-varying estimates are in Appendix B.

report standard errors clustered by FPS.[25] We report results for both reconciled commodities (rice and wheat), and unreconciled ones (salt, sugar, and kerosene), with the latter providing a plausible placebo group.

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, by including an indicator for November we do not impose that outcomes revert immediately to what they would have been absent the intervention. Doing so would significantly improve power if true, but seems 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 that theory suggests reconciliation should generate transitional dynamics.

5 Results: Aadhaar-based biometric authentication

5.1 Value transfer

We measure value (V) as the sum across commodities c 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). Formally,

$$V_{ht} = \sum_c Q_{cht}(p^m_{ht} - p^s_{ht}) \tag{5}$$

Entitlements are meaningful: their mean value evaluated using Equation 5 is Rs. 595 per month, 14% of the national rural poverty line for the average household in our sample.[26]

We first examine leakage in the control group to better understand the counterfactual and map it to our conceptual framework. The government recorded disbursing commodities worth approximately the full entitlement amount: an average of Rs. 579 per month, or 98% of mean value entitled. Beneficiaries received Rs. 463 on average, which is less than their entitlement and implies that roughly 20% by value of the commodities the government did disburse did not reach them. Of this 20%, fairly little is attributable to outright "ghost" beneficiaries.

[25]This is as specified in our pre-analysis plan, and differs from the experimental results where we cluster at the block-level as the identifying variation is at the block level. In A.14, where we do use block-level randomization as an instrument, inference is substantively the same if we cluster at the block level.

[26]Ration limits depend on the size of the household and category of ration card it holds. 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.

To anticipate the effects of authentication and reconciliation, our framework suggests that we should examine how many of these households did not actually demand transfers (and thus contribute to γ) and what was the distribution of b_i among those that did. The bulk of the leakage occurs on the extensive margin, i.e. is attributable to the 15% of households who do not collect benefits in any given month (Figure 2). However, few of these are ghosts: recall that we successfully located 97% of sampled households; of the remaining 3% we confirmed that 1.5% were ghosts and cannot be sure of the status of the other 1.5% (Figure A.2).[27]

Among the remaining households that repeatedly did not collect transfers, we asked *why* this was the case. Only 1 percentage point said that they did not attempt to collect their benefits and that this was because they did not need or require their ration; the rest either attempted to collect their benefits or did not because they did not think they would be successful. We interpret these as beneficiaries that demanded transfers but captured a low share b_i of them. Conditional on receiving positive benefits, mean under-provision was under 6% (5%/85%), implying a mean b_i of 0.94. Overall, these facts suggest that γ was low but the initial distribution of b_i was very heterogeneous, close to 1 for many beneficiaries but close to 0 for a minority. This heterogeneity in turn implies that the effects of authentication and reconciliation are theoretically ambiguous.

5.1.1 Value disbursed

Table 2 summarizes impacts of ABBA on value transfer during January-March 2017, beginning in Panel A with value disbursed by the government. We observe this outcome for the universe of FPSs in our study area (8,924 shops) and therefore use all of these data, with outcomes expressed per ration card $\times$ 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 nets out to a small but significant increase in total value disbursed of Rs. 12 per ration-card month, or around 2%.[28] In any case, there is no evidence that ABBA by itself saved the government money.

[27] We classify a household as a ghost if (i) the survey team cannot locate it after at least two attempts, and (ii) two neighbors warrant that no such household exists. Details are in Appendix C.

[28] By default the government provided rice to rural blocks, but prior to the reform it made exceptions for those that expressed a desire for wheat, providing them with rice and wheat in 3:2 proportions. After the reform it appears to have reduced such exceptions.

5.1.2 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. $[-25.2, 22.8]$, ruling out decreases greater than 4.3% and increases greater than 3.9% of value disbursed. Any effects on value received by the average household were thus small in economic terms.

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. We examine this possibility in Panel B of Table A.5 reports, and see no significant changes in the mean quantity of any commodity received, though 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.6, Panel A).[29]

The quality of goods could decrease if PDS dealers tried to substitute for an anticipated loss of rents on the quantity margin by further adulterating goods (e.g. by adding sand or stones to wheat) or selling spoiled goods (e.g. rotten grains). We test for this in two ways. First, we asked respondents who had completed purchases whether they received adulterated or low-quality goods. Overall, few beneficiaries report experiencing these issues and rates are unaffected by treatment (Table A.7). 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. Second, we elicited respondents' (stated) willingness to accept value of income in lieu of the bundle of goods they purchased at the FPS in each month, a metric which should capture all aspects of both quantity and quality as perceived by beneficiaries. We estimate an insignificant effect on this measure of value received, and can rule out effects outside of $[-5.5\%, 3.6\%]$ at the 95% level.[30]

Given these results, we would not expect to see impacts on food security outcomes.[31]

[29] Interestingly, dealers do report facing lower prices, notably for rice (Table A.6, 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.

[30] Two caveats are that this metric measures stated as opposed to revealed preferences, and that this question appeared to confuse a number of respondents, as 48% gave at least one inconsistent answer to our series of binary choice WTA questions.

[31] Any treatment effects through indirect impacts on access to other, non-PDS benefits which also required Aadhaar are likely to be minimal since the difference in Aadhaar registration rates between treatment (96%) and control (92%) was only 4%.

Table A.8 confirms this. We examine two measures of household 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.[32] We see a tightly estimated null effect of treatment, with 95% confidence intervals expressed in control group standard deviations of $[-0.11\sigma, 0.12\sigma]$ and $[-0.11\sigma, 0.09\sigma]$ respectively.

5.1.3 Leakage

Given that value disbursed increased slightly while value received was unchanged, we do not expect to find reductions in leakage. Panel C of Table 2 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.[33] 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 $[-1.7\%, 6.6\%]$.

The figures in Table 2 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, as they are based on the difference between market and statutory ration shop prices. We examine overcharging separately in Panel D of Table A.6. 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 as the intervention did not directly change marginal (dis)incentives for over-charging.

5.2 Transaction costs

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

[32]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`

[33]This approach lets us exploit potential efficiency gains due to covariance in the error terms in the two equations, which differ because they includes the baseline values of the different outcomes.

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. Treatment increased these transaction costs by a small but significant amount: Rs. 7, or around 1.5% of value received and 17% of the control mean (Table 3, Column 1). In Table A.9 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 doubled from 0.13 per household per month to 0.26, 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.

Using dealer survey data, we reject economically meaningful treatment effects on dealer costs of storing and transporting grains (Table 3, Columns 2-3), which is what we would expect given the lack of an impact on quantities. Finally, using official data and budgetary records, we calculate that the cost of ePOS deployment was Rs. 6.2 per ration card per month, which was a 5% increase on GoJH's base cost of Rs. 144 per ration card per month operating the PDS.[34] Thus, overall transaction costs across the government, dealers, and beneficiaries increased by Rs. 13.6 per ration card per month, which represents a non-trivial 7.8% increase on a base of Rs. 175.

5.3 Distributional and heterogeneous effects

Though average effects on beneficiaries are not significantly different from zero, the main risk involved in rolling out stricter identification requirements is that a subgroup of beneficiaries is unable to meet these and loses access to their benefits entirely. The distributional effects of treatment suggest this was the case. Figure 2 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 4, Column 1), significant at the 10% level.

For a sharper test, we examine how impacts differed for the 23% of households that were "unseeded" at baseline, meaning that no member's Aadhaar number had been linked to

[34]The government paid around Rs. 1,600 per month per ePOS machine to an IT provider inclusive of equipment rental, maintenance, and training. The average FPS in our data has 257 households, yielding an incremental cost 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), 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.

their ration card. We expect these households to be at greatest risk of losing their benefits. Relative to "seeded" households they are also poorer and less educated (Figure A.3), and 5% less likely to be upper caste ($p < 0.01$).

Losses in value received are indeed concentrated among unseeded households (Table 5). The reform lowered value received by this group by Rs. 49 per month, equivalent to 12.6% of their control group mean and significantly different both from zero and 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 the experiment does not identify specific households that counterfactually would not have been excluded, this decrease fully accounts for the overall decrease in the share of households receiving any benefits. Treatment effects on stated willingness to accept are also significantly lower for unseeded households, though not significantly different from zero. Transaction costs, on the other hand, increase slightly more for seeded households, consistent with the idea that they are able to continue transacting with the system, albeit at a higher cost.

Overall, the results suggest that the reform did cause a significant reduction in value received for the households least prepared for the reform, likely driven by the total loss of benefits of a subset of these households. Multiplying the 2.4 percentage point increase in the likelihood that a household in a treated block received no benefits (Table 4, Column 1) by the total number of PDS beneficiaries in treated blocks (6.25 million), we estimate that around 150,000 beneficiaries were likely denied benefits in treated blocks alone. If we extrapolate to include the 7 non-study districts (that our study sample was representative of, and which rolled out ePOS everywhere) in addition to treated study blocks, we estimate that 296,000 beneficiaries were denied benefits due to ABBA alone.[35]

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, income level, and baseline satisfaction with the PDS, 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).[36] In general we find limited evidence of

[35]A potential longer-term consequence of not seeding a ration card is that the government might remove the card from the roster of beneficiaries entirely. Based on non-experimental analysis using data collected after our experiment, we find that unseeded cards were more likely to belong to "ghosts", and were much more likely to be deleted than seeded ones (36% vs. 2%). Yet, because the overall number of ghosts were small, a large fraction (88%) of deleted cards belonged to genuine non-ghost households. While purely descriptive, the deletions data illustrate another way in which attempts to reduce leakage may have come at the cost of exclusion error of genuine households. See Muralidharan et al. (2020b) for further details.

[36]Device mode was not randomly assigned; we identify heterogeneous effects under the assumption that if

heterogeneity along these dimensions (Tables A.10, A.11, A.12, and A.13.). There is some evidence that wealthier and better-educated households receive differentially more value and that wealthier households incur larger increases in transaction costs.

5.4 Impacts on dealers

We asked dealers about their current profitability and expectations of future business prospects (Table 6). Since disbursals did not change, we interpret these effects as reflecting the anticipation of subsequent reforms such as reconciliation that build on it. Dealers report a significant 18% fall in profits from selling rationed commodities. This most likely reflects the imminent onset of reconciliation, which may have *already* reduced treated dealers' incentive to divert grain if they expected (correctly) that doing so would increase their future liabilities under reconciliation. 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 72% (Columns 3-4). We interpret this result cautiously given that it is a sensitive question and only a minority of dealers provided an answer. That said, it again suggests 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. Finally, we also asked about expected payments to renew a license, but these turn out to be negligible and unaffected by treatment (Columns 5-6).

6 Reconciliation

GoJH began allocating its disbursals of rice and wheat based on reconciliation of authenticated transaction records from FPSs in July 2017. Anecdotally, reconciliation was only partially implemented. FPS dealers pressed for and were often granted adjustments or exceptions to offset its effects (we discuss reasons further below). In Appendix D 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 disbursed grain somewhat more leniently than a strict "by the book" implementation would have implied. The results that follow should thus be interpreted as the effect of the rules as actually implemented, net of various adjustments and exceptions.

treated earlier a control FPS would have used the same mode to which it was ultimately assigned. See notes to Table A.12.

6.1 Effects on value transfer

The onset of reconciliation coincided with a sharp drop in both disbursements and receipts of reconciled commodities (rice and wheat), but not of unreconciled ones (sugar, salt and kerosene). Figure A.4 illustrates this, plotting the evolution of value disbursed (Panel (a)) and received (Panel (b)) separately for reconciled and unreconciled commodities. It also overlays the raw data with the fit we obtain from estimating Equation 4 and 95% confidence bands around this fit.[37] For both series our pre-specified functional form fits the temporal patterns quite well. 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 October, before GoJH suspended reconciliation in November (see more below). For unreconciled commodities both value disbursed and received drift slightly downwards over time without any substantial change during the period of reconciliation; if anything they are somewhat higher than trend.

The dynamics of value disbursed are in line with the government's plans. GoJH's view was that dealers should have been holding grain stocks equivalent to the opening balances recorded on the ePOS machines at the start of the reconciliation period and would initially be able to meet obligations to beneficiaries by drawing these down, after which they would require fresh disbursements. As a corollary, disbursements from GoJH to FPSs would fall initially but then gradually rebound, which is what we see in the data. As per the government's intentions, beneficiaries should not have been affected in this scenario as dealers would be able to fully supply grains in spite of the temporarily reduced flow of new disbursements from GoJH by drawing down their retained stocks.[38]

In other words, GoJH hoped to squeeze rents out of the system without adversely affecting beneficiaries. In the data, however, we see that a meaningful share of the drop in disbursements was passed through to beneficiaries in the form of lower value received. In Table 7 we quantify these effects by treatment arm. Figure 3 provides the corresponding plots split by treatment arm.

[37] 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 thus reflects the evolution of quantities.

[38] To illustrate, consider a dealer who distributes 85% of his allocation each month, but continues receiving a 100% grain allocation each month. If ePOS records were available for 8 months before reconciliation, the records would show an opening balance of $8 \times 15\% = 120\%$ of a month's supply and the government would not disburse any grains in month 1 of reconciliation, as the dealer's existing stocks should be sufficient for all registered beneficiaries. If the dealer distributes 85% from those stocks in month 1, his opening balance for month 2 would be $120\% - 85\% = 35\%$, and the government would disburse 65% of a month's stock to bring the stock to 100%. By month 3, a new steady state would be reached in which the government disburses 85% each month as opposed to 100% and the dealer continues to distribute 85% to beneficiaries.

6.2 Why did value received fall?

As outlined in our framework, this fall potentially represents a combination of two effects: an opening balance effect, and a bargaining effect. The former was greater in the treated areas, as they had had ePOS devices for longer and were thus held responsible for a larger opening balance of undisbursed grains as per the ABBA records. Figure A.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 7,716 kg of undistributed grain as opposed to 3,346 kg for control shops ($p < 0.0001$). Thus, disbursals should have fallen more in treated areas, and if treated dealers had already diverted these stocks, beneficiaries would likely get less grain as well.[39] The bargaining effect is ambiguous (as seen in the framework), but should apply equally in treated and control areas.

Consistent with this reasoning, the effects of reconciliation were more pronounced in treated areas (Table 7). Value of reconciled commodities (rice and wheat) disbursed to treated blocks fell by Rs. 182 (or 37% of the pre-reconciliation mean in Table 2), in the first month of reconciliation (Panel A). Of this, value received by legitimate beneficiaries fell by 34% (Rs. 62) (Panel B), while the remaining 66% represents a reduction in leakage (Panel C). On the extensive margin, the share of beneficiaries receiving no value increased by 13 percentage points. Multiplying this figure by the number of beneficiaries in treated areas for which our data are representative, we estimate that *1.6 million* more people did not receive PDS benefits during the first month of reconciliation.[40] Averaging over the full 4-month period of reconciliation, we estimate that disbursals were Rs. 86 (17%) lower per month than they otherwise would have been, and that of this drop 49% was passed on as a reduction in benefits, with the remaining 51% representing a reduction in leakage.

In the control group, on the other hand, value disbursed fell by Rs. 92, or 19%, in the first month of reconciliation (Panel A). Of this, an estimated 22% (Rs. 20) represents a reduction in value received by legitimate beneficiaries (Panel B), while the remaining 78% represents a reduction in leakage (Panel C). In each case these figures are significantly smaller than the corresponding figures for treated areas. On the extensive margin, the share of beneficiaries receiving no value increased by 4.3 percentage points, implying that another 119,000 people lost their benefits in the first month of reconciliation (4.3% * 2.77 million beneficiaries = 119,000 excluded). Averaging over the full 4-month period of reconciliation, we estimate that disbursals were Rs. 46 (9%) lower per month than they otherwise would have been,

[39] In principle, dealers could have purchased grains to make up for stocks they had diverted. In practice, they would have had to pay out of pocket for this and we did not hear any anecdotes of it happening.

[40] This figure includes the 7 non-study districts which also received treatment: 13% * 12.32 million beneficiaries = 1,602,000 excluded.

and that of this drop 34% was passed on as a reduction in benefits, with the remaining 66% representing a reduction in leakage.

To further isolate the effect of the bargaining channel, we can also use our experimental variation as an instrument to extrapolate and predict the impacts of reconciliation if ration shops had held *no* opening balances (or equivalently, if reconciliation had been introduced on a "clean-slate" basis). Specifically, we estimate Equation 4 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 endogenous, but vary systematically with experimental assignment to early versus late ABBA treatment status, we instrument for the opening balance (and its interactions) 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 of the main effect of reconciliation, with opening balances set to zero, as corresponding to the bargaining power effects described in Section 3.

The estimates from this procedure suggest that reconciliation weakly *increased* beneficiaries bargaining power vis-a-vis dealers (Table A.14). 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. Turning to value disbursed, the estimates suggest that value disbursed would have declined by Rs. 59 per ration card (Column 1), substantially less than the drop we see in our data although still greater than zero. In short, this exercise suggests the government could have achieved more modest reductions in leakage without adversely affecting beneficiaries.

These results suggest that the long-run steady state of ABBA and reconciliation may yield positive social returns by squeezing intermediary rents without decreasing beneficiary bargaining power. In the longer run authentication and real-time transaction data also allows the government to make PDS benefits portable and accessible at *any* FPS and not just the assigned one, which would further enhance beneficiary bargaining power.[41] Making the PDS nationally portable is a current priority of the Govt. of India. At the same time, our results illustrate how well-intentioned anti-corruption policies can also generate substantial social shadow costs, as seen by the denial of benefits to 2 million beneficiaries at some point during the reforms (0.3 and 1.7 million during ABBA and reconciliation respectively).

[41]Under portability, the beneficiary's outside option if he cannot agree to a division of surplus with his assigned dealer is to go to another dealer and report receiving his transfer from her. Competition should thus strengthen his bargaining position, as argued by Shleifer and Vishny (1993).

7 Conclusion

Navigating the trade-offs between reducing fraud and corruption on the one hand, and exclusion on the other is a common problem in social programs worldwide. This is especially so in developing countries where programs are implemented by potentially corrupt intermediaries, state capacity is limited, and lives may literally be at stake. In this paper we examine the effect on government expenditure, leakage, and beneficiary receipts of stricter identification requirements, and do so in the context of the world's largest biometric ID system (Aadhaar), and India's largest welfare program (the PDS).

On its own, Aadhaar-based authentication of transactions had no measurable benefit; it slightly increased mean transaction costs for beneficiaries, excluded a minority who did not have IDs "seeded" to their ration cards at baseline from their benefits altogether, and did not reduce leakage. When paired with the new reconciliation protocols, ABBA facilitated a meaningful reduction in government expenditure and leakage but at the cost of concurrent reductions in value received by legitimate beneficiaries. Thus, both the supporters and critics of ABBA and reconciliation in the PDS were correct to some extent. Leakage did fall meaningfully, but at the cost of increased exclusion of genuine beneficiaries.

These results illustrate how the costs of controlling corruption may include indirect "collateral damage" beyond the direct costs of intervening. They also illustrate, as our framework emphasizes, how a government's ability to squeeze rents from corrupt intermediaries without unduly harming beneficiaries themselves depend on the information structure and penalties available to it and on the specific nature of the underlying leakage.

Juxtaposed with prior evidence, and in particular our own earlier work in Andhra Pradesh (Muralidharan et al., 2016), the results in this paper demonstrate that "biometric authentication" can have varying impacts depending on the specific ways it is used. Both deployments of biometric authentication reduced leakage and rents of intermediaries, but in the case of AP the money thus saved was passed on to beneficiaries rather than recouped by the 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 meaningful ways). Biometric authentication thus "worked" in a purely technical sense in both settings, but the distributional consequences varied as a function of design choices. These in turn ultimately reflected political choices: the reforms in AP focused more on improving the beneficiary experience and less on fiscal savings (as seen for example in their relatively generous manual override provision intended to minimize exclusion errors), while those in Jharkhand focused more on reducing fraud and generating fiscal savings.

One broader lesson is the importance of measuring results on the ground during major welfare reforms. In the AP case, it would have been easy to think that there was no impact on leakage because there was no change in government expenditure. Only with matched administrative records and household surveys (and better yet, a control group) could we see that leakage had fallen sharply and that more benefits were reaching people. Conversely, in the Jharkhand case, it would have been easy to interpret the reduction in disbursals as evidence of reduced leakage (and indeed, officials often made this claim). Only with matched administrative data and household surveys could we see that at least some of these reductions were coming at the cost of an increase in exclusion errors.

Our results finding evidence of increased exclusion do not imply that it was a mistake to use ABBA and reconciliation for reducing leakage in the PDS, since leakage did go down meaningfully. However, they do highlight the importance of both building in procedures to guard against exclusion error at the program design phase, and to monitor this during implementation. In the context of the PDS, one could authenticate beneficiary lists periodically rather than authenticating every transaction. Alternatively if authenticating transactions is necessary to enable benefit portability then one could create fallback methods of authentication or override mechanisms.[42] To improve real-time visibility on the last-mile beneficiary experience and enable rapid course correction of policies that may be hurting vulnerable populations, one promising approach is to use outbound call centers to call representative samples of beneficiaries regularly and simply measure whether 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).

[42] Authenticating every interaction may also be appropriate in cases where physical attendance is an important margin; for example, Bossuroy et al. (2019) find that requiring (non-Aadhaar) biometric authentication in health clinics increased adherence, and reduced over-reporting of adherence, to a tuberculosis treatment regimen. Regarding fallback methods, Aadil et al. (2019) find few reports of exclusion from PDS benefits due to Aadhaar authentication in Krishna district in Andhra Pradesh, which they attribute to the availability of several override mechanisms.

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.

Abrevaya, Jason and Stephen G Donald, "A GMM Approach for Dealing with Missing Data on Regressors," *Review of Economics and Statistics*, July 2017, *99* (4), 657–662.

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.

Barnwal, Prabhat, "Curbing Leakage in Public Programs: Evidence from India's Direct Benefit Transfer Policy," Technical Report, Michigan State University 2019.

Becker, Gary S and George J Stigler, "Law Enforcement, Malfeasance, and Compensation of Enforcers," *The Journal of Legal Studies*, January 1974, *3* (1), 1–18.

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.

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.

Di Tella, Rafael and Ernesto Schargrodsky, "The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires," *The Journal of Law Economics*, 2003, *46* (1), 269–292.

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).

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.

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.

Klitgaard, Robert, *Controlling Corruption*, University of California Press, 1988.

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 _, "General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India," Working Paper 23838, National Bureau of Economic Research March 2020.

_, _, and _, "Identity Verification Standards in Welfare Programs: Experimental Evidence from India," Working Paper 26744, National Bureau of Economic Research February 2020.

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

Niehaus, Paul, 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," *The Quarterly Journal of Economics*, August 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.

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

Table 1: 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	.05	.85***	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 FPSs. "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 2: Effects on value disbursed, value received, and leakage

	Total	Rice	Wheat	Sugar	Salt	Kerosene
	(1)	(2)	(3)	(4)	(5)	(6)
Mean entitlement	595	492	18	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	579	417	72	26	9.4	55
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	116	68	19	12	2.1	15
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 FPS $\times$ month and we use the universe of FPSs; in Panels B and C the observation is the ration card $\times$ 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 3: Effects on transaction costs

	Beneficiary costs	Dealer costs		Government costs
	(1)	(2)	(3)	(4)
Treatment	6.9*	.51	.65	6.2
	(3.8)	(.95)	(.63)	–
Adjusted R^2	.09	.10	.28	–
Reference group mean	41	6.8	5.9	144
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 FPS 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 × month incurred by the state government to administer ABBA based on vendor costs in treatment areas only (not a treatment effect). 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 4: 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 5: 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 6: Effects on dealer expectations

	Intends to continue running FPS?		Expected bribes to obtain license?		Expected bribes to renew license?		Profit (self-reported)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	.054	.033	-56,816**	-58,393*	-111	-83	-1,244**	-1,121*
	(.055)	(.059)	(28,561)	(33,370)	(123)	(147)	(531)	(583)
Adjusted R^2	.093	.13	.31	.27	.053	.035	.07	.051
Control mean	.73	.71	76,590	81,188	565	555	5,891	6,113
Observations	437	366	150	127	370	307	445	370
% of sample		92		32		78		93
Sample	Full	Restricted	Full	Restricted	Full	Restricted	Full	Restricted

This table reports estimated treatment effects on measures of FPS dealers' expectations in March about the future. The unit of analysis is the FPS. 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 7: 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)
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)
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)
Observations	25,469	13,447		25,349	13,334	
% of sample	89	91		88	90	

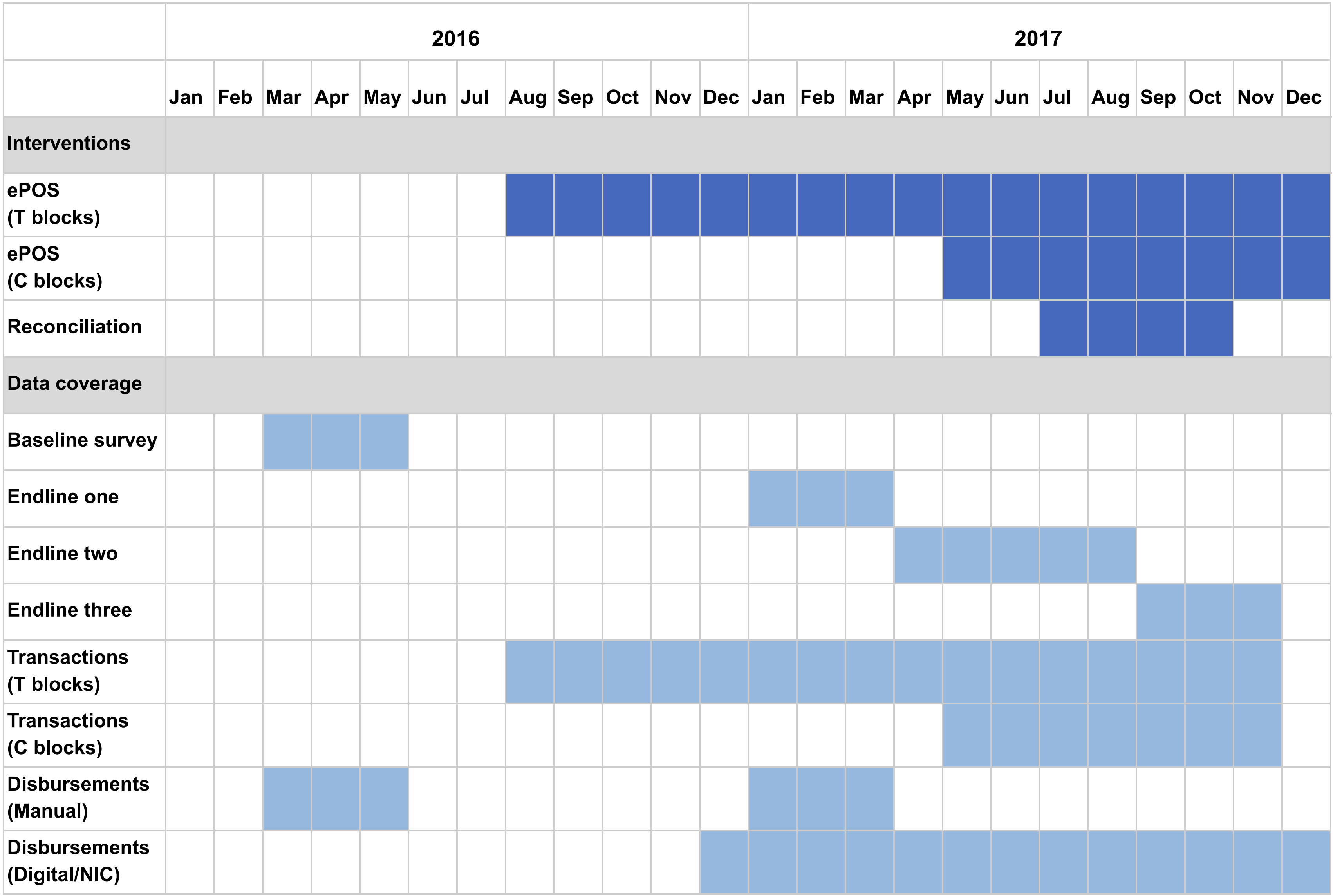
This table reports estimates of the effect of reconciliation on measures of the value disbursed by the government (Panel A), received by recipients (Panel B), and the difference (Panel C) separately for treatment and control areas using data from all three endlines. The unit of analysis is the FPS-month 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 FPSs 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 2 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 8: 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	93.1%
Total	144,161	2,305,449	2,449,610	100%
% of overall total	5.9%	94.1%	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%	

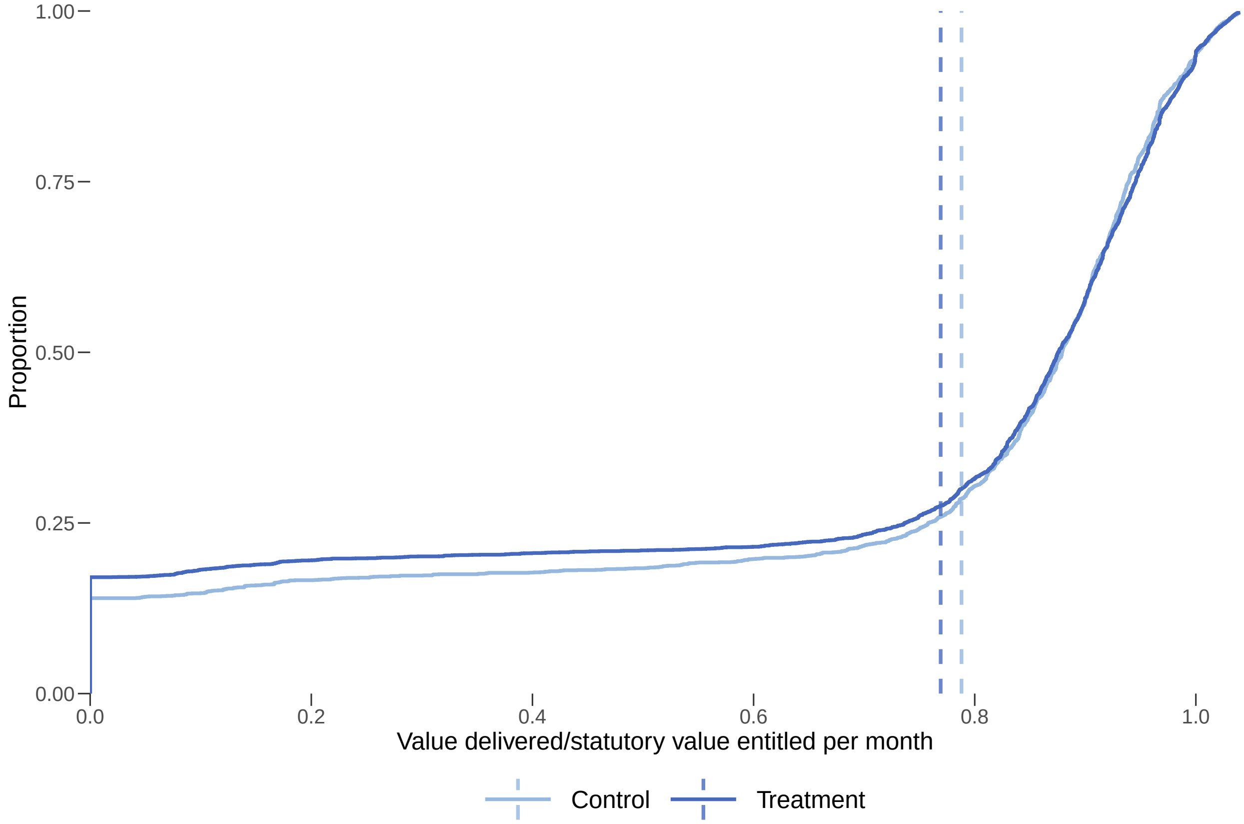
This table reports the decomposition of ration card deletions by Aadhaar seeding status. The top panel shows the results from the universe of ration cards in our 10 study districts, and the bottom panel shows results from our sampled ration cards, for which we show counts adjusted by sampling probability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

Figure 1: 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.

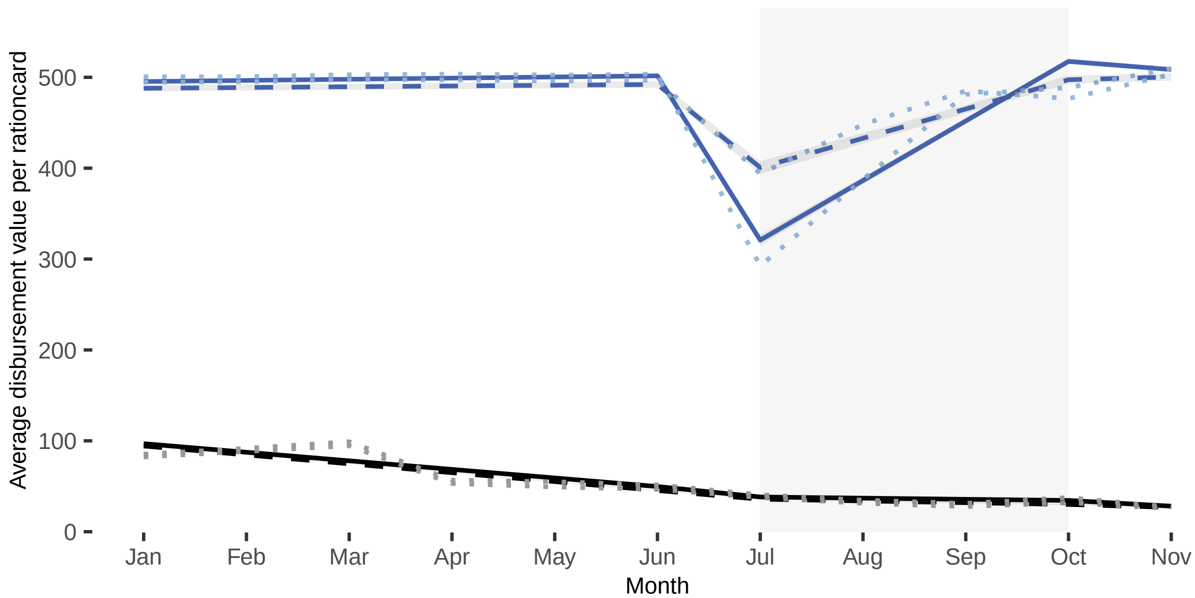
Figure 2: Value received as a proportion of entitlement



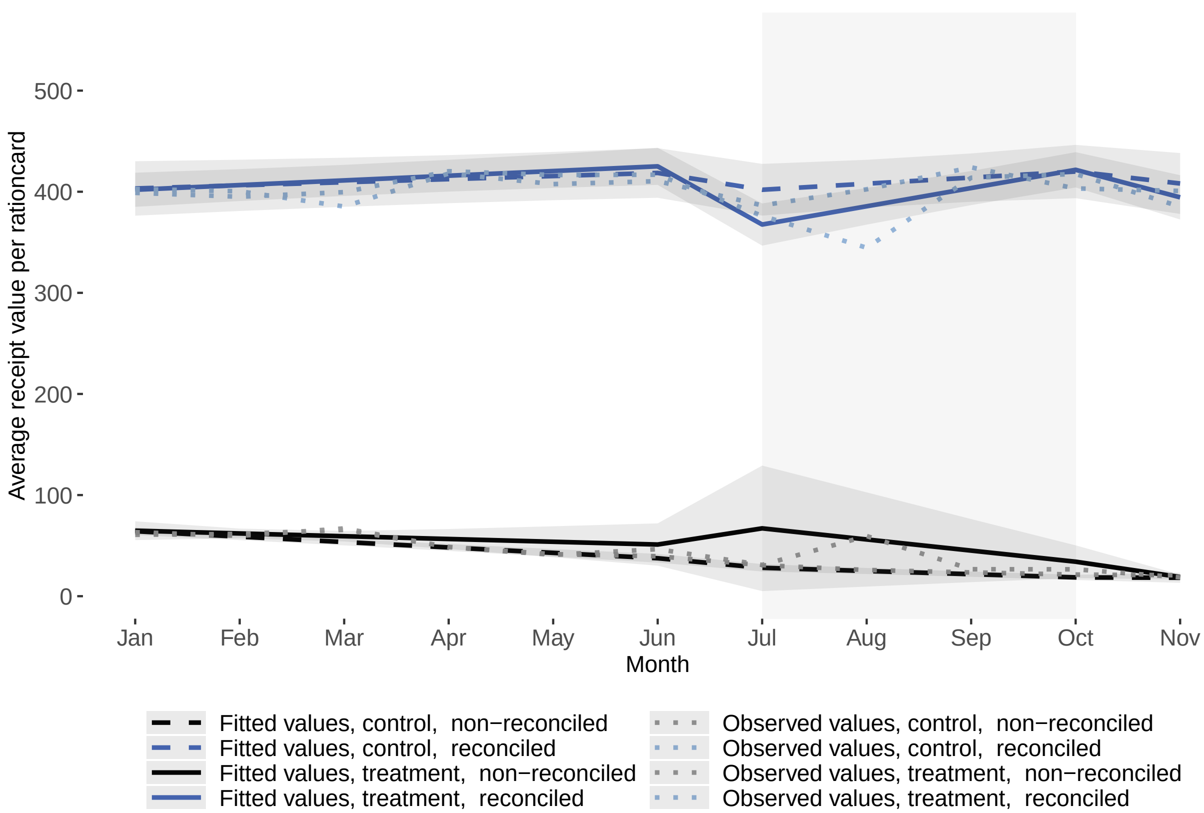
This figure plots the empirical cumulative distribution, separately for households in treatment and control blocks, of value received divided by value entitled per month, pooling the months of January, February, and March 2017. To improve legibility we right-censor the distributions at the 90th percentile.

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

(a) Value disbursed



(b) Value received



This figure plots the evolution of the average value of commodities disbursed (Panel A) and received (Panel B) by treatment status from January to November of 2017. The unit is the ration card-month. Value disbursed and value received are as described in Figure A.4. Dotted lines represent the raw data, while solid lines and dashed lines represent fitted values obtained by estimating Equation4 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.

A Supplemental exhibits

FOR ONLINE PUBLICATION ONLY

Table A.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				
Ration cards per FPS	308	293	15	0.45
Beneficiaries per FPS	981	1,041	-60	0.33
% 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. "Ration cards per FPS" is the ratio of PDS ration cards to the number of FPSs. "Beneficiaries per FPS" is the ratio of PDS ration cards to the number of FPSs. "% FPS rural" is the share of FPSs 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 A.2: 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.3: 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 difference 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: 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 difference 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.5: 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	3.4	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	.034
	(.57)	(.32)	(.081)	(.063)	(.047)
	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
Control mean	3.4	.86	.61	.23	.68
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 FPS-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 FPSs 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.6: 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 (Rs)						
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 FPS 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.7: 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 A.8: 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 A.9: 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 3, 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. Note that our survey asked about both 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 alternatives. 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) + transportation 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.10: 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.11: 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.12: 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). We assign counterfactual machine modes to control FPS's by assuming they could have received machines operating in the same mode as it was ultimately assigned once treated. This assumption 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. We define the mode in which a machine operated from transaction data as 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. 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.13: Heterogeneous effect by subjective FPS rating

	HH is Above Median of Implementation Quality		
	No	Yes	Δ
Value Received (market prices)	-4.84	-17.24	-12.40
	(15.85)	(14.75)	(16.04)
Value Received (WTA)	8.70	-7.48	-16.18
	(19.98)	(21.80)	(23.27)
Transaction Costs	8.05	3.59	-4.46
	(5.77)	(5.89)	(7.31)

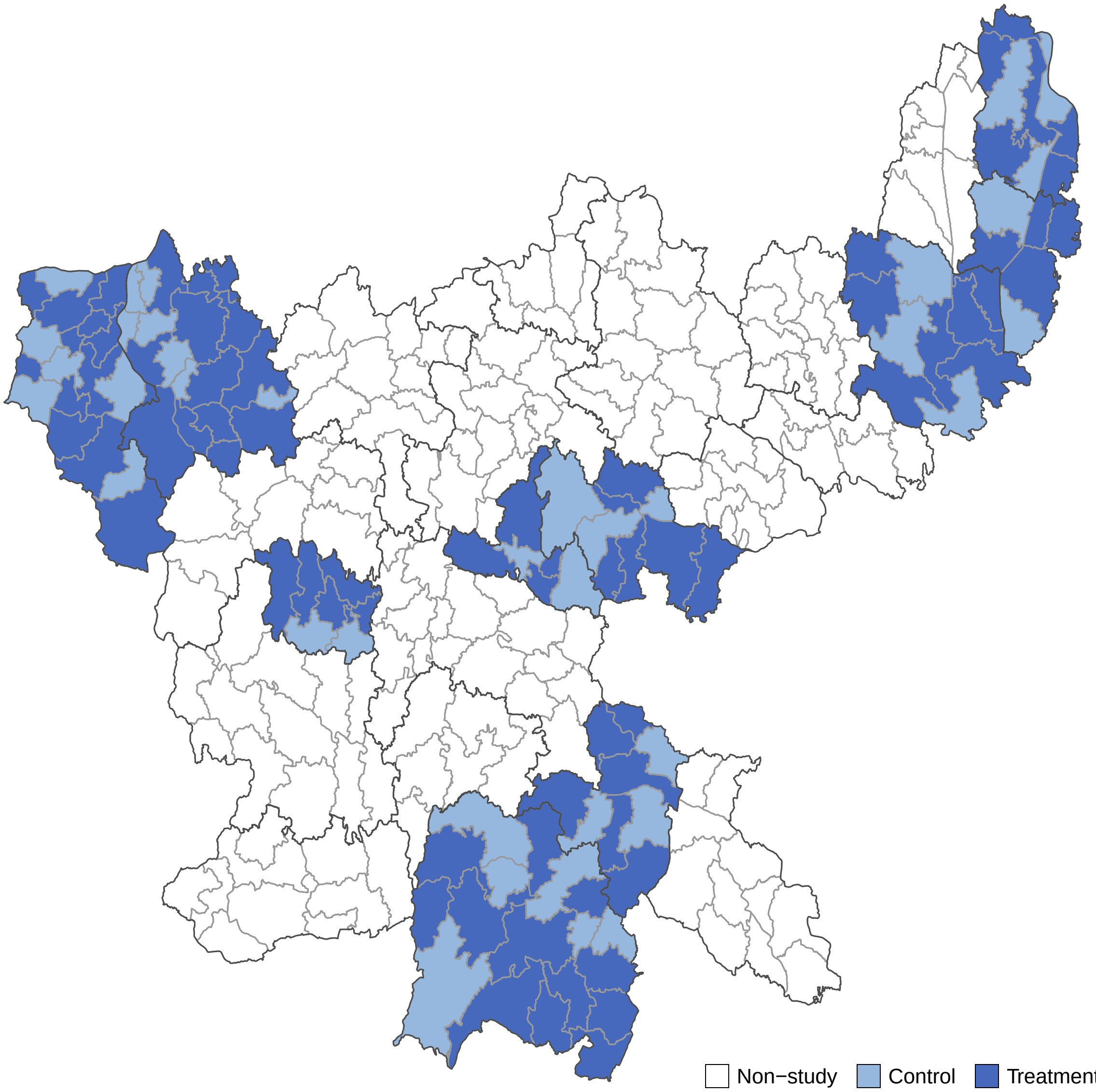
This table reports differential estimated treatment effects along a dimension of implementation quality characteristics for ration card-level outcomes in endline one (January - March). Each row represents a different primary outcome, and the column grouping represents an indicator for if the household is predicted to have a subjective FPS rating above the median. Within each column group, the first column reports the average treatment effect on households that do not satisfy this 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. 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.14: Heterogeneous effects of reconciliation on value disbursed and received by FPS balance

	Value disbursed			Value received		
	(1)	(2)	(3)	(4)	(5)	(6)
	Total	Rice	Wheat	Total	Rice	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	493	439	54	407	363	44
Observations	92489	95581	93258	35518	36269	36046
% of frame/sample	91	94	92	82	83	83

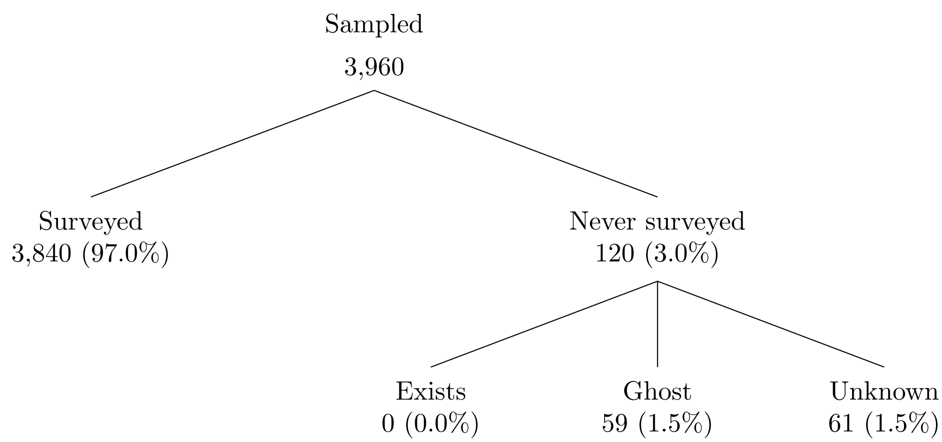
This table reports differential effects of accumulated stock balance at the time of reconciliation onset (July 2017) on value disbursed by dealers (columns 1 - 3) and received by beneficiaries (columns 4 -6). The unit of analysis is the FPS-month for columns 1-3 and ration card - month for columns 4-6. Observation counts vary because we use the universe of FPSs 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 2 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 by the FPS level (as pre-specified in our PAP) are reported in parentheses; clustering at the block level only affects significance levels for some coefficients in column 3, 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: Blockwise treatment assignment



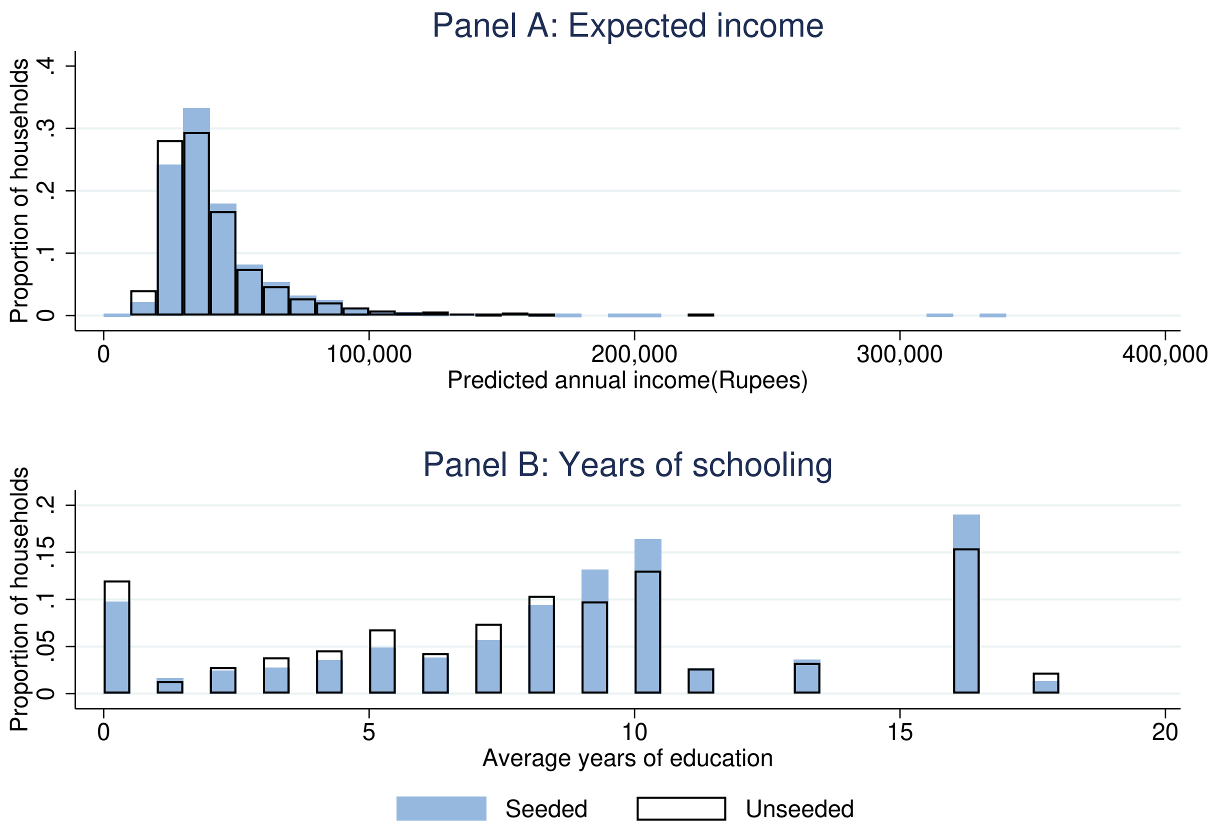
This figure shows the assignment of districts within Jharkhand to study (10) and non-study (14) status, and the assignment of blocks within these districts to treatment and control. Note that four of the census blocks depicted here are further sub-divided for the purposes of PDS administration into an 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 A.2: 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 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.

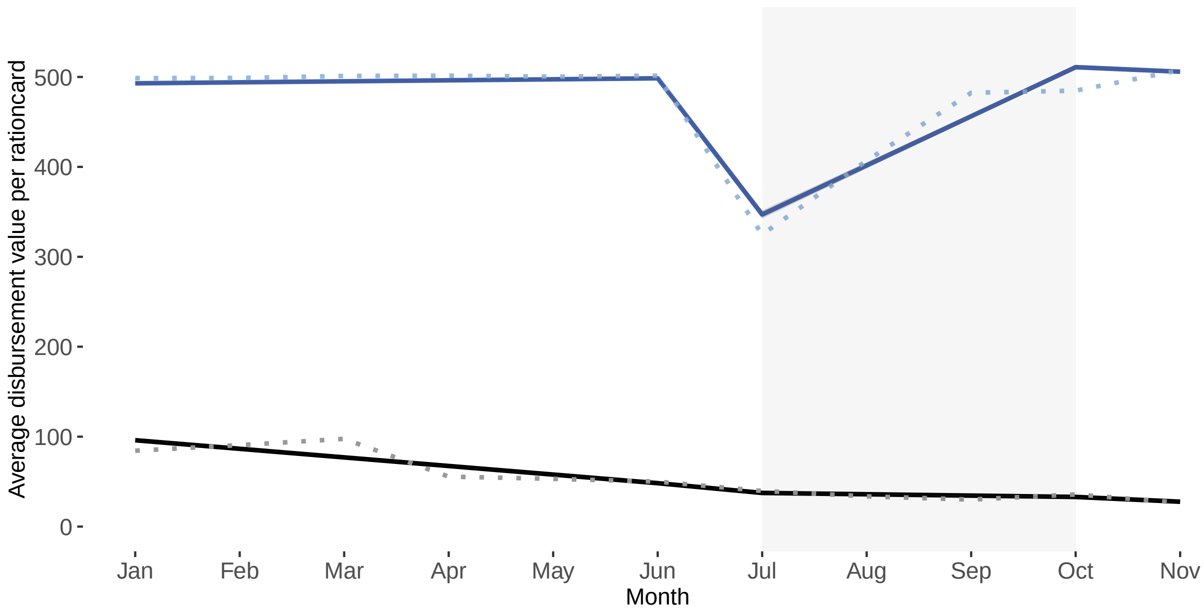
Figure A.3: Household expected income and years of schooling by seeding status



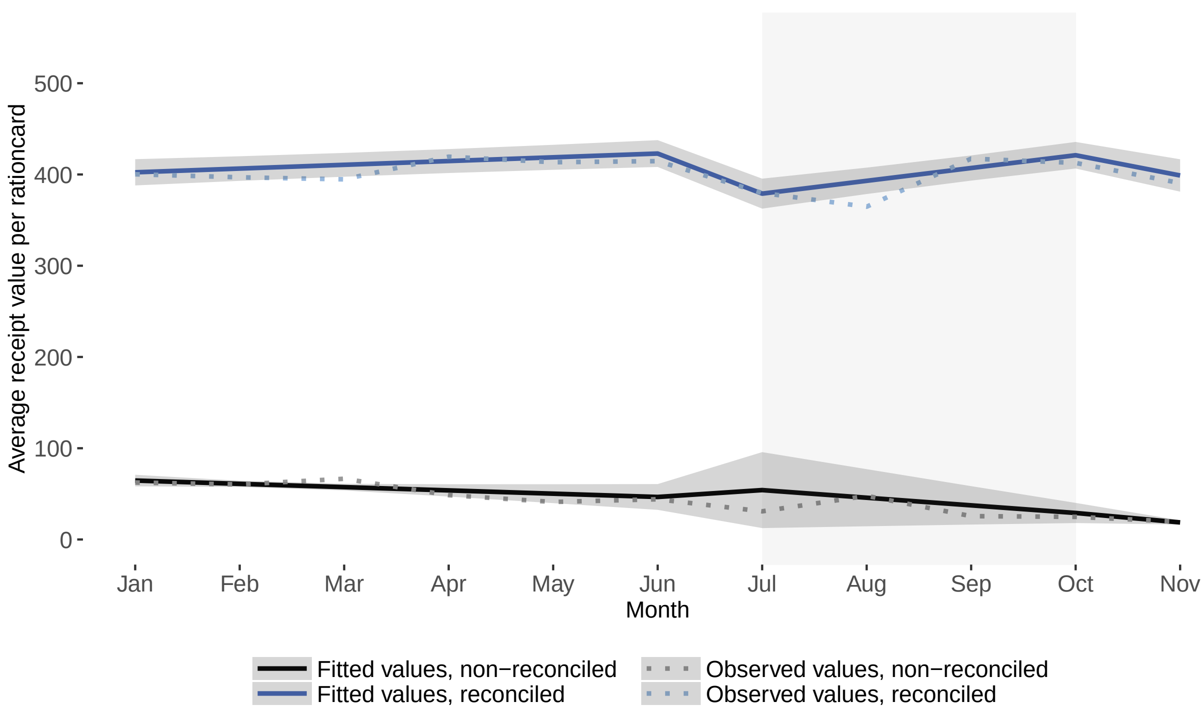
This figure shows the distribution of measures of household income (Panel A) and education (Panel B) by whether the ration card that household was attached to had at least one Aadhaar number seeded at baseline. In Panel A, the outcome is the linear prediction of annual income based on assets and household 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.4: Effects of reconciliation on value disbursed and received

(a) Value disbursed

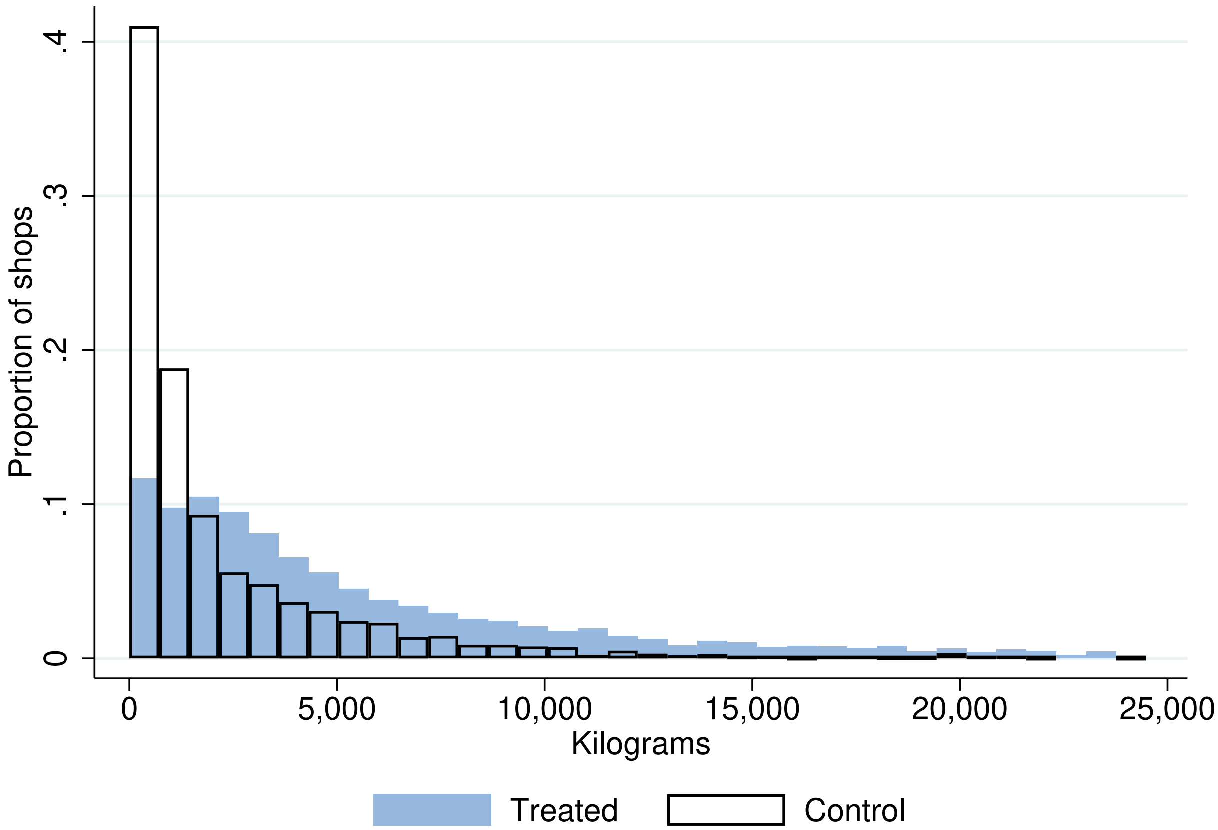


(b) Value received



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 2. Dashed lines represent the raw data, while solid lines represent fitted values obtained by estimating Equation 4. 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 A.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 FPSs at the end of June 2017 according to government records, separately for shops in treated and control blocks. The unit of observation is the FPS. To increase legibility the distributions are right-censored at the 95th percentile.

B Pre-analysis plan crosswalk

This appendix reports additional pre-specified analysis that was not reported in the main paper, and a list of exhibits in the main paper that are additional to those we pre-specified.

B.1 Additional pre-specified analysis

- Tables B.1, B.2, B.3 and B.4 examine temporal heterogeneity in the impacts of ABBA. We generally see little evidence of trends in the treatment effects.
- Table B.5 reports impacts of ABBA on FPS dealer outcomes, controlling for the baselines values of analogous outcomes. The results are very similar to those in Table 6.
- Table B.6 reports households' and FPS dealers' stated preferences related to the ABBA intervention, captured for the treatment group in March. Both sets of preferences are highly polarized.
- Table B.7 reports the impact of ABBA on government allotment by block, comparing control blocks in treated districts to blocks in non-study districts in order to check if there are any spillover impacts. We do not see any evidence of spillovers through administrative channels.

B.2 Additional analysis conducted

- Table A.1, which reports the representativeness of our study area with Jharkhand;
- Panel B of Table 1, which compares measures of program implementation in treated and control areas;
- Table 4, which examines impacts on the extensive margin of value received;
- Table 5, which reports heterogeneous effects by Aadhaar seeding;
- Figure 2, which plots the distribution of value received in treated and control areas;
- Figure A.3, which compares household income and education levels between seeded and un-seeded households;
- Table A.2, which compares the dealers we originally sampled to those we added to the sample as households were re-assigned to them;
- Figure 3, which plots the evolution of value disbursed and received in treated and control areas;
- Figure A.5, which plots the distribution of grain stocks as of June 2017 by treatment status.

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 2 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.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.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.6 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 2 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: $^{*}p < .10,^{**} p < .05,^{***} p < .01$.

Table B.5: Effects on dealer expectations controlling for baseline values

	Intends to continue running FPS?		Expected bribes to obtain license?		Expected bribes to renew license?		Profit (self-reported)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	.047	.024	-56,813*	-59,841*	-111	-83	-1,288**	-1,215**
	(.054)	(.059)	(28,922)	(33,854)	(123)	(147)	(511)	(558)
Adjusted R^2	.098	.14	.3	.27	.053	.035	.098	.083
Control mean	.73	.71	76,590	81,188	565	555	5,891	6,113
Observations	437	366	150	127	370	307	445	370
% of sample		92		32		78		93
Sample	Full	Restricted	Full	Restricted	Full	Restricted	Full	Restricted

This table reports estimated treatment effects on measures of FPS dealers' expectations in March about the future. The unit of analysis is the FPS. 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. The dependent variable in columns 7-8 is the dealer's self reported profits from ration sales. For most dealers, profits from ration sales is the same as total profits. In columns 1, 3, 5 and 7 the sample includes all dealers surveyed, including those to whom sampled households switched between baseline and endline; in columns 2, 4, 6 and 8 it includes only dealers drawn in the original sample. All specifications include strata fixed effects and the baseline value (except columns 5 and 6). Standard errors clustered at the block level are reported in parentheses with statistical significance denoted as: $^{*}p < .10, ^{**}p < .05, ^{***}p < .01$.

Table B.6: 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 B.7: Spillover effect of ABBA on allotment

	Dependent variable:
	Quantity Alloted
Treated District	31,653.34
	(22,028.09)
Observations	618
R^2	0.94
Adjusted R^2	0.94
Note:	*p<0.1; **p<0.05; ***p<0.01

This table the impact of being in a treated district on government allotment, comparing control blocks in treated districts to non-study blocks. The dependent variable is quantity allotted by block. Standard errors are clustered at the district level.

C Empirical methods

C.1 Randomization

As described above, our study takes place in 10 districts out of the total 24 in Jharkhand. We exclude 1 which had already started ABBA and 6 in which the government was rolling out another related reform - Direct Benefit Transfers for kerosene, which also involved the PDS system and FPSs. We randomly sampled 10 of the remaining 17 within which to randomize the rollout of the intervention. We used stratified random sampling to classify the 17 available districts into 8 (2x2x2) distinct categories using 3 binary variables related to geography and socio-economic status: an indicator for being above/below the median of district (centroid) latitude, an indicator for being above/below the median of district (centroid) longitude, and an indicator for above/below the median of the first principal component of a number of additional variables. We then sampled half of the districts in each category, rounding down to the nearest integer and using probability proportional to size (measured as the number of FPSs) sampling, and lastly sampled additional districts without stratification to reach our target of 10. This design ensures representativeness of the 17 districts in our frame.

Our unit of randomization is the sub-district ("block"), which on average covers 73 FPSs and 96,000 people. 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. Within each study 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: the average number of unique Aadhaar IDs per ration card, the average amount of ration claimed per PH household according to administrative records, and the average amount of ration claimed per AAY household. Within each group of 3 blocks we randomly assign 2 to treatment and 1 to control.

C.2 Sampling

Our sampling procedure for dealers and households uses the administrative database of eligible PDS beneficiaries and their assignment to FPSs from GoJH. We first sampled 3 FPSs in each study blocks for a total of 396 shops, via PPS sampling with "size" defined as the number of ration cards assigned to that FPS. We then sampled households using the list of ration cards assigned to sampled FPSs.

For each sampled FPS we sampled 10 households from the government's list of PDS beneficiaries. We define a household here as those individuals listed on 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. This generated a target sample of 3,960 households. We attempted to interview these households for baseline and three follow-up surveys to create a household-level panel. We ultimately identified and interviewed the corresponding household at least once in 97% of cases. 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.

C.3 Field data collection protocols

To determine whether a sampled household was a "ghost" (non-existent) household, our field teams followed a stringent procedure. We obtained address details for sampled ration cards from administrative data. Survey enumerators additionally obtained the full set of household member names from the Jharkhand Government PDS System SECC Cardholders list to help distinguish and ascertain the sampled household.

Enumerators first identified and reached the households' listed village. Many villages are divided into sections that are some distance from each other, called "tolas;" enumerators visited each tola of the village looking for the household. Upon reaching each tola, enumerators asked locals for help in locating the particular household, using the address and the full list of household member names to identify unique households. If they could not find the household, the team would revisit another day to search for the household. If the household could still not be found, the team supervisor would consult with the research monitor and make a subjective determination on whether it is reasonable (according to geographic spread and population density) that all probable hamlets/habitations in the village/urban area were searched to find the household. If all areas could not be searched, we labeled the household as "Not found" (this is *distinct* from a "ghost" categorization).

If all areas had been searched and the household not located, a separate enumerator was assigned the task of contacting one or more of these village leaders to attempt to locate the household: ward member (lowest level elected representative), worker in Anganwadi (childcare) centre, mukhiya (elected head of village). The enumerator spent half a day dedicatedly searching for this household. If no one had ever heard of this household, we had two neighbors (from the local tola/hamlet) record their names, addresses and mobile numbers for confirmation purposes. We then marked this household as a "ghost" household. Note that we attempted to survey at endline all households selected at baseline, including those we were unable to survey at baseline.

After identifying a sampled household as above and upon approaching a household member, the first thing enumerators did after obtaining verbal consent was to confirm that the sampled ration card actually belonged to the household. Enumerators asked beneficiaries to produce the ration card in order to verify it, and also asked for any other ration cards that the household might own. We captured details of household members listed on the sampled ration card, and clarified that all questions pertained to the sampled ration card only.

After confirming ration card details, we asked for the households' PDS purchase history from a member with knowledge of these purchases. We asked for purchase records for the previous three months at baseline, endline 1, and endline 3, and the previous five months at endline 2. While ration cards were available as per above, we do not necessarily rely on these for obtaining beneficiary purchase records, since any listing on the ration cards could have easily been manipulated by the dealer at the time of purchase.

Instead, we ask for each commodity, and for each month: the quantity actually purchased, the quantity they are officially entitled to, and how much money they paid for the purchase. We allowed beneficiaries to report either price per unit or the total paid, clearly marking which of the two the amount corresponded to. We also asked about the number of months worth of entitlement this month's purchase accounted for. Our questions are thus far more detailed than, for example, the National Sample Survey (NSS) that is used for most analysis related to the PDS in India, which only asks about the quantity and value of PDS commodities consumed in the last 30 days.

Note that PDS purchases are extremely salient; beneficiaries go once a month, usually around a particular time of month if not a specific date each month, and buy their entire month's ration at that time. We therefore did not encounter any significant recall issues.

D 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

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

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.

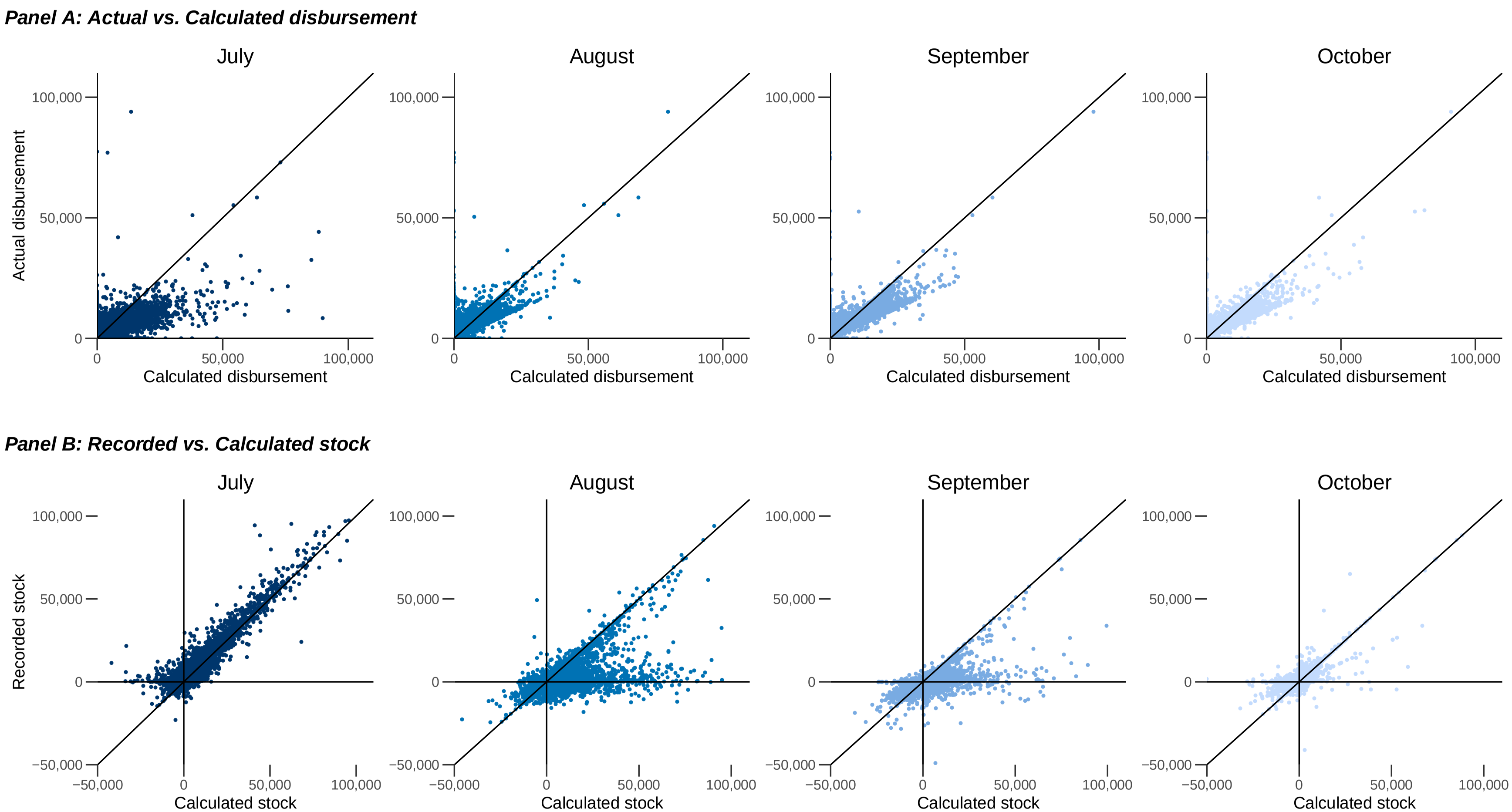
To examine how closely the government adhered to this protocol we use two sources of data: measures of the aggregates S_t, E_t, C_t, D_t and O_t which the National Informatics Commission maintained for each FPS and month, and transaction-level data directly from the ePOS devices themselves which record $O_{i,t}$ for each household and month (also provided by the National Informatics Commission).

The government was fairly flexible in its implementation of reconciliation. Figure D.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 re-calculating 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 D.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 D.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.

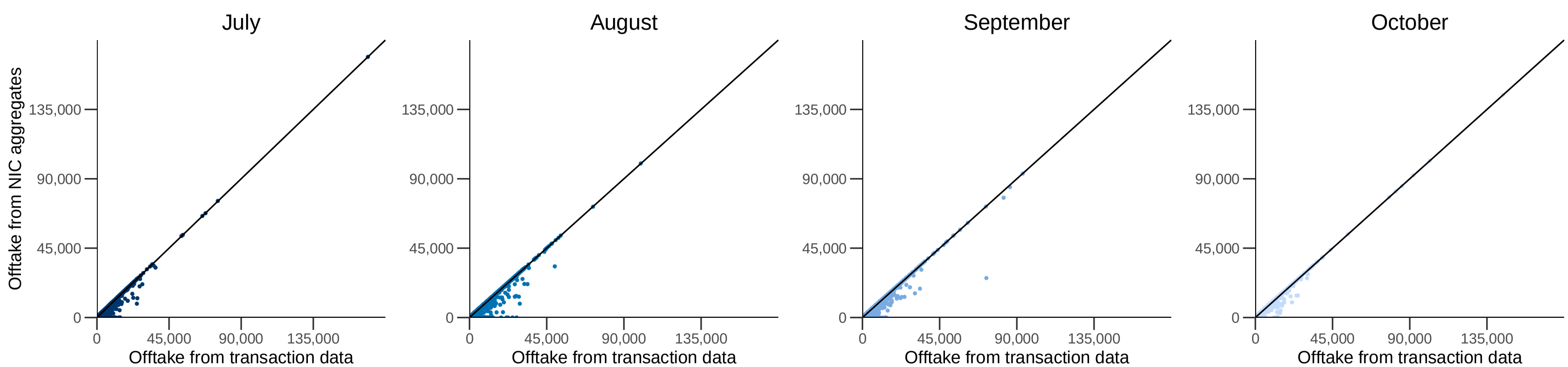
Figure D.1: Adherence of disbursement and stock to reconciliation policy, by month



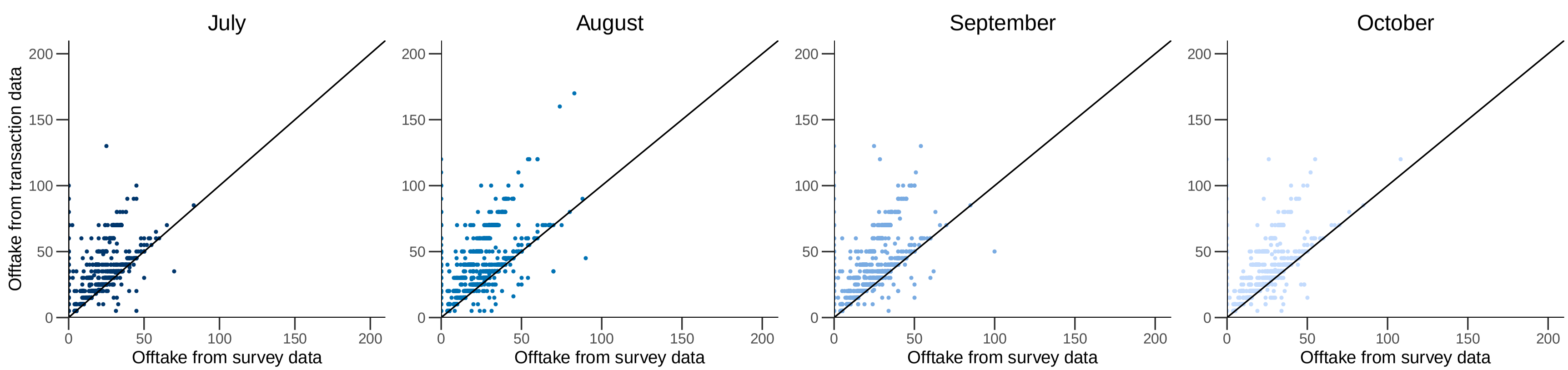
This figure shows the scatterplots of the left- and right-hand sides of Equations 1 and 2for the months July-October 2017 using aggregate measures obtained from NIC data in kilograms. Panel A has actual disbursements on the y-axis and uncensored calculated disbursements on the x-axis, and the Panel B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

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

Panel A: Offtake from NIC aggregates vs. Offtake from transaction data

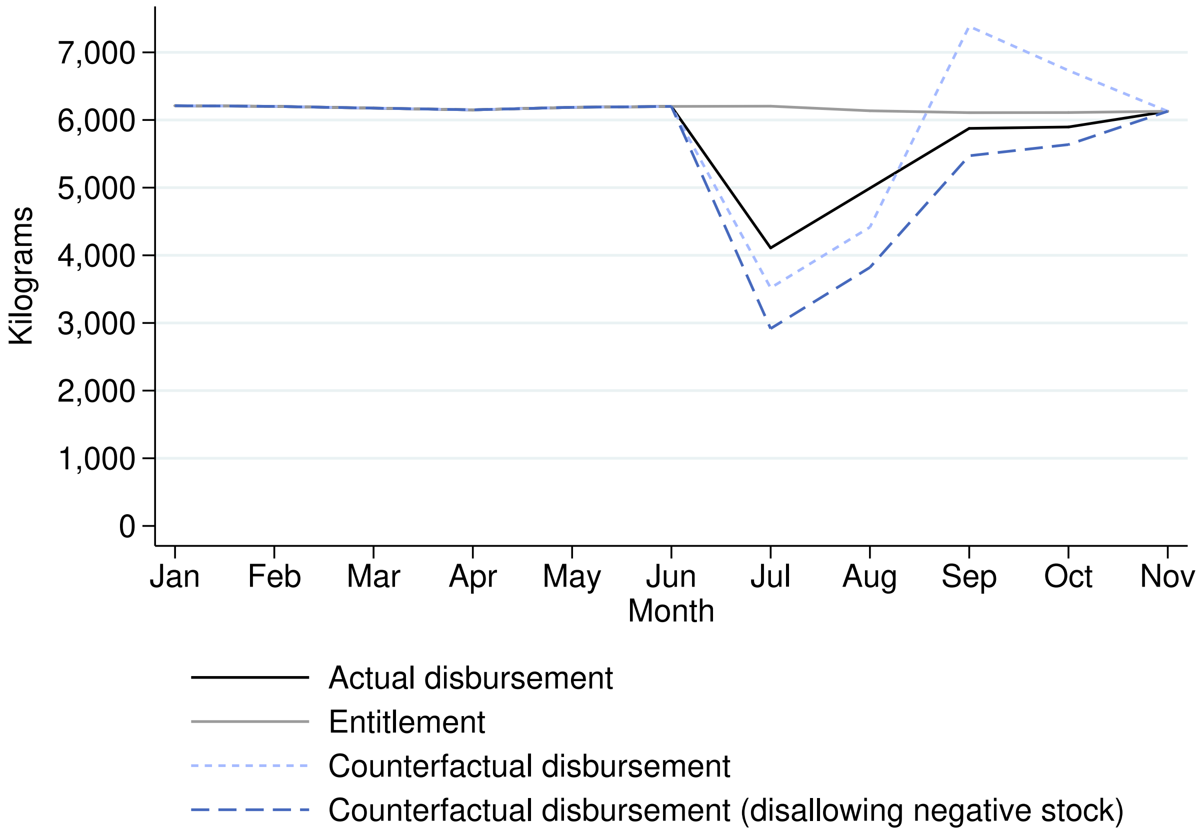


Panel B: Offtake from transaction data vs. Offtake from survey data



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 D.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.