Operational Impact of Technological Advancements in Public sector Supply chains – Evidence from India's Food Security Program

by

Maya Ganesh

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Fellow Program in Management
at The Indian School of Business
2021

Doctoral Committee:

Professor Sarang Deo, Co-Chair
Associate Professor Sripad Devalkar, Co-Chair
Professor Kamalini Ramdas

To my parents & all my gurus, with infinite gratitude.

Acknowledgements

I would like to express my boundless gratitude and indebtedness to Professors Sarang Deo and Sripad Devalkar, who I have been very fortunate to have as advisors. They did a great bit of hand-holding during the initial phases of my doctoral program, mentored me, believed in me and constantly pushed me to do better. If today, my goal is to strive towards excellence and make perfectionism my muscle memory, it is because of them. They were always available for discussions and it is their inputs and advice that have shaped me as a researcher. Their clarity of thought, thorough research process and precise communication is what I aspire to attain. Professor Sarang has been immensely patient and taught me the value of perseverance and thinking from first principles. Professor Sripad's advise of staying with a problem long enough that it becomes tamed, has helped me time and again to push and expand my boundaries of both understanding and intuition. I would also like to thank my thesis committee member, Professor Kamalini Ramdas for providing invaluable career advice. Her constructive comments and feedback greatly improved my dissertation.

I want to thank all the OM faculty for their constant support and for putting up with my questions owing to my initial unfamiliarity with academia which I always garbed under the "I am the first OM FPM student" coat! I cannot forget to thank Professors Bhimasankaram and Manasa Mandava, who taught majority of my foundational courses with umpteen patience, but are unfortunately not here today to see me graduate. My thanks to Professors Ashwini Chhatre and Shilpa Aggarwal whose initial encouragement helped me develop interest in both policy related research and food and agricultural systems.

My understanding about the Public distribution system supply chain would have been incomplete without my multiple visits to the Food and civil supplies office. I want to thank all the officials, especially Senior Deputy Director, Shri

M.C.Gangadhara and office manager, Ms.Subhashini who entertained all my questions and arranged field visits which helped strengthen my understanding of the context. My thanks to the LRC team who have always been responsive, even when I have requested access to papers, books and articles in the middle of the night. I also want to thank the FPM office for their help and support, especially Sireesha for always extending a helping hand and taking care of program related administrative issues.

I want to thank a number of persons who have been companions at different stages during my stay at ISB - Bhavya for always being there, patiently listening to my incessant ranting and celebrating my little milestones with me, however seemingly trivial; Rakesh for the innumerable discussions spanning research, academia and philosophy; Senthil, Neha, Ramya, Moumita and Shashin for making this journey enjoyable with delightful conversations that helped maintain my sanity; my fellow students in the FPM program especially Saipriya, Arkaja and Jeramia for many useful discussions and interactions; OM and Economics research associates, especially Nabin, Hemanth, Chandru and Shubham for helping me with my research at multiple occasions.

This journey would have been impossible without the unconditional love and support of my family. My parents have stood by me every time I have wanted to explore something new, however unconventional it might have seemed to them. They have been the constant pillar of strength from which I have derived renewed energy at the face of every obstacle. My sister's doctoral student journey gave me a sense of what to expect from the life of a researcher. While many might think that I jumped into the ocean of a doctoral program without having any prior experience of swimming, this inexplicable confidence, I largely owe to her, whose journey I was fortunate to witness and from whose experiences I greatly benefited.

Table of Contents

List of Tables

List of Figures

ABSTRACT OF THE DISSERTATION

Operational Impact of Technological Advancements in Public sector Supply chains – Evidence from India's Food Security Program

by

Maya Ganesh

Large government managed public sector supply chains are often plagued with inefficiencies which undermine their effectiveness to deliver public goods. New technologies have the potential to improve delivery effectiveness by addressing these concerns. I study the impact of two such technology enabled interventions in the context of India's food security program, the Public Distribution System (PDS). The PDS is one of the largest food security systems of its kind in the world. It delivers subsidized food grains to nearly 160 million low-income households through approximately 500,000 Fair Price Shops (FPS). There are many sources of inefficiency associated with the PDS – diversion of food grains at different stages of the supply chain, beneficiaries not receiving their complete entitlement, FPSs not being open regularly, mistreatment of beneficiaries by the FPS owner, overcharging of grains etc. In order to address these issues, governments and policy makers have introduced multiple technology enabled interventions across the PDS supply chain such as digitisation of beneficiary records, implementation of biometric authentication devices, introduction of choice of FPSs, grievance redressal mechanism and transparency portals, etc. I study the operational impact of two process changes in the downstream part of the PDS supply chain (monitoring at FPSs and collection

of entitlements by beneficiaries), which were enabled by the technological intervention of introduction of biometric authentication devices.

In the first essay, I examine the impact of installing Biometric Authentication (BA) based monitoring devices at more than 3,300 FPSs in the Indian state of Karnataka between 2013 and 2015 on diversion of food grains at the last mile. Using a difference-in-differences estimation technique, I find that diversion of food grains at the FPSs installed with BA devices was lower than the baseline diversion quantity by 4%. I also find that additional value can be unlocked by using the sales information captured by the BA devices for replenishment planning. I find that the value of better planning can be substantially higher than that of better monitoring (up to four times) depending on the relative magnitudes of pre-intervention levels of diversion and beneficiary demand. I find that the value of better planning is particularly high when the average demand and average diversion level before BA installation is neither too high nor too low.

In the second essay, I study the utilisation of a novel reform, termed *portability*, that was enabled by the use of BA devices. *Portability* allows beneficiaries to collect their entitlements from any FPS within the state, based on convenience, and not necessarily the FPS they were originally allotted to. I use large-scale program data from the state of Andhra Pradesh in India to empirically quantify the relationship between the utilization of alternate FPSs by beneficiaries and determinants of *portability* including availability of alternate shops near them, FPS characteristics and household characteristics. Using binary logistic and poisson regression models along with associated variance decomposition methods, I find that: i) addition of one extra FPS within 0.5 km radius of the beneficiary household increases the likelihood of using *portability* by 6.8%, ii) variation in the availability of alternate FPSs explains more than 60% of the variation in utilization of *portability* and the number of alternate FPSs used, and iii) vulnerable households (poor, socially disadvantaged, elderly, rural) have lower rate of utilization of *portability* than non-vulnerable households. The findings imply that the potential benefit from utilization of *portability* is less likely to be realised unless fundamental changes are made to the FPS network, with special attention being paid to regions with lower FPS density.

In the third essay, I develop a structural estimation model to understand the impact of *portability* on beneficiary welfare. I conjecture that the realised improvement in welfare could be less than the maximum possible improvement if there are instances of stock outs at the beneficiaries' most preferred FPS. This is because providing choice increases demand uncertainty at FPSs and can lead to unintended negative consequences like stock outs if operational decisions such as replenishment policies are not suitably modified to account for this uncertainty. Using a multinomial logit model, I find that a 0.1 km increase in distance to an FPS leads to a 18.2% decrease in the likelihood of its usage. I also find that for every extra day that an FPS is kept open, the likelihood of its usage increases by 1.6%. Using the estimates from the choice model to construct the baseline scenario (no choice), I find that provision of choice increases the proportion of beneficiaries purchasing from the PDS by 5.4% and average beneficiary utility by 12.04%. I find that the maximum attainable utility increase is 25.4%, which is not realised due to stock outs at the beneficiaries' most preferred FPS, which are to the tune of 5.96%. These stock outs are a result of continuing the replenishment policy defined during the pre-choice period, which does not account for increased demand variability experienced at the FPS after the provision of choice. I find that the coefficient of demand variation within FPSs increases by 62.1% after providing choice. The findings imply that a large portion of potential welfare gain is not recovered due to the absence of complementary replenishment policy modifications which can account for the demand variation resulting from the provision of choice.

The evaluation of the installation of BA devices illustrates how operations and supply chain frameworks can complement traditional impact evaluation approach. These frameworks facilitate a careful analysis of the interactions between technology interventions and underlying processes, which can unlock additional sources of potential value that are not explicitly captured in the traditional approach. The results from the analysis of utilization of *portability* indicate that the number of alternatives in a household's vicinity is a much more important determinant of utilization of choice in comparison to other program and household characteristics. Finally, the analysis of the impact of *portability* on beneficiary welfare suggests that a large portion of potential welfare gain could remain unrealised due to the absence of associated supply side modifications.

Chapter 1

The Public Distribution System

India's food security program known as the Public Distribution System (PDS) is one of largest food security programs in the world (Bajaj, 2012). Food grains worth approximately 1.8 Trillion INR are distributed annually to around 160 million economically weaker households through a supply chain comprising of a wide network of more than half a million fair price shops (FPS)[1]. The FPS constitutes the last mile delivery of the PDS supply chain. The expenditure on distribution of grains constitutes close to one percent of India's gross domestic product and results in a dependence of more than half of India's population on the PDS. Each household, based on its economic status, is entitled to receive a fixed quantity of food grains every month FPSs at heavily subsidized prices ($\approx$ INR 1 - 3 per kg compared to market prices of $\approx$ INR 28 - 40 per kg). Typically, private dealers or cooperative societies are issued a license to manage the FPS for a fixed period of three years and are paid a commission of about INR 0.70 per kg of grains distributed to beneficiaries.

There are many sources of inefficiency associated with the PDS - (i) Diversion of food grains at different stages of the PDS supply chain and (ii) low quality of service which manifests as FPSs not being open regularly, mistreatment of beneficiaries by the FPS owner, overcharging of grains, non-availability of commodities etc.

[1] `http://epds.nic.in/`

1.1 Diversion of food grains

PDS diversion refers to the proportion of grains released by the Food Corporation of India (FCI) that fails to reach consumers (Dreze and Khera, 2015). Diversion at the FPS can be defined as the gap between the quantity of grains that enter the shop and the sum total of all grains that is issued to the beneficiaries. Diversion has been a prevalent issue of the PDS, at all stages of the supply chain, including at the FPS, which is the last mile delivery (Lal, 2015; Kulkarni, 2014; Chandra, 2014). This results in beneficiaries not receiving their complete entitlement. In fact, several estimates suggest that only 50% of the grains entering the PDS reach the intended beneficiaries (Government of India, 2015). One of the key drivers of leakage is the willful diversion of grains by FPS owner to the open market. The primary reason for diversion is that the commission gained by issuing grains is less compared to the earnings realised by diverting the grains to open market and selling them at the market price; This presents an opportunity for arbitrage.

1.2 Service quality issues

Traditionally, each FPS has a set of beneficiary households affiliated to it and issues grains to only those set of households. This system accords monopoly power to the FPS dealers over beneficiaries, which in turn led to inefficient and poor quality of service manifesting in terms of frequent FPS closures, mistreatment of beneficiaries, long queues, adulteration of grains, overcharging and under-weighting (Sati, 2015; Sargar et al., 2014; Vaidya et al., 2014; Dreze and Khera, 2015; Dhanaraj and Gade, 2012; Sharma and Gupta, 2019). Despite the presence of grievance redressal mechanisms and vigilance committees, only an estimated 1.5% of the beneficiaries across the country were aware of them (NCAER, 2015).

1.3 Technological solutions to address PDS inefficiencies

Increasingly policy makers and governments are turning towards technology to address these issues. Specifically, a biometric authentication (BA) based system has been used to address the issue of diversion. Further, technology enabled programmatic solution of providing choice of FPSs (*portability*) has been introduced to empower beneficiaries and with the aim of improving quality of service at FPSs by inducing competition among them. In this dissertation, I study the operational impact of these two technology enabled interventions on the performance of the PDS. In the first essay, I study the impact of introduction of BA devices on the performance of PDS, with respect to monitoring and replenishment decision making. In the second essay, I study the utilisation of portability, the introduction of which was made possible by the earlier technological intervention of intallation of biometric authentication devices. I conduct a a detailed empirical investigation of the uptake of portability by beneficiaries and study the operational factors that drive this uptake. In the third essay, I build on the findings from the second essay and develop a structural estimation model to evaluate the operational impact of offering the choice of retail outlets to beneficiaries. I investigate if providing such choice increases demand uncertainty at FPSs and can lead to unintended negative consequences if operational decisions such as replenishment policies are not suitably modified to account for this uncertainty.

Chapter 2

Leveraging Digital Technology to Improve Monitoring and Planning in Public Sector Supply Chains: Evidence from India's Food Security Program

2.1 Introduction

Public sector supply chains in many developing countries routinely experience diversion of subsidized goods to private markets across a wide range of sectors such as food and nutrition, health, fuel and agricultural inputs (The Hindu, 2015; Khera, 2011; Vian et al., 2009; Saha, 2015; Dash, 2015). For instance, only 50% of the food grains in India's food security program (known as the *Public Distribution System or PDS)* comprising 160 million households and 500,000 *Fair Price Shops (FPSs)* reach the targeted beneficiaries (Government of India, 2015). Incentive and/or audit contracts used for monitoring quality and regulatory compliance in private sector supply chains (e.g., Plambeck and Taylor, 2015; Cho et al., 2019) may not be feasible or effective in these settings due to: (i) the scale and the distributed structure of their operations, (ii) the political influence wielded by the supply chain players to evade monitoring, and (iii) the slow and ineffective legal enforcement of penalties arising from violations discovered during audits.

Recent advances in digital technology offer alternative, cost-effective monitoring approaches to reduce the extent of diversion in such contexts. For instance, biometric authentication (BA) devices that require validation of identity of beneficiaries have been installed at thousands of FPSs across India to prevent fraudulent sales transactions by FPS owners (Sen, 2012; Masiero, 2017; Allu et al., 2019). However, the potential *value of better monitoring* through these technologies may not be fully realized if associated processes are not appropriately designed. For instance, allowing manual workaround in the event of failure of BA devices (e.g., heat, dust, interruptions in power and network availability) may be exploited by FPS owners to continue diversion (Sachdev, 2018).

In this paper, we undertake this quantification in the context of the PDS in the Indian state of Karnataka, where BA devices were installed in more than 3,300 FPSs in a phased manner between January 2013 and December 2015. We use publicly available data and conduct a quasi-experimental (Difference-in-Differences) analysis to compare the change in recorded sales in FPSs receiving a BA device with similar FPSs that did not receive one. We find that BA devices reduced diversion by an average of 36.3 kgs per FPS per month, which translates to ≈INR 36.2 million per year (≈USD 550,000) in terms of *value of better monitoring* (when calculated for 3,075 FPSs which received BA devices, in our analysis data set). Extrapolated to the entire state (23,241 FPSs), this would result in savings of ≈INR 301.2 million per year (≈USD 4.6 million). Combined with the investment required for installation, these savings yield a payback period of under two years.

In our study setting, the BA devices were equipped with the additional capability of recording real-time information on sales and inventory at the FPSs. This timely information provided by the BA devices (not available in the erstwhile manual process) was not used by central planners to improve the monthly inventory replenishment decisions. Hence, to estimate the potential *value of better planning*, we conduct an extensive simulation study, which is calibrated using the results from our empirical study and supplemented with other operational data from multiple

Indian states. We estimate that the average additional *value of better planning* is between 100%–300% of the *value of monitoring* and is particularly high when the average demand from genuine beneficiaries and the average diversion level before BA installation is neither too high nor too low. We find that a large portion of this value is driven by reduction in open market purchases by households due to reduced FPS stock-outs, which cannot be realized through better monitoring alone.

Findings from our econometric and simulation analysis, when put together, have important implications for implementation and evaluation of digital technologies in large public welfare programs. First, although the main motivation in implementing these technologies is typically to capture the *value of better monitoring,* policy makers should be cognizant that the same information can be leveraged to unlock additional *value of better planning.* This is true in many public welfare programs because both inadequate monitoring and inefficient planning can be attributed to the same root cause—lack of timely and accurate operational information—which is resolved by the new technology. Second, conventional impact evaluation approaches (Gertler et al., 2016) should be complemented with careful operational analysis to understand the relative magnitude of these two drivers of value. Apart from the immediate policy relevance, our findings advance the discussion at the confluence of two streams of supply chain literature: (i) emerging analytical work that proposes incentive contracts to improve socially responsible behavior in decentralized supply chains (e.g., Cho et al., 2019; Babich and Tang, 2012), and (ii) well-established analytical work that characterize the benefits of sharing information in decentralized supply chains by mitigating the Bullwhip effect (e.g., Lee et al., 1997, 2000).

We expand on these contributions to the relevant streams of prior literature in §4.2. We describe the empirical setting of PDS in Karnataka and the specific implementation of BA devices in §2.3. The accompanying data and our econometric approach to estimate the value of better monitoring using BA devices is described in §2.4. We present our main results in §2.5. In §2.6, we validate the mechanism of reduction in diversion by estimating heterogeneous effects. In §2.7, we describe the simulation study to estimate the value of better planning using timely inventory information provided by BA devices. Finally, §2.8 provides concluding remarks.

2.2 Literature review

Our work contributes to two sub-streams of the supply chain management literature. The first, emerging, stream of work develops analytical models to evaluate the (in)effectiveness of various incentive and monitoring contracts in making supply chain partners comply with ethical and environmentally sustainable practices. Cho et al. (2019) evaluate how pricing and inspection strategies can be used to restrict a supplier's use of child labor and find that reducing the cost of inspection does not necessarily reduce the use of child labor. Plambeck and Taylor (2015) show that increased audit frequency of supplier facilities can backfire, i.e., may reduce supplier's compliance effort and increase effort toward hiding non-compliance to prevent detection during audit. Huang et al. (2017) consider managing social responsibility in a three-tier supply chain where Tier 2 suppliers can potentially violate standards. They characterize conditions for optimality of either delegation (only Tier 1 supplier works with Tier 2 supplier) or control (the buyer or Tier 0 works directly with Tier 2) strategies. They also show that external pressure from NGOs or government on Tier 0 (buyer) or Tier 1 suppliers can sometimes backfire and lead to lower levels of social responsibility.

A related substream analyzes incentive contracts to deter product adulteration by suppliers. Mu et al. (2016) show, in the context of milk supply chains in emerging markets, that subsidizing the cost of product testing or investing in better testing infrastructure can lead to lower quality due to competition among milk suppliers and consequent free-riding. Levi et al. (2018) also study the impact of accuracy of testing on adulteration decisions in the presence of exogenous quality uncertainty in a distributed food supply chain. Babich and Tang (2012) show that a deferred payment contract, wherein a part of the payment to the supplier is made with delay only if no adulteration is discovered by the customers, can perform better than the conventional inspection mechanism in deterring product adulteration. Rui and Lai (2015) extend this analysis when the procurement quantity is endogenously determined along with the optimal contract.

Our work complements this literature along three dimensions. First, in contrast to the existing literature that focuses on monitoring and auditing of suppliers, our focus is on retailers. This increases the scale of the problem substantially (many more retailers than suppliers) making audits and inspections prohibitively expensive. Second, misaligned incentives in our context lead to quantity distortion rather than quality distortion. As a result, inspection mechanisms proposed in the literature to address quality distortion may not be applicable in our context. Third, barring a few exceptions (e.g., Mayer et al., 2004), existing literature is focused on developing qualitative insights based on analytical models whereas our focus is on rigorous empirical quantification. Specifically, we measure the value of installing a BA device at the last mile delivery (FPS) and show a statistically significant positive impact in reducing leakages. We also find that workarounds can reduce the ability of BA mechanisms in reducing leakages. Our findings show that the value of an imperfect monitoring (owing to the coexisting exception handling mechanism) on supply chain performance in terms of leakage is positive, and indicates that there is value, even if monitoring is imperfect.

The second, well established, stream of literature is comprised of analytical and empirical models that quantify the value of using downstream information for better upstream decision-making in supply chains to overcome the "Bullwhip effect" (Lee et al., 1997). Lee et al. (2000) find that a manufacturer can reduce its inventory cost by using the downstream sales and inventory information shared by the retailer in addition to the retailer's order quantity, especially when the demand is correlated over time. Subsequent papers show that a large part of this incremental value of shared information can be extracted from historical data on retailer's orders but the magnitude of this effect depends on the autoregressive structure of the demand distribution (Raghunathan, 2001; Gaur et al., 2005). Cachon and Fisher (2000) investigate various sources of value from implementing information technology in supply chains. They conduct an extensive numerical study to show that the value from improved decision making based on timely sharing of inventory and demand information is much lower than that from reduction in lead time and batch sizes. In contrast, Cui et al. (2015) use data from a consumer packaged goods company to empirically show that substantially reduction in forecast errors of future orders can be achieved if historic order data is augmented with downstream sales data from the retailers. The main driver behind this improvement is that the retailer does not follow the optimal inventory policy, as is typically assumed in the literature, but deviates systematically from it and these deviations are further propagated in the ordering data.

Our empirical context differs significantly from previous studies in a few crucial aspects. The upstream echelon in our supply chain already has visibility into actual sales at the downstream echelon, albeit delayed, and makes all replenishment decisions based on them instead of retailer orders. Thus, the value of timely information (for increasing retail sales and reducing inventory) estimated in our analysis is over and above that captured in the existing literature on decentralized supply chains. Our study shows that even in a centralized supply chain, such as the PDS, using timely information on sales (or, ending inventories) for planning replenishments can unlock substantial value.

We also acknowledge the extensive literature that studies how information technology applications (For example RFID, MEMS, POS and EDI) have led to multiple operational benefits (Rekik, 2010; Fan et al., 2015; Atalı et al., 2005; Lee et al., 2004; Smith, 2014). Although at a high level, our work adds to this large body of inquiry on value of information technology applications, the technology adopted in our context has different capabilities. The exact source of the incremental value of information technology in their context is reduced inventory shrinkage, increased productivity, decreased effort in information recording and processing, ease of computing, ease of communication, reduced discrepancies and reduced defects in received shipments vis-a-vis their respective orders and such and not improved supply chain decision making as in our context.

2.3 Empirical Setting

We study the operations of the PDS in the state of Karnataka, administered by the Department of Food, Civil Supplies & Consumer Affairs (DFCSCA). During our study period, September 2013 to December 2015, the PDS provided subsidized food grains to around 14 million beneficiary households. Each household was classified into one of the two categories–Antyodaya Anna Yojana (AAY) and Priority Households (PHH)–and was issued a corresponding *ration card*, which established their identity and entitlement. AAY consisted of households that faced extreme socioeconomic vulnerability and met a stringent set of inclusion criteria (e.g., homeless households, households headed by a minor or a disabled person) and were entitled to 35 kgs per month per household. PHH consisted of all other low-income households unless they met specific exclusion criteria based on their socioeconomic characteristics (e.g., ownership of assets and land) and were entitled to 5 kgs of subsidized grains per month per household member (Department of Food and Public Distribution, India, 2013).

Rice, the staple food grain of Karnataka, contributed to 75%–85% of the total entitlement and was distributed at the subsidized rate of INR 3 per kg. Wheat, which made up the remaining 15%–25% of the entitlement, was distributed at INR 2 per kg (Raju et al., 2018; NCAER, 2015). Other commodities such as millets, pulses, sugar and oil were also distributed occasionally but were not part of the regular entitlement.

During our study period, the PDS comprised of a network of more than 23,000 FPSs across 35 districts, each serving around 600 households on average. Roughly half of these were licensed to private individuals whereas the rest were licensed to cooperative societies (Government of Karnataka, 2015). Each household was assigned to an FPS from where it could collect its monthly entitlement. Next, we describe the process of sales and inventory replenishment at the FPSs (§ 2.3.1), the process of auditing and monitoring FPSs' performance (§ 2.3.2) and the impact of the intervention, i.e., installation of BA devices, on these processes (§ 2.3.3).

2.3.1 Sales and replenishment processes at the FPSs

In the absence of BA devices, at the beginning of each month (t), the DFCSCA provided each FPS with a paper copy of an updated *eligibility list* of households affiliated with it. During the course of the month, the FPS owners verified the identity of beneficiaries visiting the FPS by matching the names on their ration cards against those in the *eligibility list*. Upon verification, the FPS owner issued food grains to the beneficiaries as per their entitlement, and recorded the transaction in a paper-based register. The FPS owner also collected the beneficiary's signature or thumb impression against their name in the eligibility list as acknowledgement of the receipt of grains. At the end of the month, a copy of all sales transactions recorded in the register and the eligibility list containing beneficiary signatures was collected from all FPSs. These paper records were then consolidated, digitized and uploaded to a central database in the DFSCA headquarters by around the 15^{th} of the next month $(t+1)$.

The replenishment quantity for the next month $(t+1)$ at each FPS was calculated by the DFSCA centrally at the end of the current month (t). However, owing to the delayed processing of physical sales records described above, this calculation could not be based on the most updated information on ending inventory at the FPS and used the ending inventory at the end of the previous month $(t-1)$ instead. It calculated the replenishment quantity at each FPS for the next month (t+1) as the difference between the *Gross Requirement (GR)* of that FPS for that month (defined as the sum of entitlements of all households in its eligibility list) and the ending inventory of the previous month $(t-1)$. In summary, the inventory replenishment for each FPS was centrally determined and was characterized by an "order up to level," which was based on the total entitlement of the FPSs' beneficiaries and lagged information on its ending inventory.

2.3.2 Monitoring of the FPSs

The FPS owners were paid a commission ($\approx$ INR 1 per kg) based on the quantity of grains distributed to eligible beneficiaries, which was substantially lower than the market price of food grains ($\approx$ INR 25–30 per kg). This created a strong incentive for the FPS owners to exploit the weakness in the manual process of recording sales and to divert food grains away from the eligible beneficiaries to the open market for private gains. Key mechanisms deployed for this purpose included: (i) distributing lesser than the recorded quantity to beneficiaries who transacted in a given month, (ii) recording false transactions against beneficiaries who did not transact in a given month, (iii) recording transactions against fraudulent ration cards that do not belong to any genuine beneficiary (NCAER, 2015; Dreze and Khera, 2015).

To curtail diversion, FPSs were periodically audited by *food inspectors* through a manual and labor intensive process. It involved verification of ration cards of beneficiaries, signatures of beneficiaries in the sales registers and/or eligibility lists, and physical verification of the stock of food grains. Our field visits and discussions with the DFCSCA officials revealed that each food inspector was responsible for auditing 50 to 60 FPSs but was able to audit only 4 to 5 FPSs each month. As a result, each FPS was audited roughly once a year. The manual nature of the auditing process, along with its low frequency, severely compromised its ability to detect violations.

2.3.3 The Intervention: Installation of Biometric Authentication devices at the FPSs

Starting from January 2013, BA devices were installed in a phased manner in 3,332 out of 7,313 FPSs across 11 districts and were operational until December 2015 (Figure 2.1). The proportion of FPSs receiving the BA devices varied significantly across districts (7% to 84%). Our discussions with the DFCSCA officials indicated that FPSs receiving the BA devices were chosen based on availability of power, transportation and telecommunication (mobile and internet) connectivity, and proximity to *taluk* (sub-district) headquarters.

The BA device at each FPS had an electronic database of eligible beneficiary households comprising names of their members, their biometric identifiers (e.g., finger prints) and their entitlements. It enabled biometric authentication of the beneficiaries at the time of transaction by capturing their physical biometrics and matching it against those stored in the device. A sales transaction was recorded only upon a successful verification. This biometric authentication process was aimed at reducing the FPS owners' ability to divert food grains by recording fraudulent sales transactions as described above. Sales transactions were electronically transmitted to the central database thereby enabling real-time monitoring of inventory at the FPS. However, the upstream planning and replenishment policy was not modified to utilize this updated information and replenishment quantities continued to be determined using lagged inventory information collected through the manual process described in Section 2.3.1. Furthermore, to ensure that genuine beneficiaries were not denied grains in the event of a technical failure in the biometric authentication process, FPS owners were allowed to use the traditional paper-based sales register and eligibility list as a workaround mechanism.[1]

2.4 Data and Econometric approach

2.4.1 Data description

We use publicly available data related to PDS operations in Karnataka across 11 districts from September 2013 to December 2015. After excluding observations with missing data and outliers, we obtain an unbalanced panel dataset comprising 153,028 FPS-month observations corresponding to 6,934 FPSs (of which 3,075 FPSs received BA devices) over a period of 28 months.[2] Each observation includes gross requirement and ending inventory for rice and wheat for AAY and PHH households. In addition, it also includes the *Sales transaction transmission status* indicating whether a BA device was used to record sales transactions at a particular FPS in a particular month. For a given FPS, the first month with an active transmission status (if any) represents the month of installation of a BA device at that FPS.

Table 2.1 provides summary statistics on key operational variables of interest. We find that PHH beneficiaries contribute $\approx 90\%$ of the gross requirement. As mentioned

[1] A variety of reasons contributed to these failures. First, the quality of biometric images stored in the BA devices could be poor. Second, the sensors in the BA devices may not accurately capture finger prints that are worn out (e.g., among elderly beneficiaries and those involved in manual labor) or temporarily compromised (e.g., due to oily or sweaty hands). Third, sensors may experience failure due to extended exposure to dust and heat (Rosamma, Thomas, 2016; Ananda, Jonathan, 2017; Lakhani, Somya, 2018; Yadav, Anumeha, 2016). Our interaction with beneficiaries during our field visits confirmed the existence of authentication failures and the use of the workaround mechanism.

[2] We describe the details of the data cleaning procedure in § 2.9.1 of the Appendix and show that our results are robust to several alternative inclusion/exclusion criteria as shown in Table 2.16.

earlier, $\approx 80\%$ of the gross requirement is satisfied by rice, which is the staple food grain in Karnataka whereas the remaining $\approx 20\%$ is satisfied by wheat. Across all categories of households, FPSs (ownership and location) and grains, more than 99% of the gross requirement is recorded as sales with less than 1% remaining as inventory at the end of the month. Furthermore, observations with BA installation have a higher closing inventory (0.71%) than their Non BA counterparts (0.27%). We also find that, in observations with BA installation, 37% have closing inventory value of zero while in observations without BA installation, 64% have closing inventory value of zero. Roughly 60% of the FPSs in the dataset were in rural locations, while 45% of all the FPSs were managed by cooperative societies.

2.4.2 Econometric approach

The main outcome variable of interest, diversion of food grains, cannot be directly observed in the available data. We use change in recorded sales as a proxy for change in quantity of grains diverted. The validity of using recorded sales as a proxy is strengthened by the features of BA implementation briefly described earlier. First, one of the mechanisms used by FPS owners to divert food grains is recording false sales transactions against beneficiaries who do not transact in a given month. Second, the manual workaround mechanism ensures that any reduction in recorded sales cannot be attributed to genuine beneficiaries not receiving their entitlements due to technical failures of the BA devices. Finally, the unchanged replenishment policy after the BA installation ensures that the change in sales is not due to change in availability of food grains at the FPSs.

We use difference-in-differences (DID) methodology (see Imbens and Wooldridge, 2009; Angrist and Pischke, 2008; Lechner et al., 2011, e.g.) to compare the change in recorded sales at the FPS-month level between the treatment (FPSs with BA devices) and control (FPSs without BA devices) groups as an estimate of the *treatment effect*, i.e., the impact of installing BA devices on diversion at the FPSs. As mentioned in § 2.3.3, selection of FPSs into the treatment group was based on availability of power, transportation and telecommunication connectivity, and proximity to taluk headquarters. Hence, the composition of treatment and control groups may be different and a direct comparison of changes in recorded sales across these groups may not capture the true treatment effect.

We use a two-step approach that accounts for the difference in pre-treatment characteristics of FPSs across the two groups to obtain an unbiased estimate of the treatment effect (Austin, 2013; Rosenbaum and Rubin, 1983). In the first step, we estimate the probability that an FPS will be selected for installation of BA device (a propensity score), and assign weights to each observation in our data based on the FPS's propensity score. In the second step, we estimate the difference-in-differences (DID) model using the weighted observations across the two groups. This two-step procedure yields a "doubly robust estimator" that remains unbiased even if one of the models is misspecified (Funk et al., 2011; Słoczyński and Wooldridge, 2018).

We find that the range of propensity scores that are common to both control and treatment groups, is satisfactory (between 0.014 and 0.99) and comprises approximately 99% of our observations. We also find substantial reduction in the values of the standardized differences between the control and treatment groups after weighting of observations. This indicates that the weighting procedure improves the comparability of the treatment and control groups on observed covariates and hence improves the ability to draw causal inferences about the treatment effect.[3]

[3]See Appendix § 2.9.2 for details on propensity score estimation and weighting

2.5 Value of better monitoring using Biometric Authentication devices

As described earlier, we use change in recorded sales as a proxy measure for the outcome variable of interest – change in diversion of food grains. We calculate recorded sales for each FPS for each month ($S_{i,t}$) using the replenishment quantity ($Q_{i,t}$) and ending inventory for the current ($CB_{i,t}$) and previous period ($CB_{i,t-1}$) as follows:

$$S_{i,t} = CB_{i,t-1} + \underbrace{Q_{i,t}}_{=\ GR_{i,t}-CB_{i,t-2}} - CB_{i,t}. \tag{2.1}$$

As described in § 2.3.1, for both before and after installation of BA devices, the replenishment policy aims to bring the inventory available at the beginning of the current month t to be equal to $GR_{i,t}$ but using lagged ending inventory information ($CB_{i,t-2}$), resulting in $Q_{i,t} = GR_{i,t} - CB_{i,t-2}$.

To control for the size of the FPS, we normalize recorded sales by the gross requirement of the FPS ($SGR_{i,t} = \frac{S_{i,t}}{GR_{i,t}}$) and use the normalized value, $SGR_{i,t}$, as the dependent variable in our econometric analysis.

2.5.1 Overall treatment effect

We estimate the average treatment effect using the following DID specification:

$$SGR_{i,t} = \alpha_i + \beta_t + \delta BA_{i,t} + \varepsilon_{i,t}, \tag{2.2}$$

where $BA_{i,t} = 1$ if a BA device was in use at FPS i in month t and 0 otherwise. The variable α_i captures time-invariant FPS fixed effects such as location and accessibility of the FPS and the attitude of the FPS owner whereas β_t captures shop-invariant month fixed effects such as festival seasons, agricultural activity and seasonal migration for work. The average treatment effect, change in diversion as a fraction of the gross requirement due to BA installation, is given by δ.[4] We calculate cluster-robust standard errors (Bertrand et al., 2004) with two-way clustering at the *taluk* and month level to account for correlation in unobserved household (e.g., dietary preferences), administrative (e.g., inspection intensity), and market (e.g., prices) characteristics that vary over time and geography.[5]

Estimating the above model using gross requirement and sales transactions for rice, we find that installing BA devices reduced diversion by 0.39% of gross requirement (Column (1) in Table 2.2).[6] Note that this does not mean the magnitude of diversion before the intervention was 0.39% of gross requirement and that it has been completely eliminated after the intervention. In fact, the extent of diversion at the FPS level in Karnataka was estimated to be 9.5% of the entitlement (NCAER, 2015), which implies that the intervention was able to reduce the baseline diversion quantity by about 4% ($\approx 0.39/9.5$).[7] A possible explanation for the limited impact lies in the capabilities of the BA devices, the mode of their implementation and other contextual factors. First, BA devices cannot prevent FPS owners from exploiting the uneducated, unaware and disempowered beneficiaries by physically disbursing less than the weighed and recorded quantity. Second, FPS owners could exploit the manual workaround mechanism (originally intended to address technical failures) thereby limiting the effectiveness of the intervention in eliminating fraudulent transactions.

[4]We verify that *parallel trends*, a key identifying assumption in DID estimation, is satisfied for our model. See § 2.9.3 in the Appendix for details. We also ascertain that no other significant event that would bias the estimate occurred in Karnataka during the study period. We do this by carrying out an extensive review of news and media articles. See § 2.9.4 in the Appendix for details.

[5]All our results continue to hold if the errors are clustered at the district and month level.

[6]We also estimated the above model using total recorded sales for rice and wheat and found a similar effect (0.42% of the gross requirement). However, we found no significant effect on diversion of wheat. This could be due to smaller quantities of wheat distributed through PDS compared to rice, which may limit the statistical variation required for identification.

[7]We conduct falsification checks (Gertler et al., 2016; Kennedy, 2003) by randomly assigning FPSs to treatment and control groups, along with randomly chosen intervention periods. The insignificant results of these falsification checks confirm that the actual estimated effect of BA installation is not due to pure chance. See § 2.9.5 in the Appendix for details.

We conducted extensive field visits and interviews with beneficiary households, FPSs and Office of the food and civil supplies ministry to understand the factors which explain the limited impact we find. Several beneficiaries reported receiving less grains (than their entitlement quantity) even after the implementation of BA devices, due to the control and power that the FPS owners wield over beneficiaries. Furthermore, FPS owners seemed reluctant to use BA devices and cited multiple reasons like BA device breakdown, requirement of maintenance and power failures for the prevalent use of workaround mechanism. All the above combined together offer possible explanation for the lower effect size and the underwhelming impact of BA devices.[8]

Given the scale of the PDS, however, even this limited relative impact can be substantial in absolute terms when aggregated across the entire system. Given the average monthly gross requirement of 9.32 MT (i.e., 9,320 kg) of rice per FPS per month and 3,075 FPSs with BA devices installed, 0.39% translates to a reduction in diversion of 111.77 MT per month. Assuming that this would result in an equivalent reduction in procurement of grains by the state of Karnataka, the value of better monitoring is approximately INR 36.3 million per year ($\approx$ USD 0.56 million per year at an exchange rate of $\approx$ 65 INR per USD).[9] The potential value of monitoring if the intervention had been implemented in the entire state of Karnataka (23,241 FPSs and average gross requirement of 10.25 MT of rice) would have been USD 4.6 million. Accounting for the cost of implementation of around INR 20,000 per device (Business Today, 2016), these savings translate to a payback period of less than 21 months.[10]

[8] We have included a more detailed description of interviews with various stakeholders and the associated insights in Appendix § 2.9.6

[9] Economic cost of procurement, storage, transportation and distribution of grains during the period of analysis was INR 27.02 per kg of rice as per statistics published by the Government of India in the Food grains bulletin. See `http://dfpd.nic.in/writereaddata/images/pdf/food-grain/MAY-2014-040614.pdf` for details.

[10] The estimate may be lower than the actual savings as it does not include the reduction in diversion of other high-value commodities such as sugar and kerosene, which were not included in our analysis due to unavailability of data.

We also perform a variety of additional analyses to test the robustness of our results to changes in methodology and variable definitions and to rule out other plausible explanations. We verify that our results are not driven by anticipatory behavior of FPS owners (anticipation of implementation of BA devices) or due to the presence of FPSs without pre-intervention data. We also test the robustness of our results to different matching techniques. Finally we also confirm that our results do not change when we include market price as a covariate. See Appendix § 2.9.7 for details.

2.5.2 Evolution of treatment effect over time

In the preceding analysis, we estimated the average treatment effect over the implementation period. However, there may be systematic changes in the effect over time due to contextual factors. For instance, if FPS owners realize that using the paper-based workaround mechanism to record false sales transactions poses minimal risk, they may increase its usage over time. As a result, the impact of BA devices on curbing diversion may decrease over time. On the other hand, if true technical glitches reduce due to technological improvements and adjustments, the PDS administration may increase scrutiny of transactions conducted through the manual workaround mechanism. As a result, the impact may increase over time.

To investigate the time-varying impact of BA devices, we estimate the following specification:

$$SGR_{i,t} = \alpha_i + \beta_t + \sum_{j=0}^{10} \delta_j \, Short\_term_{i,j,t} + \delta_r \, Long\_term_{i,t} + \varepsilon_{i,t}. \tag{2.3}$$

$Short\_term_{i,j,t}$, for $j = 0, 1, \ldots, 10$, takes the value 1 if t represents the j^{th} month after installation of BA device at FPS i and 0 otherwise. Similarly, $Long\_term_{i,t}$ takes the value 1 if t represents 11^{th} or later month after installation of a BA device at FPS i and 0 otherwise.

Figure 2.2 shows that the impact of BA devices is statistically significant both in the short term (from 2 months to 7 months after installation) as well as in the longer term (beyond 10 months after installation), Moreover, the long term impact of BA devices in reducing diversion (0.5% of GR) is similar in magnitude to the average impact (0.4% of GR).[11]

2.6 Estimating heterogeneous effects to validate the mechanism of reduction in diversion

In this section, we validate the mechanism of reduction in decrease in diversion by estimating differential impacts based on various underlying behavioural channels which drive the quantity of diversion. Goldfarb and Tucker (2014) mention how mechanism checks are important in making causal identification claims more convincing. The paper suggest to estimate the effect separately based on whether an individual is a member of a group. This definition of the grouping and suggestion of which group experiences a bigger effect, comes from theory. If the effect is larger when theory suggests it should be, then this helps identify the mechanism. In the following subsections, we test three such mechanisms, differential impact based on - household category, FPS location and ownership and proportion of vulnerable households served by the FPS.

2.6.1 Heterogeneous effect across household categories

The treatment effect is likely to depend on the level of empowerment, education and awareness levels of the beneficiary households. To investigate this formally, we use the household category as a proxy, because AAY households are amongst the poorest

[11]These results are robust to alternate definition of "long-term" (beyond 5 months and 15 months instead of beyond 10 months). They are shown in Table 2.13 of the Appendix.

of the poor and are likely to be less educated and empowered compared to the PHH households. We suspect that such households are prone to exploitation and hence the treatment effect on these households is likely to be lower. We recast our data at the FPS-household category-month level and estimate the following triple difference model (2.4):

$$SGR_{i,j,t} = \alpha_i + \beta_t + \gamma\, AAY_j + \delta_p\, BA_{i,t} + \delta_a\, AAY_j \times\, BA_{i,t} + \varepsilon_{i,j,t}, \tag{2.4}$$

where $SGR_{i,j,t}$ represents the normalized sales value at FPS i for household category j in month t, AAY_j is a binary variable that takes the value of 1 for AAY entitlements and 0 otherwise. Thus, γ represents the difference in diversion for AAY entitlements before the intervention, δ_p represents the treatment effect on PHH entitlements and δ_a represents the incremental effect on AAY entitlements compared to PHH. We also a estimate a variant of this model without the interaction term to check the robustness of the average treatment effect of BA installation estimated by (2.2).

Column (2) of Table 2.2 shows that the average impact of BA installation remains virtually unchanged if the model is estimated at the FPS-category-month level and controlling for household category. The average treatment effect is comprised of effect on PHH and AAY household entitlements, as seen from Column (3) of Table 2.2. The impact on PHH entitlements (δ_p) is 0.17% and not statistically significant, whereas the impact on AAY households ($\delta_p + \delta_a$) is 0.53% and statistically significant (with $p < 0.05$ and std. error of 0.002). Interestingly, the difference between the two, δ_a, itself is not statistically significant. The findings from this analysis precludes our initial suspicion that the effect of BA installation on diversion differs across AAY and PHH households.

2.6.2 Heterogeneous effect by FPS location and ownership

The underlying benefit and risk for FPS owners from diversion of food grains is likely to differ across FPSs depending on their geographic location (rural vs. urban) and ownership (private individual vs. cooperative society). Urban FPSs may have easier access to markets and incur lower transactions cost in disposing of the diverted food grains compared to rural FPSs. Also, FPSs owned by cooperative societies may be subject to stricter community monitoring thereby making it riskier to divert food grains compared to those with private ownership (Nagavarapu and Sekhri, 2016). We use the following fourth order differences model (Imbens and Wooldridge, 2007) to estimate the differential treatment effect by FPS category:

$$\begin{aligned} SGR_{i,t} = \alpha_i + \beta_t + \delta_{u-c}\, BA_{i,t} + \delta_{u-pr}\, Private_i \times\ BA_{i,t} + \delta_{r-c}\, Rural_i \times\ BA_{i,t} \\ + \delta_{r-pr}\, Private_i \times\ Rural_i \times\ BA_{i,t} + \varepsilon_{i,t}, \end{aligned} \tag{2.5}$$

where $Private_i = 1$ if FPS i is operated by a private individual (and 0 otherwise), and $Rural_i = 1$ if FPS i is located in a rural area (and 0 otherwise).

Figure 2.3 (bars 1 to 4) shows the impact of BA installation on different FPS types based on location and ownership. For ease of interpretation, we calculate the net impact of the intervention for different types of FPSs using the appropriate linear combination of the coefficients in the above equation. The effect on urban privately owned FPSs, $\delta_{u-c} + \delta_{u-pr} = -0.24\%$, is not statistically significant while the effect on rural cooperative FPSs, $\delta_{u-c} + \delta_{r-c} = -0.38\%$, is statistically significant (with $p < 0.05$ and std. error equal to 0.0014). Finally, the impact on privately owned rural FPSs, $\delta_{u-c} + \delta_{u-pr} + \delta_{r-c} + \delta_{r-pr} = -0.36\%$ of GR, is also statistically significant (with $p < 0.01$ and std. error equal to 0.0012). In summary, we find that installing BA devices had a statistically significant impact on reducing diversion for all categories of FPSs except those operated by private individuals in urban locations. Column (1) of Table 2.11 in the appendix shows the coefficient estimates for this model.

To further elucidate the role of market access in moderating the treatment effect, we focus on the sub-sample of all rural FPSs in our dataset. We use the road distance between the village of the rural FPS and the nearest urban center as reported in the *Village and Town Amenities Dataset* as a proxy for market access.[12]

We use the following triple differences (difference-in-difference-in-differences) model to estimate the differential treatment effect by distance to market:

$$SGR_{i,t} = \alpha_i + \beta_t + \delta_n \, BA_{i,t} + \delta_f \, Farther_i \times \, BA_{i,t} + \varepsilon_{i,t}, \tag{2.6}$$

where $Farther_i$ is an indicator variable that takes the value 1 if the distance from a rural FPS i to its nearest urban market is greater than the median distance over all rural FPSs in our data ($\approx$ 20 kms) and 0 otherwise.

Figure 2.3 (bars 5 and 6) shows the impact of BA installation on different FPS types based on distance to market. For FPSs in rural locations that are less than 20 km from a urban center, installing BA devices reduces diversion by 0.33% of the gross requirement while the impact on FPSs that are farther away is significantly lesser. The net impact on FPSs that are more than 20 km away from an urban location is 0.15% of gross requirement, which is statistically significant (with $p < 0.05$ and std. error equal to 0.0007). This translates to reduction in diversion of 14 kgs per FPS per month, which is less than half of the impact on FPSs that are closer (31 kgs per FPS per month). Column (2) of Table 2.11 in the appendix shows the coefficient estimates for this model.[13] The findings from this analysis corroborate the role of market access in moderating the effect of BA installation on diversion.

[12]Market access is generally good for FPSs in urban areas. Moreover, our proxy measure of distance to market is zero for all urban FPSs by definition. As a result, market access is unlikely to offer meaningful explanation of the variation in the impact of the intervention across urban FPSs, both statistically as well as substantively.

[13]Our results are robust to alternate definitions of $Farther_i$ as being greater than the 60$^{\text{th}}$ and 75$^{\text{th}}$ percentile of distance between the rural FPS and its nearest urban market (25 kms and 31 kms, respectively). The results are shown in table 2.14 of the appendix.

2.6.3 Heterogeneous effect by proportion of vulnerable households in the FPS

Some studies and anecdotal evidence suggests that vulnerable households are more prone to being cheated by FPS owners. This vulnerability could be due to several reasons such as economic status, social status (caste and tribe) (Newman and Thorat, 2010; Sabharwal, 2011; Vaidya et al., 2014).

We use the following proxy variables of vulnerability and test for the behavioral mechanism:

1. Proportion of economically vulnerable households served by the FPS
2. Proportion of socially vulnerable class population with respect to caste and tribe in the village where the FPS resides [14]

We estimate a variant based on combination of the two dimensions of vulnerability. We define four such indicator variables where the two dimensions of economic and social vulnerability take two levels - low and high. We estimate the following equation:

$$\begin{aligned} SGR_{i,t} = \alpha_i + \beta_t + \delta_{ll}\, BA_{i,t} + \delta_{hev}\, FPS\ economic\ vulnerable\ group_i \times\ BA_{i,t} \\ +\delta_{hsv}\, FPS\ social\ vulnerable\ group_i \times\ BA_{i,t} \\ +\delta_{hh}\, FPS\ combined\ vulnerable\ group_i \times\ BA_{i,t} + \varepsilon_{i,t}, \end{aligned} \tag{2.7}$$

where $FPS\ economic\ vulnerable\ group_i = 1$ if FPS i belongs to a group with proportion of AAY households greater than the defined threshold (and 0 otherwise), $FPS\ social\ vulnerable\ group_i = 1$ if FPS i belongs to a group with proportion of socially vulnerable group population greater than the defined threshold (and 0 otherwise) and $FPS\ combined\ vulnerable\ group_i = 1$ if FPS i belongs to a group with proportion of both AAY households and socially vulnerable group population

[14] For details on how we define the proxy variables, see Appendix § 2.9.9

greater than the defined threshold (and 0 otherwise). We define threshold proportions as the values associated with 90^{th} percentile.

Figure 2.4 shows the impact of BA installation on different FPS groups based on proportion of vulnerable households. We find that the impact on FPSs with low proportion of vulnerable households on both dimensions, high proportion of economically vulnerable households and FPSs with high proportion of socially vulnerable households are all statistically significant. The impact on both high proportion of economically vulnerable households and high proportion of socially vulnerable households is lower compared to FPSs with low proportion of vulnerable households on both dimensions. The impact on FPSs with high proportion of vulnerable households on both dimensions is insignificant. Column (2) of Table 2.12 in the appendix shows the results from estimating the above equation. The findings from this analysis confirm our initial conjecture that the impact of BA installation is lower in FPSs with high proportion of vulnerable households indicating that vulnerable households are indeed at a disadvantage and more prone to cheating by FPS owners.

2.7 Value of better planning using timely inventory information

The BA devices in our context were equipped with the additional capability of recording real-time information on sales and inventory at the FPSs. This information captured by the BA devices can also be utilized to determine replenishments to the FPSs and create additional *value of better planning*. As discussed in § 2.3.1, in the absence of BA devices, lagged sales and inventory information captured manually at the FPSs, was used to create a centralized plan of inventory replenishments to FPSs. This method can lead to demand-supply mismatches even without misaligned incentives and concomitant distortion and incompleteness of information commonly observed in decentralized private sector supply chains (Lee et al., 1997, 2000; Cui et al., 2015). Implementation of BA devices could provide the central planner with access to timely information on sales and inventory and thereby improve replenishment planning to FPSs.

We are unable to empirically quantify this value of better planning in our study setting as the Government of Karnataka continued to use lagged inventory information to determine replenishment quantities even after the implementation of BA devices. Hence, we conduct an extensive simulation study toward this end under different values of key input parameters that are representative of the characteristics of the PDS in various Indian states. The effect of monitoring that we find from our empirical analysis directly feeds into the design of our simulation analysis and connects the two analyses of monitoring and planning evaluations.

Next, we describe our study design (§ 2.7.1), choice of input parameters (§ 2.7.2) followed by the discussion of our results and associated insights (§ 2.7.3).

2.7.1 Simulation design

To quantify the incremental value of better planning, we compare the performance of the PDS supply chain under three scenarios: (i) *Baseline*, which simulates the planning and monitoring processes in the absence of BA devices, (ii) *Monitoring*, which simulates improved monitoring (reduced diversion) due to BA devices but unchanged planning process, and (iii) *Planning*, which simulates improved planning (calculation of replenishments using timely inventory information) in addition to improved monitoring. In the absence of granular operational data on the temporal co-evolution of the demand fulfilment (DF) and diversion of grains (DG) at the FPS, we simulate each of the above three scenarios under these two extreme cases and obtain upper and lower bounds on the value of better planning.

In the *Baseline* scenario, for each FPS $i \in \{1, 2, \ldots, I\}$ and month $t \in \{1, 2, \ldots, T\}$, we simulate genuine demand for food grains from beneficiaries and the quantity of grains that the FPS owner intends to divert as proportions of the gross requirement. We denote these by $D_{i,t} = p_{i,t}\, GR_{i,t}$ and $\hat{L}^{(b)}_{i,t} = l^{(b)}_{i,t}\, GR_{i,t}$, respectively. We assume that $p_{i,t}$ and $l^{(b)}_{i,t}$ follow a uniform distribution with support on subsets of $(0, 1)$. The replenishment quantity for FPS i and month t is calculated as $Q^{(b)}_{i,t} = GR_{i,t} - CB^{(b)}_{i,t-2}$ denoting the use of lagged inventory information under the *Baseline* scenario. This brings the inventory of food grains available at the beginning of month t to $OB^{(b)}_{i,t} = CB^{(b)}_{i,t-1} + Q^{(b)}_{i,t}$. Under the $DF \rightarrow DG$ assumption, sale against demand from genuine beneficiaries occurs first and is given by $\hat{S}^{(b)}_{i,t} = \min\left\{D_{i,t}, OB^{(b)}_{i,t}\right\}$, followed by diversion of grains given by $L^{(b)}_{i,t} = \min\left\{\hat{L}^{(b)}_{i,t}, OB^{(b)}_{i,t} - \hat{S}^{(b)}_{i,t}\right\}$. In contrast, under the $DG \rightarrow DF$ assumption, diversion of grains occurs first given by $L^{(b)}_{i,t} = \min\left\{\hat{L}^{(b)}_{i,t}, OB^{(b)}_{i,t}\right\}$, which is followed by sales against demand from genuine beneficiaries given by $\hat{S}^{(b)}_{i,t} = \min\left\{D_{i,t}, OB^{(b)}_{i,t} - L^{(b)}_{i,t}\right\}$. In both cases, the recorded sales is a sum of genuine sales and diversion and is given by $S^{(b)}_{i,t} = \hat{S}^{(b)}_{i,t} + L^{(b)}_{i,t}$, while the ending inventory is given by $CB^{(b)}_{i,t} = OB^{(b)}_{i,t} - S^{(b)}_{i,t}$.

Under the *Monitoring* scenario, all calculations are similar to those described above except the quantity of grains that the FPS owner intends to divert. It is given by $\hat{L}^{(m)}_{i,t} = l^{(m)}_{i,t}\, GR_{i,t}$, where $l^{(m)}_{i,t}$ follows a uniform distribution whose mean is lower than that of $l^{(b)}_{i,t}$ to reflect the value of better monitoring (lower diversion of grains) due to BA devices. Similarly, all calculations in the *Planning* scenario are identical to that in the *Monitoring* scenario except that the replenishment quantity is calculated using updated inventory information and is given by $Q^{(p)}_{i,t} = GR_{i,t} - CB^{(p)}_{i,t-1}$.

We use the results from the above simulation to calculate two intermediate outputs: (a) average reduction in the inventory levels, $\frac{\sum_{i,t}\left(CB^{(m)}_{i,t} - CB^{(p)}_{i,t}\right)}{I \times T}$ and (b) average increase in genuine sales to beneficiaries, $\frac{\sum_{i,t}\left(\hat{S}^{(p)}_{i,t} - \hat{S}^{(m)}_{i,t}\right)}{I \times T}$. We use these intermediate outputs to calculate the reduction in inventory holding cost for the PDS and financial cost of open market purchase for households and characterize their sum as the incremental value of better planning. Finally, we calculate the ratio of the value of better planning estimated from the simulation model and the value of better monitoring (reduced diversion) estimated from the empirical models (§ 2.5) to compare their magnitudes.

2.7.2 Simulation parameters

For each of the 6,934 FPSs ($i = 1, \ldots, 6,934$) in our dataset described in §2.4.1, we initialize the simulation model with ending inventory values for the first two months $t = 1, 2$. We then use these in conjunction with the gross requirements for 28 months to simulate the inventory, sales and diversion dynamics for subsequent months $t = 3, \ldots, 28$ as described in §2.7.1.

We choose different supports for demand ($p_{i,t}$) and diversion in the *Baseline* scenario ($l^{(b)}_{i,t}$) such that all feasible combinations of their means approximately match the high, medium and low levels of the reported values of these entities across Indian states (Khera, 2011; NCAER, 2015; Dreze and Khera, 2015; Gulati and Saini, 2015).

In particular, we consider all feasible combinations of $\mathbf{E}(p_{i,t}) \in \{0.45, 0.75, 0.90\}$ and $\mathbf{E}(l^b_{i,t}) \in \{0.05, 0.20, 0.40\}$ such that $p_{i,t}+l^b_{i,t} \leq 1$. For each combination, we choose the support for $l^{(m)}_{i,t}$ such that the difference in the average quantity actually diverted between the *Baseline* and *Monitoring* scenarios, $\frac{\sum_{i,t}\left(L^{(b)}_{i,t}-L^{(m)}_{i,t}\right)}{I\times T}$, is approximately equal to our empirical estimate of the value of monitoring from BA devices, i.e, 0.4% of $GR_{i,t}$ (Table 2.2).

We assume an inventory holding cost of 8% per year and open market price of INR 28 per kg of rice to compute the monetary value of operational improvements.[15] For each combination of input parameters, in each scenario and each sequence of events,

[15]For cost of capital, we use the average term deposit bank interest rates of State Bank of India during our study period. See `https://www.sbi.co.in/portal/web/interest-rates/old-interest-rates-last-10-years` for details. For open market price, we use the average wholesale market price of rice in Karnataka during our study period, available on the AGMARKNET portal which collects, analyses and disseminates market information to farmers, traders and Policy makers. See `http://agmarknet.gov.in/Default.aspx` for details.

we conduct 1,000 simulation runs and report the average value of the performance metrics described in §2.7.1.

2.7.3 Simulation results

Table 2.3 shows the ratio of *value of better planning* and *value of better monitoring* for various combinations of demand and diversion stated above under $DF \rightarrow DG$ and $DG \rightarrow DF$ cases. This ratio ranges from around 0.1 to around 4.5 depending on the magnitude of the genuine demand from beneficiaries and the extent of diversion in the *Baseline* scenario and is non-monotone with respect to demand and diversion. On average, for the former, the value of better planning is almost equal to that of the value of better monitoring whereas for the latter it is thrice as much. To better understand the mechanisms behind these findings, we analyze the value from reduced inventory holding costs and that from reduced lost sales separately.

Figure 2.5 shows that, for a given level of demand (diversion), the benefit from reduced inventory is lower if the extent of diversion (demand) in the *Baseline* scenario is greater. All else being equal, greater sum of demand and diversion results in ending inventory being closer to zero in consecutive time periods and displaying lower variability across periods. In such a situation, i.e., if $|CB_{i,t-2} - CB_{i,t-1}|$ is small, using lagged ($CB_{i,t-2}$) or updated ($CB_{i,t-1}$) inventory information (corresponding to *Monitoring* and *Planning* scenarios) results in similar replenishment quantities. Since this mechanism is driven almost exclusively by the sum of demand and diversion in *Baseline* scenario, the magnitude of this effect is not different between $DF \rightarrow DG$ and $DG \rightarrow DF$ assumptions. To demonstrate the underlying mechanism more clearly, we plot the value from reduced inventory for each combination of demand and diversion against the average value of $(\max_t CB_{i,t} - \min_t CB_{i,t})$ in the *Monitoring* scenario (Figure 2.6). As expected, we observe a monotone behavior between the two, which is almost identical for both $DF \rightarrow DG$ and $DG \rightarrow DF$ assumptions.

Figure 2.7 shows that the benefit from reduced lost sales, for low level of demand (diversion), is non-monotonic in the extent of diversion (demand). To understand this observation, consider the following two extreme scenarios. On the one hand, when the sum of demand and diversion is high, as explained above and seen in Figure 2.6, variation in ending inventory across periods is low. Hence, the value obtained from reduction in lost sales is low too. On the other hand, when demand and diversion are both low, although there is greater variation in ending inventory across periods, the probability of occurrence of lost sales itself is low due to low levels of demand. Hence the value from reduction in lost sales is again low. For intermediate levels of demand and diversion, both variability in ending inventory across periods as well as the probability of occurrence of lost sales are sizeable. Hence, the value obtained from reduced lost sales is high under these scenarios. Finally, comparing Figures 2.7a and 2.7b shows that the benefit from reduced lost sales due to better planning is higher for $DG \rightarrow DF$ as shortage in availability has greater impact on sales to genuine beneficiaries under this sequence.

We also find an interesting interaction between diversion and using updated information for planning. We find that the actual diversion between the *Monitoring* ($L_{i,t}^{(m)}$) and *Planning* ($L_{i,t}^{(p)}$) scenarios increases under both $DF \rightarrow DG$ and $DG \rightarrow DF$ assumptions. The average increase between *Monitoring* and *Planning* scenarios is 1.01% under $DF \rightarrow DG$ and 0.07% under $DG \rightarrow DF$. This increase is attributed to the increase in availability of grains under the alternate replenishment policy, which in turn, results in more opportunities to divert the grains. We note that average increase in actual diversion is higher under the assumption $DF \rightarrow DG$. Under the $DF \rightarrow DG$ assumption, the FPS owner diverts grains before fulfilling demand. As the quantity diverted as a fraction of GR is small ($\approx 10\%$), the improved availability due to better planning has lesser impact on quantity diverted, compared to $DF \rightarrow DG$. Figure 2.8 shows the percentage increase in actual diversion between planning and monitoring scenarios for different levels of demand and diversion.

2.8 Conclusion

This paper employs an operational lens to evaluate technology-based interventions to improve the performance of public welfare delivery supply chains along two dimensions: *better monitoring* and *better planning*. We find that installation of biometric authentication devices had an economically and statistically significant impact on diversion of grains despite the presence of a manual workaround mechanism (originally intended to prevent interruptions in grain distribution due to technology failures), which can be exploited to continue diversion. Furthermore, the timely information on sales and inventory captured due to the additional capabilities of the BA devices in our context can be leveraged to create additional value through better planning of replenishments even in a centrally managed supply chain without the traditional information distortions due to misaligned incentives. In fact, in some cases, this additional *value of better planning* can be substantially higher than the *value of better monitoring*. Findings from our study have important implications to managers and policy makers that are generalisable beyond the context of the PDS and biometric authentication. One could directly extend our results to similar scenarios where biometric authentication is used and which allows the collection and storage of real time sales information. Our estimate of the proportions of value of monitoring and replenishment decisions from BA devices can guide policy makers facing similar scenarios to make appropriate implementation decisions. We believe that our results can be generalised at multiple levels - other Indian states within India's food security program, food security programs of other countries and other welfare benefit programs and schemes.

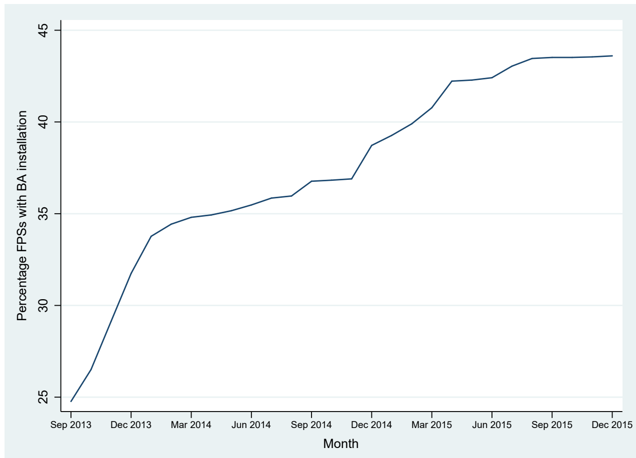
These findings illustrate how operations and supply chain frameworks can complement the traditional impact evaluation approach (Gertler et al., 2016).

First, these frameworks facilitate a careful analysis of the interactions between new technology interventions and the underlying processes, which can unlock additional sources of potential value that are not explicitly captured in the traditional approach. In our context, ignoring the additional value of better planning and focusing only on the value of better monitoring may severely underestimate the value of biometric authentication devices with the capability to record sales information, in other welfare programs that experience diversion in manual paper-based systems (e.g., Barnwal, 2016; Muralidharan et al., 2016). Second, these frameworks can provide insights into why different implementations of the same technology may result in differential effectiveness (see Allu et al., 2019, for heterogeneity in the implementation of BA in PDS).

Our paper provides early empirical evidence on the positive impact of technology on monitoring and planning in public sector supply chains suffering from diversion. These findings lead to interesting theoretical questions around the interplay of poor planning and monitoring due to lack of information and misaligned incentives. For instance, incentives may lead individual players to undertake hidden action (e.g., diversion by FPS owners), which in turn may distort the information contained in the sales signal thereby adversely affecting replenishment planning decisions. Future research should explore if the information captured by new technology (e.g., BA devices with capability to record sales information) can be used to design innovative contracts to mitigate the twin problems of poor monitoring and planning.

Tables and Figures

Figure 2.1: Scale-up of BA installation over time



Notes: This figure shows the change in cumulative fraction of FPSs with a BA device over time in the 11 districts of Karnataka.

Table 2.1: Summary Statistics on key operational variables of interest

	Household category		**FPS Location**		**FPS ownership**		**Grain type**		**Total**
	AAY	**PHH**	**Urban**	**Rural**	**Private**	**Co-op**	**Rice**	**Wheat**	
Gross Requirement (MT)	1.36	10.49	9.83	13.13	11.01	12.92	9.32	2.53	11.86
Recorded Sales (MT)	1.35	10.43	9.73	13.08	10.95	12.84	9.27	2.51	11.79
Ending Inventory (kg)	5.27	49.81	40.95	77.59	46.46	65.91	40.44	14.65	55.09
Observations	153,028	153,028	59,085	93,943	85,079	67,949	153,028	153,028	153,028

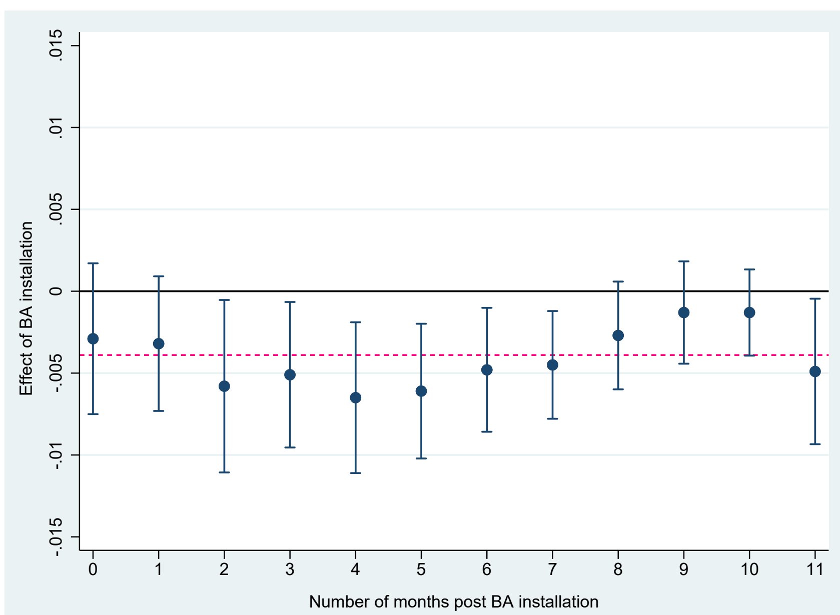
Notes: 1) MT denotes Metric Tonne (= 1000 kg). 2) All numbers for Gross Requirement, Recorded Sales and Ending Inventory represent averages over all observations at the FPS-month level.

Table 2.2: Overall impact of installing BA devices on curbing diversion of rice

	Overall impact		**Impact at**
	Without controlling for HH-category	**After controlling for HH-category**	**HH-category level**
Overall impact (δ)	-0.0039**	-0.0035**	
	(0.0017)	(0.0015)	
Impact on PHH entitlements (δ_p)			-0.0017
			(0.001)
Additional impact on AAY entitlements (δ_a)			-0.0036
			(0.0024)
R^2	0.167	0.061	0.061
Adjusted R^2	0.128	0.039	0.039
Observations	152334	284,450	284,450

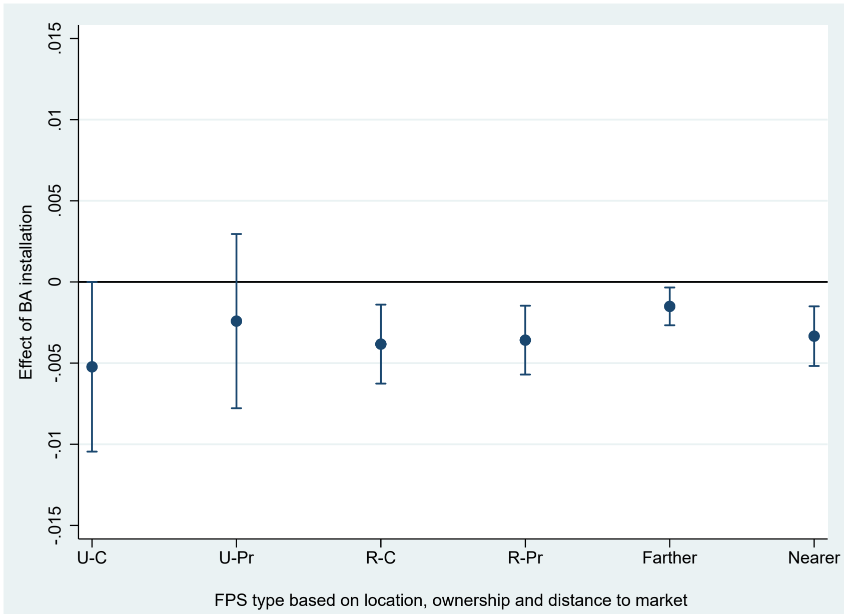
Notes: **1)** Results shown are for FPS-month panel regressions with FPS and month fixed effects and standard errors clustered at the month and taluk levels. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** Column (1) shows the results from estimating equation (2.2) with the dependent variable $SGR_{i,t}$. **4)** Columns (2) and (3) show results from models using $SGR_{i,j,t}$ as the dependent variable, where j denotes the household category (AAY or PHH). **5)** The results in Column (2) are from estimating a model $SGR_{i,j,t} = \alpha_i + \beta_t + \gamma AAY_j + \delta BA_{i,t} + \epsilon_{i,t}$, where $AAY_j = 1$ if the household category is AAY and 0 otherwise. **6)** Column (3) shows the result from estimating equation (2.4) and the impact by household category. The impact of the intervention on diversion from entitlements of AAY category households is given by $\delta_p + \delta_a = -0.0053^{**}$.

Figure 2.2: Impact of installing BA devices on curbing diversion of rice over time



Notes: **1)** The figure is based on the coefficient estimates of the model specification (2.3). **2)** The X-axis indicates the number of months post installation of a BA device, with 11 indicating 11 months or more. The Y-axis shows the impact of BA installation along with the 90% confidence intervals. **3)** The dotted line indicates the average treatment effect value of -0.0039.

Figure 2.3: Differential impact of installing BA devices on curbing diversion by FPS type based on location, ownership and distance to market



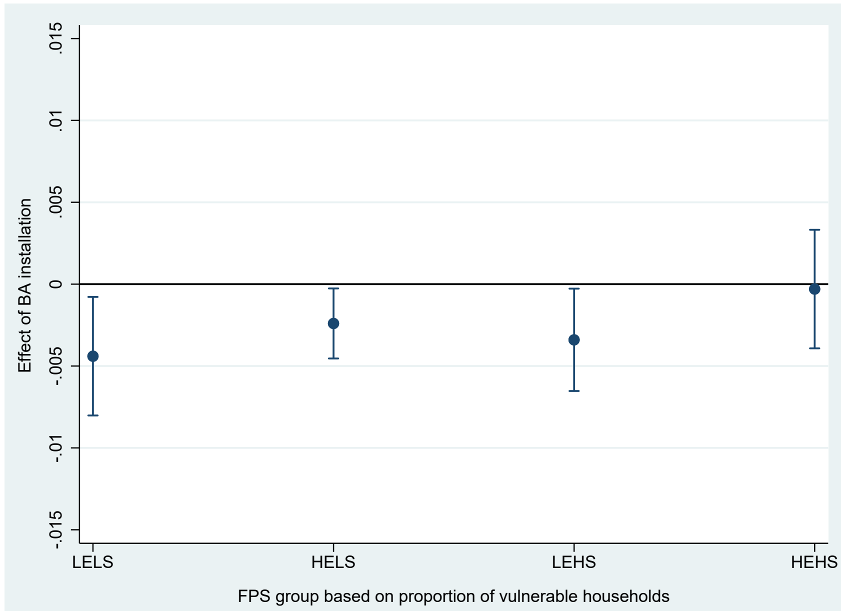
Notes: **1)** The figure is based on the coefficient estimates of the model specification (2.5). **2)** The X-axis indicates different types of FPS based on location, ownership and distance to market. U-C indicates urban-cooperative, U-Pr indicates urban-private, R-C indicates rural-cooperative, R-Pr indicates rural-private, Nearer indicates nearer to the market and Farther indicates farther to the market. The Y-axis shows the impact of BA installation along with the 90% confidence intervals.

Table 2.3: *Value of better planning* relative to *Value of better monitoring* for different levels of demand and diversion

Average diversion	**Average demand**	*Value of better planning* as a fraction of *Value of better monitoring*	
		$DF \rightarrow DG$	$DG \rightarrow DF$
0.05	0.45	0.61	0.89
0.05	0.75	3.23	4.63
0.05	0.90	0.78	2.27
0.20	0.45	1.00	4.00
0.20	0.75	0.11	2.24
0.20	0.90	NA	NA
0.40	0.45	0.24	3.92
0.40	0.75	NA	NA
0.40	0.90	NA	NA
	Average	**0.99**	**2.99**

Notes: **1)** Average demand and average diversion are expressed as a proportion of gross requirement. **2)** Average demand refers to the demand from genuine beneficiaries while average diversion refers to the level of diversion in the *Baseline* scenario. **3)** NA denotes that the particular combination of demand and leakage values is not possible as it exceeds 100% of gross requirement. **4)** $DF \rightarrow DG$ denotes demand fulfillment precedes diversion of grains while $DG \rightarrow DF$ denotes diversion of grains precedes demand fulfillment.

Figure 2.4: Differential impact of installing BA devices by FPS group based on proportion of vulnerable households

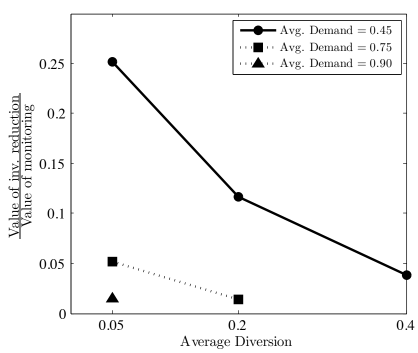


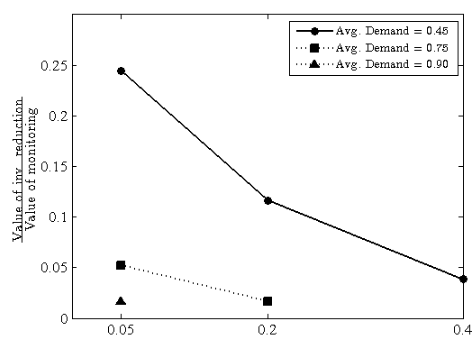
Notes: **1)** The figure is based on the coefficient estimates of the model specification (2.6). **2)** The X-axis indicates different types of FPS based on proportion of vulnerable households. LELS indicates low proportion of both economic and socially vulnerable households, HELS indicates high proportion of economic and low proportion of socially vulnerable households, LEHS indicates low proportion of economic and high proportion of socially vulnerable households and HEHS indicates high proportion of both economic and socially vulnerable households. The Y-axis shows the impact of BA installation along with the 90% confidence intervals.

Figure 2.5: *Value of better planning* in terms of lower inventory cost relative to *Value of better monitoring* for different levels of demand and diversion

(a) ($DF \rightarrow DG$)

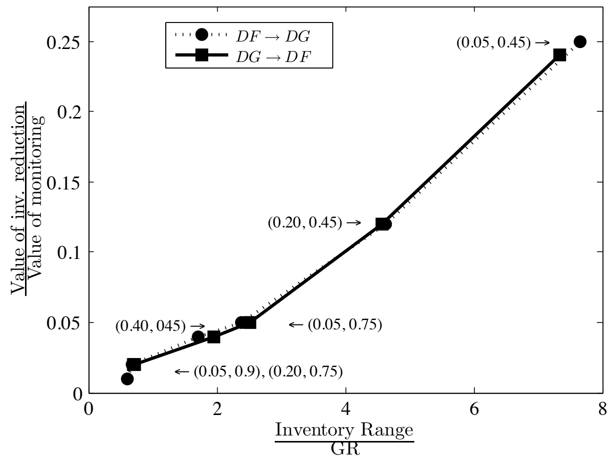
(b) ($DG \rightarrow DF$)





Notes: **1)** The X-axis corresponds to the average diversion level under the *Baseline* scenario while the different lines correspond to different average demand from genuine beneficiaries, with diversion and demand expressed as a fraction of the gross requirement. **2)** The Y-axis indicates the value from lower inventory holding cost as a fraction of the value from better monitoring for different (feasible) combinations of average demand and diversion levels.

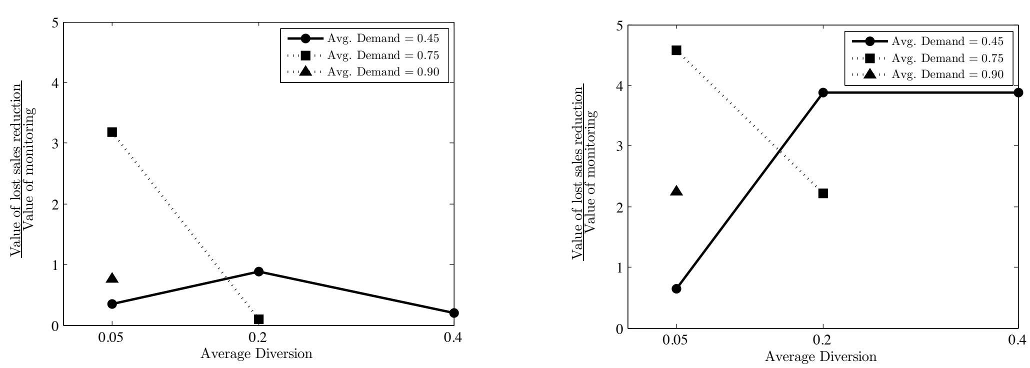
Figure 2.6: *Value of better planning* in terms of lower inventory cost relative to *Value of better monitoring* as a function of inter-temporal variability of ending inventory levels



Notes: **1)** The X-axis shows the average value of $(\max_t CB_{i,t} - \min_t CB_{i,t})$ as a fraction of the gross requirement for different demand and diversion combinations. **2)** The labels in the figure refer the average diversion and demand values corresponding to the point; e.g., the $(0.05, 0.45)$ in the top right hand corner refers to the case where average diversion and demand in the *Baseline* scenario are equal to 5% and 45% of gross requirement respectively.

Figure 2.7: *Value of better planning* in terms of lower lost sales cost relative to *Value of better monitoring* for different levels of demand and diversion

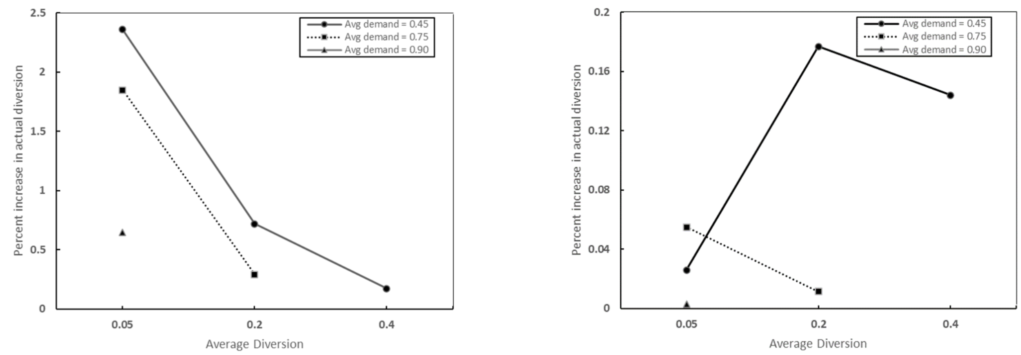
(a) $(DF \rightarrow DG)$ (b) $(DG \rightarrow DF)$



Notes: **1)** The X-axis corresponds to the average diversion level under the *Baseline* scenario while the different lines correspond to different average demand from genuine beneficiaries, with diversion and demand expressed as a fraction of the gross requirement. **2)** The Y-axis indicates the value to beneficiaries from reduction in lost sales as a fraction of the value from better monitoring for different (feasible) combinations of average demand and diversion levels.

Figure 2.8: Percentage increase in actual diversion between planning and monitoring scenarios for different levels of demand and diversion

(a) ($DF \rightarrow DG$) (b) ($DG \rightarrow DF$)



Notes: **1)** The X-axis corresponds to the average diversion level under the *Baseline* scenario while the different lines correspond to different average demand from genuine beneficiaries, with diversion and demand expressed as a fraction of the gross requirement. **2)** The Y-axis indicates the percentage increase in actual diversion between planning and monitoring scenarios for different (feasible) combinations of average demand and diversion levels.

2.9 Appendix

2.9.1 Data cleaning procedure and model variants based on treatment of outliers

The original data set scraped from publicly available source consists of 172,305 FPS-month observations corresponding to 6,934 FPSs. We check for various data errors to determine the final set of observations to use in the analysis.

We find 2,246 observations in the overall data set that have $GR_{i,t}$ equal to zero. As $GR_{i,t}$ represents the maximum possible demand that a FPS can experience, it is unlikely an operational FPS will have $GR_{i,t}$ equal to zero. Therefore, we discard these observations.

By definition of $S_{i,t}$ and $GR_{i,t}$, we have $\frac{S_{i,t}}{GR_{i,t}} \leq 1$. However, in the data set we find observations with $\frac{S_{i,t}}{GR_{i,t}} > 1$. We verify that observations with $\frac{S_{i,t}}{GR_{i,t}} > 1$ do not follow any pattern and are randomly distributed across FPSs, districts, months, treatment and control groups and in periods before and after BA installation. Further, the mean value by which $S_{i,t}$ exceeds $GR_{i,t}$ across these observations is 0.011 quintals

(1.1 kg) which is insignificant compared to the magnitude of average $GR_{i,t}$ and $S_{i,t}$. The distribution of $\frac{S_{i,t}}{GR_{i,t}}$ values for all observations with $\frac{S_{i,t}}{GR_{i,t}} > 1$ is shown in Figure 2.14, while the descriptive statistics of these observations is given in Table 2.15. These observations are likely because of data entry errors and we do not consider them in our analysis. Of the remaining 153,028 observations, 152,334 pertain to FPSs that belong to the common support and form the final data set for our analysis.

We define different model variants based on several alternate inclusion/exclusion criteria with respect to observations that have $\frac{S_{i,t}}{GR_{i,t}} > 1$. As recorded sales and gross requirement data is available at the FPS-household category-month level, we can have observations with $\frac{S_{i,t}}{GR_{i,t}} > 1$ and/or $\frac{S_{i,j,t}}{GR_{i,j,t}} > 1$ where j represents the household category. The results estimating equation (2.2) under different inclusion/exclusion criteria are summarized in Table 2.16, and indicate that our results are robust to different combinations of inclusion/exclusion criteria.

2.9.2 Propensity score estimation and weighting

Logistic regression

We use logistic regression to calculate the propensity score for each FPS. We use three sets of variables from the *Village and Town Amenities Dataset* of the district census handbook, published by the Office of the Registrar General & Census Commissioner, India, as proxies for the criteria used by the DFCSCA for selecting FPSs for BA installation. These include: (i) higher education institutions to proxy for socioeconomic development, (ii) mobile phone coverage and power availability to proxy for technological connectivity, and (iii) public transportation and roads to proxy for logistical connectivity.

We use the following logistic regression model to calculate the propensity scores, i.e., probability of a FPS being selected for installation of BA device, described in § 2.9.2:

$$p_i = \frac{\exp(\alpha\, SD_i + \beta\, TC_i + \gamma\, LC_i + \varepsilon_i)}{1 + \exp(\alpha\, SD_i + \beta\, TC_i + \gamma\, LC_i + \varepsilon_i)}, \tag{2.8}$$

where p_i is the probability that FPS i is selected for treatment, i.e., $BA_i = 1$. In the above specification, SD_i comprises of all covariates related to size and development, TC_i comprises of all covariates related to technological connectivity and LC_i comprises of all covariates related to logistical connectivity. The results of the logistic regression and the model statistics are shown in Table 2.4. Results indicate that all variables except availability of private polytechnic college, availability of telephone and availability of district roads significantly predict the chances of getting selected into treatment.

Figure 2.9 shows the Receiver Operator Characteristic (ROC) curve for the logistic regression model to assess the goodness of fit. The area under the ROC curve is 0.73, which indicates that the model's ability to discriminate between FPSs in the treatment and control group is acceptable (Hosmer Jr et al., 2013). We also find that the common support (Austin, 2009), range of propensity scores that are common to both control and treatment groups, is satisfactory (between 0.014 and 0.99) and comprises approximately 99% of our observations. Restricting attention to only FPSs with propensity scores in the common support region results in 152,334 observations for our empirical analysis. Figure 2.10 shows the distribution of propensity scores for the control and treatment groups. To ensure comparability between the treatment and control groups, we only include FPSs that have propensity scores in the common support; the range of propensity scores that are common to both control and treatment groups.

Inverse probability weighting

Despite restricting observations to only FPSs in the common support region, the treatment and control groups may still not be directly comparable. In general, the propensity scores of FPSs in the treatment (control) group are likely to be high (low). Hence, observations in the treatment (control) group with low (high) propensity score are more similar to those in the control (treatment) group. The observations can be made comparable across the two groups if we assign low weight to observations in the treatment group with high propensity score and those in the control group with low propensity score. The *inverse probability weighting* method (Austin, 2011; Austin and Stuart, 2015) achieves this by weighing observations belonging to FPS i in the treatment group by $w_i = \frac{1}{p_i}$ whereas those related to FPS k in the control group by $w_k = \frac{1}{1-p_k}$, where $p_{(\cdot)}$ is the propensity score of an FPS.

We assess the appropriateness of this method by comparing the standardized differences in pre-treatment characteristics between the treatment and control groups before and after the weighting (Austin, 2009).[16] Figure 2.11 shows a substantial reduction in the values of the standardized differences between the control and treatment groups after weighting of observations, with the maximum value of the difference after weighting (0.091) being well within acceptable limits suggested in the literature (Austin, 2009; Garrido et al., 2014). This indicates that the weighting procedure improves the comparability of the treatment and control groups on observed covariates and hence improves the ability to draw causal inferences about the treatment effect.[17]

We illustrate the importance of using inverse propensity score weighting by considering two variations of the difference-in-differences estimation: (i) including all observations and without weighting observations, and (ii) including only FPSs

[16] Austin (2009) defines standardized difference (also referred to as Cohen's effect size) as the distance between the treatment and control group means of a covariate. It is given by $d = \frac{|\bar{x}_{treat} - \bar{x}_{control}|}{\sqrt{(s^2_{treat} + s^2_{control})/2}}$, where $\bar{x}_{treat}$ and $\bar{x}_{control}$ denote the means while s^2_{treat} and $s^2_{control}$ denote variance of treatment and control group observations, respectively.

[17] We employ two variations of our difference-in-difference approach: (i) without any matching or propensity score weighting, and (ii) excluding observations outside of common support but without propensity score weighting. The results corresponding to these variants (Table 2.5 in the Appendix) indicate that not using inverse propensity score weighting would have resulted in downward biased estimates.

in the common support but without weighting the observations. Columns (1) and (2) in Table 2.5 show the results for these two variations respectively. The results indicate that not matching control and treatment observations appropriately would have resulted in a downward bias in the estimate of the treatment effect.

2.9.3 Testing parallel trends assumption

The *parallel trends* assumption is a key identifying assumption in DID estimation (Angrist and Pischke, 2008; Lechner et al., 2011) which states that in the absence of treatment, the difference in the dependent variable between the treatment and control group would be constant over time. For example, in the current context any change in recorded sales because of improvements due to surprise visits and audits, governance mechanisms improving naturally would affect both the control and treatment group FPSs similarly in the absence of BA devices.

We verify that the parallel trends assumption holds in our context using a formal test suitable for multi-valued treatment period (Pischke, 2005; Autor, 2003) and estimate the following model:

$$SGR_{i,t} = \alpha_i + \beta_t + \sum_{j=1}^{7} \gamma_j Pre\ period_{i,j,t} + \delta BA_{i,t} + \varepsilon_{i,t}, \tag{2.9}$$

where the treatment indicator is interacted with time dummies for pre-treatment periods. Specifically, $Pre\ period_{i,j,t}$ takes the value 1 if a BA device is installed j periods after t at FPS i, and 0 otherwise. As in the main equation (2), $BA_{i,t} = 1$ if a BA device was in use at FPS i in month t and 0 otherwise. The results of the above regression are shown in Column (1) of Table 2.6. All of the seven coefficients of $Pre\ period_{i,j,t}$ are insignificant, thus verifying that parallel trends assumption is satisfied for our model.

2.9.4 Other significant events during the study period

The concern here is that if any other significant event had occurred in Karnataka during our study period (Sep 2013 to Dec 2015) that may have also affected FPS sales or diversion, it would bias the estimate of the impact of BA intervention. Such events are likely to be those that drastically impact either the demand or the supply of food grains, thereby altering the incentives for FPS owners to divert and incentives for beneficiaries to access their entitlements and can be classified under: (i) natural disasters such as droughts and floods, (ii) price shocks, (iii) political events such as elections.

We conducted an extensive survey of secondary sources such as media reports and newspaper articles.[18] We found that Government makes official declaration of drought affected areas at the district level and accordingly relief funds are allocated to drought affected districts. Table 2.7 shows the list of districts included in our analysis that were declared as drought affected. (Districts which were drought affected but not part of our analysis are not shown in the table).[19] Since in our data, every district has control and treatment FPSs, we do not have a scenario where we have only treatment (or control) FPSs from a certain district which was declared as drought affected in a certain period.

Further, our survey of media and newspaper reports does not suggest the occurrence of any other significant event. The sixteenth loksabha elections were held in 2014. But again, all districts in our data set went into polls.[20] As for the assembly elections, the 14th assembly elections happened during 2013 (before our analysis period) and again in May 2018 (much after our analysis period ends) and the

[18] http://floodlist.com/tag/india/page/14, https://www.deccanchronicle.com/nation/current-affairs/121216/list-of-major-cyclones-that-have-hit-india-over-the-last-few-years.html

[19] http://wwfenvis.nic.in/files/DROUGHT\%20IN\%20KARNATAKA.pdf, https://sandrp.in/2016/05/07/karnataka-profile-of-2015-16-drought/

[20] https://en.wikipedia.org/wiki/2014_Indian_general_election_in_Karnataka

same state government (Indian national congress and CM Siddaramaiah) remained in power during our analysis period [21].

However, even if such an event (e.g., drought, floods, cyclones) had occurred it would have to affect that affected all FPSs, time fixed effects would absorb its effects.The only events we might need to worry about are such events which are both time and FPS varying (varying at the $\{i,t\}$ level), affect sales at FPS and correlated with our variable of interest ($BA_{i,t}$). Only then, will it induce endogeneity due to omitted variable bias, thus rendering the coefficient of BA installation biased (Angrist and Pischke, 2008). To the best of our knowledge, there are no other events that have occurred during our analysis period which is both $\{i,t\}$ varying and correlated with $BA_{i,t}$.

2.9.5 Falsification checks

We conduct two falsification tests (Gertler et al., 2016; Kennedy, 2003; Shoag et al., 2010; O'Neill et al., 2016) to ascertain that the estimated effect of BA installation is not idiosyncratic.

In the first test we exclude all the treatment observations, i.e., all observations with $BA_{i,t} = 1$ in the data. In the remaining observations, we randomly assign FPSs to control and treatment groups (based on a uniform distributed random variable and a decision rule that assigns an equal probability of getting assigned to treatment). We then assign a phantom intervention period (randomly selected) and estimate the DID model specified in equation (2.2) to estimate the average treatment effect for this pseudo treatment. (denoted as δ_{pseudo}). We repeat this procedure 100 times by drawing uniform random variables and assigning FPSs to control and treatment groups based on the realisation. We find that none of the 100 δ_{pseudo} coefficient values are significant (Column (1) in Table 2.8). The p-value of the 100 coefficients vary

[21] https://en.wikipedia.org/wiki/List_of_chief_ministers_of_Karnataka

between 0.11 and 0.93, with a mean value of 0.25.

In the second falsification test (Lechner et al., 2011), we retain all the observations as is, but change the intervention period for FPSs that had a BA device installed. We assign randomly chosen phantom intervention periods and estimate equation (2.2) to estimate the treatment effect for this pseudo treatment. We estimate the model for 12 such random intervention periods. We find that none of the δ_{pseudo} coefficient values are significant. The p-value of the 12 coefficients vary between 0.18 and 0.9, with a mean value of 0.45 (Column (2) in Table 2.8). The results of these falsification tests confirm that the treatment effect we find is not the result of pure chance.

2.9.6 Description of field visits

We undertook multiple field visits (around 10 visits comprising a total of about 60 hours spent in the field) - which involved interactions with beneficiary households, FPS owners and officials of food and civil supplies ministry (Secretary, Commissioner, Deputy Director, Zonal officers, Food inspectors). Field visits were made at different stages of our project and to complement our understanding from secondary sources - to understand the details of the intervention and workaround at the initial phase, to better interpret the data and variables which were scraped from publicly available sources during the data set construction and analysis phase, to communicate and share the results of our analysis.

Our understanding of the choice of FPSs for BA installation, workaround mechanism, suspicions about possible heterogenous effects with respect to ownership and location of FPSs, type of households etc. were informed from these field interviews. We interacted with different stakeholders (Secretary, Commissioner, Deputy director, Senior deputy director, Officials at the IT cell, Officials at the allocation cell, Food inspectors, FPS owners and Beneficiaries) from multiple districts (Bangalore Rural, Bangalore Urban, Tumakuru Urban, Tumakuru Rural) during our field visits in order to understand the integrated process flows across the PDS supply chain. The following is a brief summary of our interactions with each stakeholder group.

1. **Beneficiary households** - Our interaction with beneficiaries confirm the prevalent usage of the workaround mechanism. Few beneficiaries also reported receiving less grains (than their entitlement quantity) when workaround mechanism was used. Beneficiaries in some villages were particularly furious about this and also mentioned that such a practice was not uncommon. This means that the FPS owner can use the workaround mechanism to distribute lower quantity of grains, but record the entire entitlement as sales. Therefore, a significant portion of instances of use of workaround mechanism that were noted during our interaction with beneficiaries could be misuse of the mechanism by the FPS owner, which could be the reason behind the underwhelming impact of the BA devices. Some beneficiaries also spoke about them having to make multiple trips due to irregular FPS opening times. In few cases, beneficiaries mentioned that the FPS owners forcibly sells other retail items like oil, salt etc purchasing which is like an unwritten requisite in order to be able to get their grains.(Rozindar, 2016) These complaints reflect the control and power that the FPS owners wield over beneficiaries.

2. **FPS owners** - We visited shops with different combinations of location (urban,rural) and affiliation (private, cooperative). We witnessed the working of the BA device. The BA device had features such as display of the scale of issue (quantity) and price as well as announcement of the same in the local language. The FPS owners complained about the long and time consuming process involved in the repair and maintenance of the BA device in the event of a break down. Mechanics and engineers who were trained and had the expertise had to travel from nearby cities which increased the downtime of the device.

These indicate that the FPS owners seemed reluctant to use the BA devices and often resorted to the manual workaround mechanism.

3. **Top officials of food and civil supplies ministry** - Our interaction with the Secretary, Commissioner and the Deputy Director informed us about the criteria that the government used in choosing FPSs for BA installation. We defined proxy variables which represent these criteria in calculating the propensity score. We also got details about what variables in our data set relate to affiliation of FPS, household categories etc and what they represent.

4. **Officials in the planning and allocation division, IT cell and Zonal officers** - This interaction helped us acquire details about entitlement quantities and the fact that the entitlement quantity cannot be carried forward, but lapses every month. We were also told that the entitlement quantity for AAY category is mandated by the central government while the entitlement quantity for PHH category entitlement is defined by the state government. They also told us about the delay in processing of physical sales records at the FPS which leads to lagged information on ending inventory at the FPS which is used for calculating the replenishment quantity. We have described this in detail in the manuscript in Section 3.1 on page 6. This got us thinking about how the installation of BA devices also implies recording of real time sales data, and how this timely updated inventory information could be used for replenishment planning. This later developed into the simulation study where we quantify the value of better planning from using timely inventory information.

2.9.7 Alternate explanations and robustness checks

To confirm that the results in § 2.5 reflect the impact of BA devices on diversion, and are not driven by other reasons, we test for and rule out other plausible explanations.

We also perform a variety of additional analyses to test the robustness of our results to changes in methodology and variable definitions.

Anticipatory behavior of FPS owners.

If FPS owners had prior knowledge about the intervention, they may have preemptively diverted more than the usual quantity of food grains in the months immediately before the intervention and/or reduced diversion in the months immediately following the installation of the BA devices as a precaution. In such a case, our model specifications would overestimate the treatment effect. We check for this by estimating the model in (2.2) after excluding observations for two months immediately before and after the installation of BA devices. Column (1) in Table 2.9 shows that the magnitude of the treatment effect remains similar and statistically significant (0.45% of GR or 41.9 kgs per FPS per month).

Excluding FPSs without pre-intervention data.

Our dataset includes 1,767 FPSs that were selected early for the intervention and for which we do not have any pre-intervention data. It is plausible that these FPSs were selected based on performance, i.e., these FPSs may have high (low) levels of diversion before the installation of BA devices. In such cases including only post-intervention observations for these FPSs may overestimate (underestimate) the treatment effect. To test for this, we estimate the model in (2.2) after excluding all observations for these 1,767 FPSs. Results in Column (2) of Table 2.9 indicate that the average impact of the intervention remains similar and statistically significant (0.47% of GR or 43.8 kgs per FPS per month).

Alternate matching methods.

Recall that in the main model we use inverse probability weighting method to compare the control and treatment groups. We use two alternative matching techniques to check for robustness of our results. In the first, we use the "Nearest Neighbor" technique that selects a control FPS that is nearest in terms of propensity score for each treated FPS. In the second, we use "Caliper Matching" technique (Stuart, 2010) where all control FPSs within a predefined distance, i.e., caliper size, in terms of propensity score of each treated FPS are chosen. Columns (3) and (4) of Table 2.9 indicate that our results are robust to these alternative matching techniques. The change in the magnitude of the coefficient estimates and reduction in statistical significance are due to the difference in the composition and number of control FPSs used across the different matching techniques.

Alternate definition of the outcome variable.

We estimate a variant of our main model given by equation (2.2) using recorded sales ($S_{i,t}$) as the dependent variable and including the gross requirement ($GR_{i,t}$) as a regressor. Column (5) of Table 2.9 shows that the treatment effect estimated using this specification is very similar to earlier estimates and continues to be statistically significant (32.7 kgs vs. 36.34 kgs per FPS per month as per column (1) of Table 2.2) [22]. In addition, this specification provides a much better fit with data compared to the base model (Adj. R^2 of 0.999 vs. 0.128) as sales closely track gross requirement as seen from the summary statistics in Table 2.1.

[22] We also check for the validity of parallel trends assumption for this alternate specification with outcome variable as recorded sales. The results are shown in Column (2) of Table 2.6.

Impact of market price.

We estimate a variant of our main model given by equation (2.2) by including the market price variable to test if the magnitude of diversion may be (positively) correlated with market price of rice. We find that the main result on the impact of BA installation does not change with or without controlling for the market prices. We also find that the effect of market price is both statistically and economically insignificant. We believe that the main reason for the lack of effect is that the magnitude of geographic and inter-temporal variation in market price is substantially lesser compared to the difference between the average market price (≈INR 25 - 30 per kg) and the PDS commission for FPS owners (≈ INR 1 per kg). We provide details of this analysis below:

We hypothesize that the magnitude of diversion may be (positively) correlated with market price of rice. To elaborate, if the underlying driver for diversion (incentive for FPS owner) is the difference between market price of rice and the PDS commission for distribution of grains, one should expect that the magnitude of diversion will be greater for FPS-month combinations with higher market price compared to those with lower market price. It is worth noting that following three key assumptions underlie this argument: (i) storage of grains by FPS owners is difficult so that their diversion decisions are predominantly influenced by spot market prices, (ii) FPS owners are price takers, i.e., price in the open market is not influenced by the magnitude of diversion itself, and (iii) the PDS commission itself does not change substantially, so that the variation in market prices can be assumed to be reflected in the variation in the net incentive for FPS owners to divert. Based on the contextual understanding developed from our field visits, we believe that all three assumptions are reasonable in our study setting.

Data source and identification of proxy measures

We consider two measures of market price, at the wholesale and at the retail level. We obtain the wholesale price details for rice from the agricultural market information network portal (`http://agmarknet.gov.in`. Wholesale markets are sites of aggregation and assembly, dealing with the bulk purchases and trade of agricultural commodities, before they are processed in different units and distributed through a range of retail channels. The wholesale price data is recorded at the market - day level. Modal rice prices are recorded for each market at a daily level (Modal price here refers to the most common value of the price quoted by all the sellers who transacted at the market on that day). We take an average of the daily modal prices and arrive at the monthly wholesale price for every market.We obtain the retail price details for rice from the retail price information portal maintained by the Directorate of Economics and Statistics (`https://rpms.dacnet.nic.in/QueryReport.aspx`). Retail prices of agricultural commodities are collected from various market centres by different state and central government agencies and consolidated by the Directorate of Economics and Statistics. The retail price data is available at the market - month level. The markets in the above data sets are typically defined at the village or town levels. We match the village/town name of the wholesale and retail price datasets to the corresponding villages/towns of our analysis data set and populate the price of a market (village/town) against all the FPSs that belong to that village or town.

Unfortunately, these datasets have important limitations. First, they do not include data at the level of geographic granularity (FPS) that is relevant for our analysis. Second, they do not include data for all geographic units in our dataset. The data on wholesale price is only available for 1,379 villages across seven districts, which correspond to 85,482 FPS-month observations in our main dataset. Similarly, data on retail price is available only for two districts, which correspond to 25,885 FPS-month observations in our main dataset.

Figures 2.12 and 2.13 provide a pictorial representation of the variation in wholesale and retail prices. Figure 2.12 shows the trend of wholesale prices (shown as box plots) over time. Each box plot shows the variation of wholesale price across different markets for that month. Similarly, Figure 2.13 shows the trend of retail prices (shown as box plots) over time. Each box plot shows the variation of retail price across different markets for that month. The average wholesale price of rice for the markets in our data set during our analysis period is 23.29 INR per kg while the average retail price is 34.40 INR per kg. The FPS owners were paid a commission ($\approx$ INR 1 per kg) based on the quantity of grains distributed to eligible beneficiaries, which was substantially lower than the market price of food grains ($\approx$INR 25 - 30 per kg) (Correspondent, 2018). Thus, the gap between market price and the commission gained per kg from PDS is very high relative to the variation in the market price.

Model specifications and Results

Given that we do not have access to market price data at the FPS level for the period of our study, we define two proxy measures, wholesale price and retail price, respectively, and estimate the following equations:

$$SGR_{i,t} = \alpha_i + \beta_t + \delta\, BA_{i,t} + \gamma_w\, Wholesale\ price_{i,t} + \varepsilon_{i,t}, \tag{2.10}$$

$$SGR_{i,t} = \alpha_i + \beta_t + \delta\, BA_{i,t} + \gamma_r\, Retail\ price_{it} + \varepsilon_{i,t}, \tag{2.11}$$

$$SGR_{i,t} = \alpha_i + \beta_t + \delta\, BA_{i,t} + \gamma_w\, Wholesale\ price_{i,t} + \gamma_r\, Retail\ price_{i,t} + \varepsilon_{i,t}, \tag{2.12}$$

We present the results for the above three models in Table 2.10. We find that our main result on the impact of BA installation does not change with or without controlling for the market prices (wholesale or retail). We also find that the coefficient of both wholesale as well as retail price is both economically and statistically insignificant. As stated above, we reiterate that these results are likely driven by the fact that the variation in market prices (retail as well as wholesale) is much lesser in magnitude compared to the difference between mean market prices and the PDS commission.

2.9.8 Robustness to different models of heteregeneity

In § 2.6.2, we defined the notion of market access by a proxy indicator variable, $Farther_i$ which took values 0 or 1 based on the distance of the rural FPS i to its nearest urban market. Results in Table 2.11 defined a FPS as being close to a market if the distance to the nearest urban market was less than 20 km, the median distance to market in our data. We use alternate definitions, distances of 25 km (60^{th} percentile) and 31 km (75^{th} percentile), to define if a FPS is close to a market or not. Table 2.14 shows the results from estimating equation (2.6) with these alternate definitions, and indicates that our results are robust.

In §2.5.2, we estimated the effect of installing BA devices over time by considering a horizon of 10 months from the time of installation. We consider alternate time horizons, 5 and 15 months, to check for robustness of our results. Table 2.13 shows the results for the impact of BA devices for time horizons of 5, 10 and 15 months. We find that impact of BA devices persists in the long term when these alternate time horizons are considered.

2.9.9 Definition of proxy variables of vulnerability

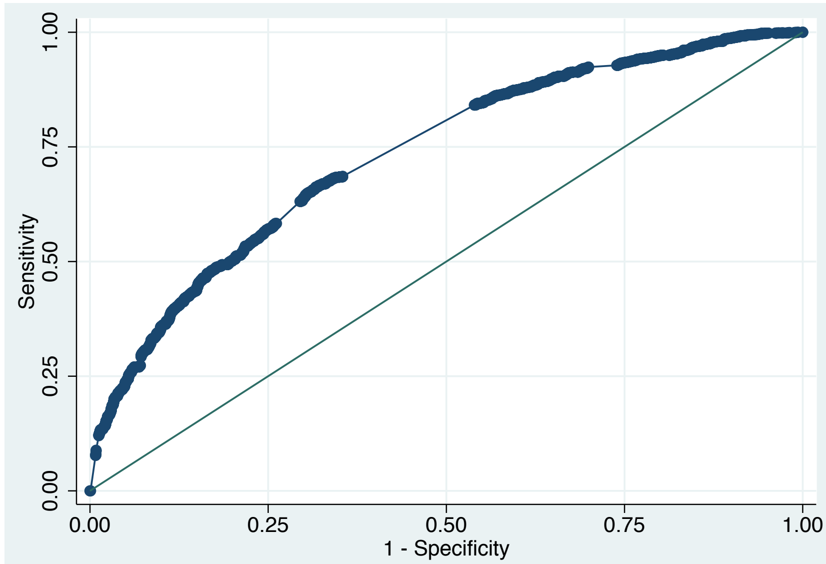
We use the following proxy variables of vulnerability and test for the behavioral mechanism:

1. Proportion of economically vulnerable households in the FPS - We calculate the average proportion of AAY cards of all months for each FPS. We

then use a threshold (90th percentile) to define the indicator variable, *FPS economic vulnerable group* that takes the value 1 if the FPS belongs to a group with proportion of AAY households greater than the defined threshold and 0 otherwise.

2. Proportion of socially vulnerable class population with respect to caste and tribe in the village where the FPS resides - We calculate the proportion of Scheduled caste (SC)and Scheduled tribe (ST) population for each FPS. (Chatterjee and Sheoran, 2007). We get this information at the village level from the Village and Town Amenities Dataset of the district census handbook, published by the Office of the Registrar General & Census Commissioner, India. (We have also used this data set to get proxy variables for the logistic regression used in calculating the propensity scores). We calculate the proportion of vulnerable class population for each village and populate the value against the FPS residing in that village. Again, we use the threshold of 90th percentile (0.45) to define the indicator variable, *FPS social vulnerable group* that takes the value 1 if the FPS belongs to a group with proportion of socially vulnerable group population greater than the defined threshold and 0 otherwise.

Figure 2.9: Receiver Operating Curve (ROC) of the propensity score model used to predict selection of FPS for BA device installation



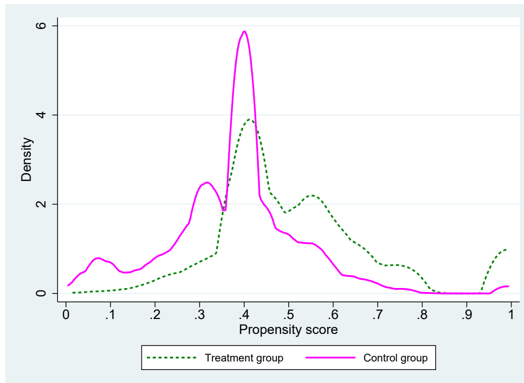
Notes: Details of the logistic regression model used to estimate the propensity scores are given in Appendix 2.9.2. Area under the Receiver Operating Curve shown is 0.73. Area under the curve (AUC) represents the discrimination ability of the prediction model, i.e., its ability to correctly classify treatment and control units. Numerically, it is equal to the probability that, for a randomly selected pair of FPSs (each pair has one from treatment and one from control group), the one from the treatment group has higher probability of being chosen in the treatment arm compared to the one from the control group. In other words, it is equal to the probability that a randomly selected pair of FPSs , one from the control arm and the other from treatment arm, are correctly classified. Thus, higher AUC denotes better discrimination ability of the model.

Table 2.4: Logistic regression for selection into treatment

Variable	**Coefficient**
Size and development	
Total geographical area in hectares	0.99***
	(2.78e-5)
Total population	1.00***
	(1.97e-07)
Availability of government engineering college	97.63***
	(30.96)
Availability of private engineering college	2.10*
	(0.97)
Availability of government polytechnic college	2.39***
	(0.74)
Availability of private polytechnic college	1.43
	(0.53)
Availability of public library	1.15**
	(0.08)
Technological connectivity	
Availability of telephone / landline	0.88
	(0.13)
Mobile phone coverage	1.67***
	(0.18)
Availability of power supply	0.32***
	(0.07)
Logistical connectivity	
Availability of public bus service	0.82*
	(0.08)
Availability of private bus service	1.71***
	(0.11)
Availability of auto rickshaws	1.55***
	(0.11)
Availability of taxis	0.56***
	(0.04)
Availability of state highways	0.74***
	(0.05)
Availability of major district road	0.99
	(0.08)
Availability of other district road	1.07
	(0.08)
Constant	2.14***
	(0.34)
L R Chi-square	1087.38
Prob >Chi-square	0.000
Pseudo R-squared	0.11
Observations	6,934

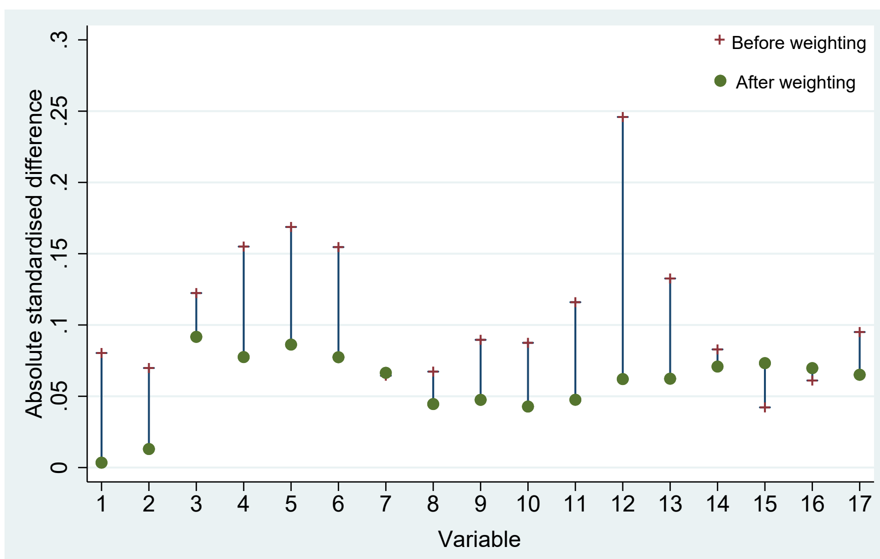
Notes: **1)** The results shown are from estimating equation (2.8) with robust standard errors. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 2.10: Distribution of propensity scores over treatment and control groups.



Notes: The common support is between propensity score of 0.0146 and 0.9903.

Figure 2.11: Standardized differences in pre-treatment characteristics used in the propensity score model



Notes: **1)** This figure shows the absolute standardized difference for each variable (variables 1 to 17 on the X-axis) used in the propensity score model, before and after using the inverse probability weights. **2)** Variables 1 to 7 relate to geographic size and the level of socioeconomic development. They include (1) Total geographical area in hectares, (2) Total population, (3) Availability of government engineering college, (4) Availability of private engineering college, (5) Availability of government polytechnic college, (6) Availability of private polytechnic college, and (7)Availability of public library. **3)** Variables 8 to 10 relate to technological connectivity and include (8) Availability of telephone / landline, (9) Availability of Mobile phone coverage, and (10)Availability of power supply. **4)** Variables 11 to 17 relate to logistical connectivity and include (11) Availability of public bus service, (12) Availability of private bus service, (13) Availability of auto rickshaws, (14) Availability of taxis, (15) Availability of state highways, (16) Availability of major district road, and (17) Availability of other district road.

Table 2.5: Estimation without Propensity score matching/weighting

	Without propensity score weighting	
	Including all observations	**Excluding observations outside the common support**
Overall impact	-0.0025*	-0.0025*
	(0.0013)	(0.0013)
R2	0.138	0.138
Adjusted R^2	0.0984	0.0983
Observations	152,958	152,334

Notes: **1)** Results shown are for FPS-month panel regressions with FPS and month fixed effects and standard errors clustered at the month and taluk levels. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** The dependent variable is $SGR_{i,t}$ for the all the models whose results are given above.

Table 2.6: Testing of parallel trends assumption

SGR ($SGR_{i,t}$)	**SGR** ($SGR_{i,t}$)	**Recorded sales** ($S_{i,t}$)
7 MBBI	0.0019	-0.144
	(0.0023)	(0.0954)
6 MBBI	0.0030	-0.0842
	(0.0031)	(0.1833)
5 MBBI	0.0026	0.0377
	(0.0032)	(0.1321)
4 MBBI	-0.0003	-0.1758
	(0.0027)	(0.1674)
3 MBBI	0.0009	0.1860
	(0.0017)	(0.1219)
2 MBBI	-0.0001	0.1902
	(0.0017)	(0.1931)
1 MBBI	-0.0012	0.1404
	(0.0024)	(0.1904)
$GR_{i,t}$		0.9955***
		(0.0022)
$BA_{i,t}$	-0.0037**	-0.277*
	(0.0017)	(0.1522)
R^2	0.167	0.99
Adjusted R^2	0.129	0.99
Observations	152,334	152,334

Notes: **1)** Results shown are for FPS-month panel regressions with FPS and month fixed effects and standard errors clustered at the month and taluk levels. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** The acronym MBBI in the variable column stands for 'Month(s) Before BA Installation'

Table 2.7: Drought affected districts in Karnataka

Districts which are part of our study	**Year 2013**	**Year 2014**	**Year 2015**
Bangalore Central	Yes	Yes	Yes
Bangalore East	Yes	Yes	Yes
Bangalore North	Yes	Yes	Yes
Bangalore South	Yes	Yes	Yes
Bangalore West	Yes	Yes	Yes
Belagavi	Yes	Yes	Yes
Chikkamagaluru	Yes	No	No
Dharwar	Yes	Yes	Yes
Kalaburagi	Yes	Yes	Yes
Mysuru	Yes	Yes	Yes
Tumakuru	Yes	Yes	Yes

Table 2.8: Results of falsification tests

	Random assignment of	
	FPSs to treatment	**Treatment period**
Mean of δ_{pseudo}	0.0001	-0.0025
Range of δ_{pseudo}	-0.0115 to 0.0111	-0.0113 to 0.0069
Mean of p values associated with δ_{pseudo}	0.2509	0.4502
Instances of δ_{pseudo} significant at $p < 0.1$	0	0

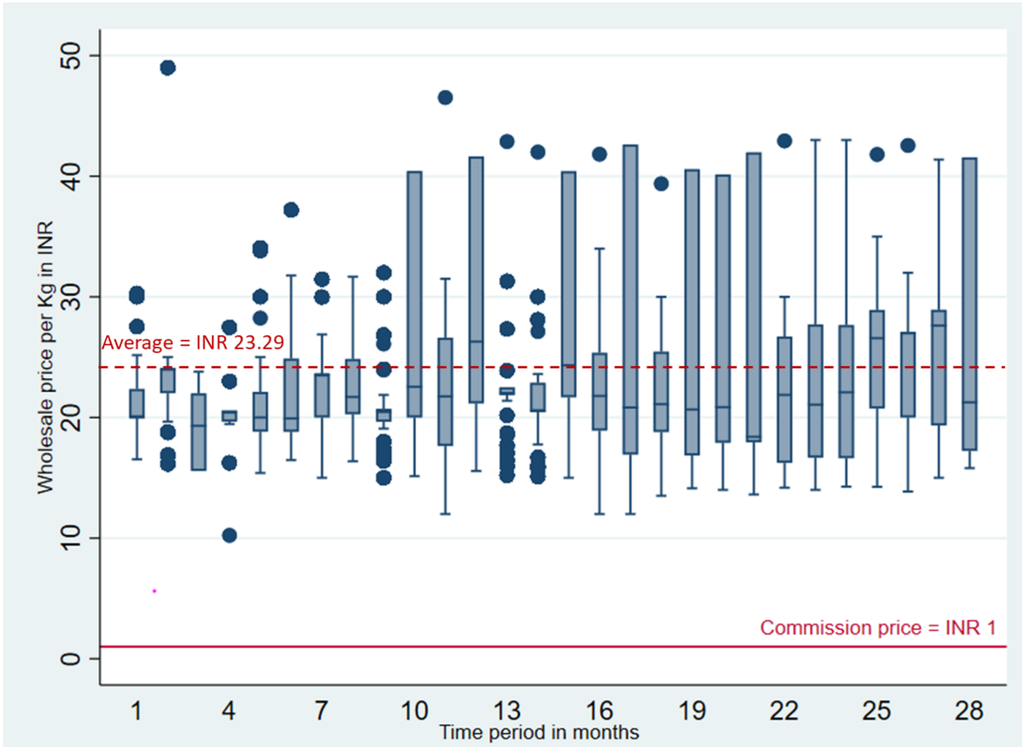
Notes: **1)** Column (1) represents the results over 100 simulation runs where FPSs were randomly assigned to treatment and control groups. **2)** Column (2) represents the results over 12 simulation runs with the treatment period being randomly assigned. **3)** The coefficient δ_{pseudo} was estimated using the DID model specified in equation (2.2). **4)** The estimated value of the original coefficient of BA_it (δ) is -0.0039^{**}.

Table 2.9: Testing for alternate explanations and other robustness checks

	Controlling for FPS owners' behavior	**Excluding FPSs with early BA installation**	**Control groups using nearest neighbor technique**	**Control group using caliper matching**	**Recorded sales ($S_{i,t}$) as dependent variable**
Overall impact (δ)	-0.0045**	-0.0047**	-0.0027*	-0.0023*	-32.7**
	(0.0027)	(0.0021)	(0.0013)	(0.0013)	(0.1430)
R^2	0.167	0.169	0.137	0.147	0.999
Adjusted R^2	0.127	0.130	0.097	0.107	0.999
Observations	146,146	112,743	133,780	98,593	152,334

Notes: **1)** Results shown are for FPS-month panel regressions with FPS and month fixed effects and standard errors clustered at the month and taluk levels. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** Column (1) shows the results from estimating equation (2.2) after excluding data for two periods immediately prior to, and after, BA installation for each FPS. **4)** Column (2) shows the results from estimating equation (2.2) after excluding FPSs that have no pre-intervention observations. **5)** Columns (3) and (4) show the results from estimating equation (2.2) when the nearest neighbor and caliper matching with caliper size of 0.1 are used to create the control groups, respectively. **6)** Column (5) shows the results from estimating equation (2.2) using $S_{i,t}$ as the dependent variable instead of $SGR_{i,t}$ and including $GR_{i,t}$ as a regressor.

Figure 2.12: Variation in wholesale market price with respect to time



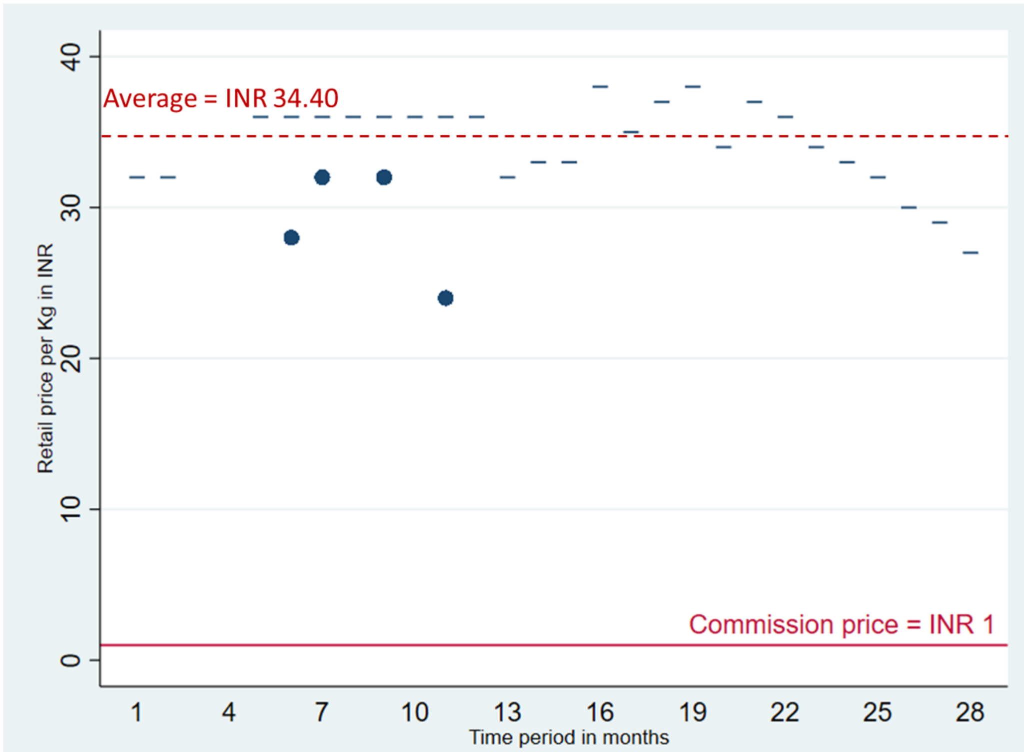
Notes: This figure shows the graph of box plots of wholesale price in different markets with respect to time (between the period September 2013 and December 2015). The red solid line represents the commission price (INR 1) and the red dotted line represent the average wholesale price for the entire dataset (INR 23.29).

Table 2.10: Model with market price proxies as covariates

	Wholesale price		**Retail price**		**Wholesale price & Retail price**	
	Without controlling	**After controlling**	**Without controlling**	**After controlling**	**Without controlling**	**After controlling**
Overall impact (δ)	-0.0058* (0.0032)	-0.0055 * (0.0030)	-0.0068 * (0.0029)	-0.0068 * (0.0029)	-0.0068 * (0.0029)	-0.0068 * (0.0029)
Wholesale price (γ)		-1.3e^{-06} (1.6e^{-06})				1.04e^{-05} (1.03e^{-05})
Retail price (γ)				-4.2e^{-06} (5.9e^{-06})		-1.03e^{-05} (9.47e^{-06})
R^2	0.178	0.178	0.130	0.130	0.130	0.130
Adjusted R^2	0.125	0.125	0.072	0.072	0.072	0.072
Observations	85,482	85,482	25,885	25,885	25,885	25,885

Notes: **1)** Results shown are for FPS-month panel regressions with FPS and month fixed effects and standard errors clustered at the month and taluk levels. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** Column (2) shows the results from estimating equation 2.10 with SGR_{it} as the dependent variable and including wholesale price as a covariate. **4)** Column (4) shows the results from estimating equation 2.11 with SGR_{it} as the dependent variable and including retail price as a covariate. **5)** Column (6) shows the results from estimating equation 2.12 with SGR_{it} as the dependent variable and including both wholesale price and retail price as covariates.

Figure 2.13: Variation in retail market price with respect to time



Notes: This figure shows the graph of box plots of retail price in different markets with respect to time (between the period September 2013 and December 2015). The red solid line represents the commission price (INR 1) and the red dotted line represent the average retail price for the entire dataset (INR 34.40).

Figure 2.14: Distribution of $\frac{S_{i,t}}{GR_{i,t}}$ for observations with $\frac{S_{i,t}}{GR_{i,t}} > 1$

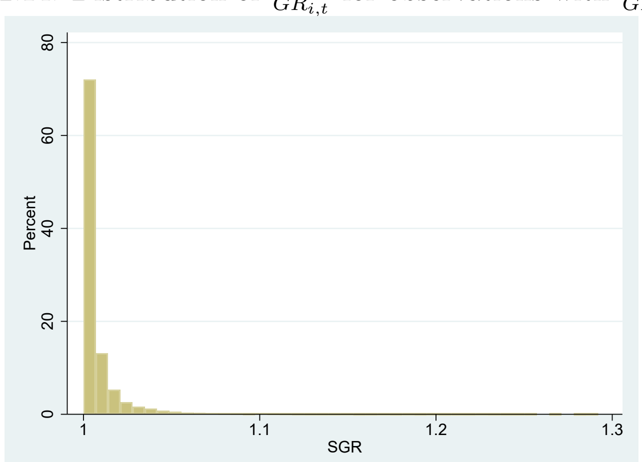


Table 2.11: Differential impact of installing BA devices on curbing diversion by FPS type

	By FPS location and ownership	**By FPS distance to nearest market**
Impact on urban cooperative FPSs (δ_{u-c})	-0.0052*	
	(0.0031)	
Incremental impact on urban private FPSs (δ_{u-pr})	0.0028*	
	(0.0016)	
Incremental impact on rural cooperative FPSs (δ_{r-c})	0.0014	
	(0.0036)	
Incremental impact on rural private FPSs (δ_{r-pr})	-0.0026	
	(0.0021)	
Impact on FPSs nearer to the market (δ_n)		-0.0033***
		(0.0011)
Incremental impact on FPSs farther from the market (δ_f)		0.0018*
		(0.0010)
R^2	0.167	0.303
Adjusted R^2	0.128	0.271
Observations	152,334	93,661

Notes: **1)** Results shown are for FPS-month panel regressions with FPS and month fixed effects and standard errors clustered at the month and taluk levels. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** Column (1) shows the results from estimating equation (2.5) with dependent variable $SGR_{i,t}$. The treatment effect on different types of FPSs are given by linear combinations of the coefficients shown in the table. We find that the impact is -0.0052* for Urban-Co-operative FPSs, -0.0024 for Urban-Private FPSs, -0.0038** for Rural-Co-operative FPSs and -0.0036*** for Rural-Private FPSs. **4)** Column (2) shows the results from estimating equation (2.6) with the dependent variable $SGR_{i,t}$. The impact on a FPS located farther from the open market is given by $\delta_n + \delta_f = -0.0015^{**}$.

Table 2.12: Differential impact of installing BA devices on curbing diversion by FPS group based on combination of different dimensions of vulnerability

	Base model	**Model based on combination of levels of different dimensions of vulnerability**
Impact on FPSs with low proportion of vulnerable households on both dimensions (δ_{ll})	-0.0041**	-0.0044*
	(0.0020))	(0.0022)
Incremental impact on FPSs with high proportion of economically vulnerable households (δ_{hev})		0.0019
		(0.0014)
Incremental impact on FPSs with high proportion of socially vulnerable households (δ_{hsv})		0.0009
		(0.0032)
Impact on FPSs with high proportion of vulnerable households on both dimensions (δ_{hh})		-0.0040
		(0.0038)
R^2	0.168	0.168
s Adjusted R^2	0.129	0.129
Observations	137,462	137,462

Notes: **1)** Results shown are for FPS-month panel regressions with FPS and month fixed effects and standard errors clustered at the month and taluk levels. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** Column(1) shows the results for the base model. **4)** Column (2) shows the results for estimating the model based on atleast one dimension of vulnerability. **5)** Column (3) shows the results for estimating the model based on vulnerability in both dimensions. **6)** The impact on FPSs with high proportion of economically vulnerable households ($\delta_{ll} + \delta_{hev}$) is -0.0024*, the impact on FPSs with high proportion of socially vulnerable households ($\delta_{ll} + \delta_{sev}$) is -0.0034* and the impact on FPSs with high proportion of vulnerable households on both dimensions ($\delta_{ll} + \delta_{hh}$) is -0.0003.

Table 2.13: Impact of BA installation on curbing diversion over time

	Definition of long term		
Variable	**> 5 months**	**> 10 months**	**> 15 months**
During the month of BA installation	-0.0026 (0.0026)	-0.0029 (0.0028)	-0.0033 (0.0030)
1 MPBI	-0.0029 (0.0023)	-0.0032 (0.0025)	-0.0036 (0.0027)
2 MPBI	-0.0054* (0.0029)	-0.0058* (0.0032)	-0.0063* (0.0034)
3 MPBI	-0.0047* (0.0025)	-0.0051* (0.0027)	-0.0057* (0.0030)
4 MPBI	-0.0061* (0.0026)	-0.0065** (0.0028)	-0.0071* (0.0031)
5 MPBI	-0.0056** (0.0023)	-0.0061** (0.0025)	-0.0067** (0.0027)
6 MPBI		-0.0048** (0.0023)	-0.0056** (0.0026)
7 MPBI		-0.0045** (0.0020)	-0.0054** (0.0022)
8 MPBI		-0.0027 (0.0020)	-0.0037* (0.0021)
9 MPBI		-0.0013 (0.0019)	-0.0024 (0.0019)
10 MPBI		-0.0013 (0.0016)	-0.0024 (0.0016)
11 MPBI			-0.0018 (0.0019)
12 MPBI			-0.0026 (0.0019)
13 MPBI			-0.0042* (0.0022)
14 MPBI			-0.0041* (0.0018)
15 MPBI			-0.0044** (0.0018)
Long term impact of BA installation	-0.0036* (0.0020)	-0.0049* (0.0027)	-0.0080* (0.0042)
R^2	0.167	0.167	0.167
Adjusted R^2	0.128	0.129	0.129
Observations	152,334	152,334	152,334

Notes: **1)** Results shown are for FPS-month panel regressions with FPS and month fixed effects and standard errors clustered at the month and taluk levels. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** The acronym MPBI in the variable column stands for 'Month(s) Post BA Installation'

Table 2.14: Robustness to alternate definitions of closer to market

	Distance to nearest market $\leq$ 25 kms	**Distance to nearest market $\leq$ 31 kms**
Impact on FPSs closer to the market (δ_n)	-0.0034***	-0.0031***
	(0.0011)	(0.0010)
Incremental impact on FPSs farther from the market (δ_f)	0.0028**	0.0018**
	(0.0012)	(0.0010)
R2	0.303	0.303
Adjusted R^2	0.271	0.271
Observations	93,661	93,661

Notes: **1)** Results shown are for FPS-month panel regressions with FPS and month fixed effects and standard errors clustered at the month and taluk levels. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** The dependent variable is $SGR_{i,t}$ for the all the models whose results are given above. Columns (1) and (2) show the results for the model in equation (2.6) for alternate definitions of $Farther_i$, as being greater than 25 kms and 31 kms respectively.

Table 2.15: Descriptive statistics of observations with $\frac{S_{i,t}}{GR_{i,t}} > 1$

Observations	**Mean**	**Std. Dev.**	**25th percentile**	**Median**	**75th percentile**
17,031	1.011	0.08	1.001	1.003	1.008

Table 2.16: Model variants based on different combinations of inclusion/exclusion criteria

Exclusion/inclusion criteria	**Statistic**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
	Coefficient	-0.0033**	-0.0035**	-0.0035**	-0.0037**	-0.0039**
Based on $\frac{S_{i,t}}{GR_{i,t}}$		(0.0013)	(0.0014)	(0.0015)	*(0.00152=)	(0.0017)
	Adjusted R^2	0.0542	0.1108	0.1264	0.1273	0.1288
	Observations	1,69,361	1,69,361	1,68,260	1,68,260	1,52,334
	Coefficient	-0.0029**	-0.0029**	-0.0028**	-0.0029**	-0.0028**
Based on $\frac{S_{i,j,t}}{GR_{i,j,t}}$		(0.0010)	(0.0010)	(0.0011)	(0.0011)	(0.0013)
	Adjusted R^2	0.0946	0.0946	0.2126	0.2183	0.2363
	Observations	160,332	160,332	157,546	157,546	142,225

Notes: **1)** Results shown are for FPS-month panel regressions with FPS and month fixed effects and standard errors clustered at the month and taluk levels. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** The dependent variable for all the models is $SGR_{i,t}$. The treatment of outliers is based on recorded sales and gross requirement at the FPS-month level for the results shown in the top panel, while it is based on recorded sales and gross requirement at the FPS-household category-month level for results shown in the bottom panel. **4)** Column (1) shows the results for the model with all observations with $GR_{i,t} > 0$ included and using $SGR_{i,t}$ values as is. Column (2) shows the results for the model with all observations included, and $SGR_{i,t}$ values truncated to one. Column (3) shows the results for the model after excluding outliers (observations with $SGR_{i,t}$ values greater than the 99th percentile) and using remaining $SGR_{i,t}$ values as is. Column (4) shows the results for the model after excluding outliers (observations with $SGR_{i,t}$ values greater than the 99th percentile) and remaining $SGR_{i,t}$ values truncated to one. Column (5) shows the results for the model after excluding all observations with $SGR_{i,t}$ greater than one (Main model in the paper).

Chapter 3

Availability and utilization of choice in food security programs: Analysis of an intervention from the Indian Public Distribution System (PDS)

3.1 Introduction

In many developing countries, government-managed food distribution program is a major policy instrument for achieving the Sustainable Development Goal of 'Zero hunger.' Approximately 92% of low-income and 73% of low- and middle-income countries operate some form of food distribution program (Gentilini et al., 2014). Despite substantial budgetary allocations (typically about 1% of the national GDP), many of these programs have not made satisfactory progress toward the 'Zero Hunger' goal (FAPDA, 2019). Most of these programs are plagued by leakage of grains to open markets, poor quality of grains, and apathetic customer service, driven in large part by the presence of monopolistic agents who are responsible for the last-mile delivery of food grains. These agents do not have strong incentives to improve efficiency and customer service but have strong incentives to divert subsidized grains to open markets (Banerjee et al., 2018; Pingali et al., 2017; Bank, 2003).

Governments have implemented several interventions in recent years to improve program efficiency. For instance, on the supply-side, Indonesia privatized the last-mile delivery of grains in its food security program (*Raskin*) and enabled the entry of new players through a competitive bidding process. While the entry of new players reduced operational cost, the prices paid by beneficiaries decreased only in regions with sufficient competition in the bidding (Banerjee et al., 2019). On the demand-side, Sri Lanka integrated its food security program with its national poverty alleviation program (*Samrudhi*) by replacing in-kind food transfers with cash transfers (Tilakaratna and Sooriyamudali, 2017). Similar designs of cash transfer have been piloted by other countries such as Bangladesh, Egypt and Ecuador (Gentilini, 2007; Gentilini and Omamo, 2011; Gentilini et al., 2014). Proponents of cash transfer argue that cash provides beneficiaries with the freedom to purchase *whatever* they want, *whenever* they want, and from *whomever* they want while simultaneously reducing the government's cost of program delivery (Bergolo and Galván, 2018; Del Ninno et al., 2007; Hidrobo et al., 2014; Lusk and Weaver, 2017; Morton, 2019). However, detractors argue that cash transfers may not be effective in eliminating hunger if: (i) beneficiaries face barriers in using cash for food purchases due to inaccessibility of markets, (ii) magnitude of cash transfer is not dynamically adjusted to match the volatility of food prices in the local markets, and (iii) beneficiaries willingly use cash for non-food purchases (e.g., alcohol or tobacco) (Harvey and Savage, 2006; Currie and Gahvari, 2008; Sabates-Wheeler and Devereux, 2010; Demirguc-Kunt and Klapper, 2012; Michelson et al., 2012; Khera, 2014; Pradhan et al., 2015; Lentz et al., 2016; Tilakaratna and Sooriyamudali, 2017; Pingali et al., 2019; Torkelson, 2020).

Recently, several state governments in India have implemented a novel intervention that leverages the ongoing digitization of India's food security program, the Public Distribution System (PDS). Under this intervention, termed *portability*, beneficiaries can use their biometric identities (fingerprints or iris scans) to digitally authenticate themselves and collect their food grain entitlements from any licensed shop within their state. This is in stark contrast to the traditional mode of operations, wherein beneficiaries could collect their entitlements only from a single, pre-assigned shop (hereafter ‘home shop’).[1] Thus, portability is designed to be a technology-driven solution to provide choice to beneficiaries and thereby decrease the monopolistic power wielded by the shop dealers.

However, the provision of choice is likely to make an actual impact only if beneficiaries have easy access to viable alternatives. In the context of PDS, where beneficiaries incur substantial (direct or indirect) cost of hauling large quantity of food grains (20-35 kgs) from the shop to their residence each month, access to viable alternatives depends on the density of shops, i.e., the number of shops within reasonable distance, in the beneficiaries’ neighborhood. Thus, beneficiaries with higher density of shops in their neighborhoods are likely to incur lower incremental cost of accessing alternative shops and, hence, are more likely to utilize portability.[2] Even when alternatives are available in close proximity, shop dealers may collude to not serve each other’s beneficiaries, thereby reducing the beneficiaries’ ability to utilize portability (Sharma and Gupta, 2019). Given that cartels with a smaller number of players are stronger and more likely to sustain for a longer period (Hamaguchi et al., 2009), the problem of collusion might be more acute in regions with lower shop density. We therefore believe that although implementation of portability is enabled by technology, its utilization by beneficiaries depends on physical characteristics of the PDS network, specifically, the density of shops in the beneficiaries’ neighborhood.

Our aim in this paper is to understand the relation between shop density (number of shops within 0.5 km) and efficacy of providing choice through portability. However, efficacy of portability—in terms of improving beneficiary welfare—is not captured in the program data available to us. Since utilization of an IT intervention in a public program has been shown to be a good proxy for its efficacy (Heeks, 2001), we quantify

[1]Food entitlements for a month are pre-determined by the government and typically include, but are not limited to, commodities such as rice, wheat, pulses and sugar.

[2]The heterogeneity in the density of shops within a reasonable distance is likely to arise due to the administrative rule of thumb of allocating one shop per 1000 beneficiaries (Allu et al., 2019) combined with heterogeneity in the population density.

the relation between shop density and three inter-related aspects of beneficiaries' utilization of portability: (i) whether portability is utilized in a given month, (ii) number of unique shops where portability is utilized,[3] and (iii) frequency of utilization of portability in a given time period. The second and third measures, in addition to whether portability is utilized in a given month, helps us to better understand potential underlying drivers of utilization. For instance, a beneficiary using portability occasionally might be indicative of short-term inconveniences at the home shop such as internet or electricity shut down whereas continued utilization of portability might be indicative of aspects such as incompatibility with the shop owner due to caste, religion or other social identity.

Our empirical context is the Indian State of Andhra Pradesh, which was one of the earliest ones to introduce portability in 2015. We use large-scale program data comprising more than 75 million transactions by 13.92 million beneficiaries at over 28,000 shops from March 2018 to August 2018. We find that on average 18% of beneficiary households utilized portability in any month. Among those who utilized portability at least once, the median number of alternate shops (a shop other than the home shop) and the median number of months in which portability was utilized are 1 and 4, respectively. Using a logistic regression model, we find that one additional shop within 0.5 km radius increases the odds of using portability by 6.8%. We also find that shop density is the most important determinant of portability utilization; it explains 68% of the variation as against socioeconomic household characteristics which together explain 22% of the variation. Using two separate Poisson regressions, we find that, among beneficiaries who utilize portability, those in regions with greater shop density use more shops. However, we do not find any association between shop density and the frequency of utilizing portability, i.e., number of months when portability was utilized by a household.

[3] We find that over 99.9% of the beneficiaries use only one shop in a month. Hence we analyze the total number of shops used in our entire analysis time period.

Our results contribute new evidence to the relatively sparse discussion on utilization of portability in PDS in India. Two previous studies on utilization of portability in the urban areas of the state of Chhattisgarh (Rajan et al., 2016; Joshi et al., 2016) find mixed evidence. Rajan et al. (2016) find that beneficiaries prefer to use shops with better service quality and road connectivity whereas Joshi et al. (2016) report that the fraction of beneficiaries utilizing portability falls to almost zero within 18 months after the launch of the intervention. Neither of these studies quantify the association between number of alternatives (shop density) and utilization of portability. Furthermore, these studies capture initial experiences with portability, which may be transitory and may differ from the steady state performance of the program. Our study period (March 2018 to August 2018) is nearly three years after the introduction of portability in Andhra Pradesh and, hence, provides a more accurate reflection of the long run utilization of portability and its key drivers. Consequently, these findings are more appropriate for evaluating technology-enabled program interventions, which involve significant investments and beneficiary welfare depends on long-term behavior.

Further, our work also contributes to the existing literature on technology-enabled interventions in public programs (Chowdhury and Koya, 2017; Banerjee et al., 2014; Elbahnasawy, 2014; Ganesh et al., 2019; Gössling and Michael Hall, 2019; Muralidharan et al., 2016; Naik et al., 2020; Tjoa and Tjoa, 2016). Several studies highlight that technology interventions in public programs are likely to fail, leading to huge costs without commensurate benefits, if their design does not adequately account for existing preconditions of the program (Heeks, 2001; Masiero, 2016; OFT, 2010; Ray and Mukherjee, 2007; Saxena, 2005). For instance, within the specific context of PDS, Masiero (2015) finds that digitizing transactions between beneficiaries and shops in PDS alone does not address the issue of grain leakage into the open market as a large portion of leakage occurs upstream in the supply chain. Similarly, Allu et al. (2019)

argue that digitizing the process of applying for identity card in PDS may not be effective in reducing the time taken to process the application if document verification continues to be a manual process. Most studies in this literature employ qualitative methods such as semi-structured interviews (Ekirapa-Kiracho et al., 2011; Heeks, 2002; Kaushik and Singh, 2004). We complement this discussion with a quantitative analysis of the role of preexisting program characteristics (e.g., shop density) on the utilization of the technology-enabled program intervention (e.g., portability).

Finally, our findings contribute to the sizeable literature on choice-based interventions in public programs (Clarke et al., 2008; Fischer et al., 2006; Fotaki and Boyd, 2005; Le Grand, 2009) due to two key contextual differences. First, unlike previous studies in this literature, providing choice to beneficiaries induces competition among pre-identified government agents without introduction of new private players. Second, commonly used levers of competition in those contexts such as product differentiation and service quality (e.g., insurance, pensions and healthcare) are not applicable in the PDS context. Consequently, the decision-making process for beneficiaries is likely to be more heavily influenced by transaction costs as difference in product or service features across shops is minimal.

The remainder of the article is structured as follows. We provide an overview of the PDS in §3.2. In §3.3, we describe our data, present measures of our outcome variables and various factors that may be associated with them. §3.4 provides descriptive statistics on spatiotemporal utilization of portability, shop density and other controls used in our analysis. We analyze the influence of shop density on our outcome variables in §3.5 and provide relevant extension and robustness checks in §3.6. We conclude by discussing policy implications of our results and streams of further inquiry in §3.7.

3.2 Background: Indian Public Distribution System

India's Public Distribution System (PDS) is one of the largest food security programs in the world. In 2018, India spent roughly 1% of the national GDP (INR 1.15 Trillion) to provide subsidized food grains to around 160 million households through government licensed outlets called Fair Price Shops (FPSs). Licenses to manage shops are issued to private dealers or cooperative societies for a fixed period of 3 years. The shop dealers are paid a commission of about INR 0.70 per kg of grains distributed to beneficiaries.[4] Each eligible household is entitled to receive a fixed quantity of food grains every month at heavily subsidized prices (INR 1 per kg compared to market prices of INR 28 - INR 40 per kg).[5] The magnitude of entitlement varies based on the economic status of a household which is identified either as AAY (Antyodaya Anna Yojana) or PHH (priority household), with the former being the poorest of the poor. AAY households receive an entitlement of 35 kgs per household irrespective of the number of individuals in the household while PHH households receive an entitlement of 5 kgs per person per household.

Traditionally, beneficiaries were allotted a specific shop from which to collect their entitlements. Every shop dealer received a paper-based roster of beneficiaries allotted to her shop and issued grains to beneficiaries only after verifying their names on a government issued identity card against the list of allotted beneficiaries. This system accorded monopoly power to the shop dealers over beneficiaries, which in turn led to inefficient and poor quality of service manifesting in terms of frequent shop closures, mistreatment of beneficiaries, long queues, adulteration of grains, overcharging / underweighting (Khera, 2011; Vaidya et al., 2014; Sargar et al., 2014; Dreze and Khera, 2015; Sati, 2015; Sharma and Gupta, 2019). Despite the presence

[4] Based on the new article - https://www.thehindu.com/news/cities/Hyderabad/ration-dealers-will-be-paid-increased-commission-eatala/article23315122.ece.

[5] This data is obtained from the Agmarknet portal - https://agmarknet.gov.in/.

of grievance redressal mechanisms and vigilance committees, only an estimated 1.5% of the beneficiaries across the country were aware of them (NCAER, 2015).

Beginning in 2010, central and several state governments have embarked on an ambitious plan of end-to-end digitization of the PDS. A prominent feature of this initiative is the use of digital authentication of beneficiaries using smart cards or biometrics at fair price shops using electronic devices that are linked to central servers Allu et al. (2019). As of September 2019, around 10 states have started to leverage this feature to allow any member of a beneficiary household to authenticate their identity and collect entitlements at any shop in the state (Ali, 2018; Today, 2019). It is expected that portability of benefits will provide convenience to the beneficiaries and cut down the monopoly power of the shop dealers. In addition, although not our main interest, it is expected to help migrant workers access their food entitlements when they are away from home for work, as long as they are within their home state (Ali, 2018; Hindu, 2018, 2019). Currently, the central government is taking measures to implement this functionality across the country under the 'One Nation One Ration Card' program (Today, 2019).

3.3 Data and Measures

3.3.1 Data

Our study is based in Andhra Pradesh, which was the first state to introduce state-wide portability in 2015. We collected publicly available program data for a period of six months (March 2018 to August 2018) from a state government website operated by The Department of Consumer Affairs, Food and Civil Supplies (`https://aepos.ap.gov.in/ePos/`). It comprises 75.57 million transactions made by 13.92 million households at 29,212 shops spread over 13 districts and 664 sub-districts. The data is organized in three parts—beneficiary dataset, FPS dataset and transaction dataset—which are described below.

Beneficiary Dataset

The *beneficiary dataset* contains the following information on each beneficiary household: a unique identification number, identification code of the FPS that it was originally allocated to (hereafter referred to as *home shop*), district/sub-district of residence, gender and name of the head of the household, and its category defined as per the National Food Security Act—PHH or AAY.[6]

Fair Price Shop (FPS) Dataset

The *FPS dataset* contains the following information for each shop: a unique identification number, its geographic coordinates (latitude and longitude), address and dealer's name.[7]

We classify the dealership of an FPS as a cooperative society if the dealer name contains key words such as Self-Help Group (SHG), Co-operative, Society.[8] The remaining are classified as private for whom we predict the gender of the dealer from their name using Naive Bayes Classifier algorithm (Langley et al., 1992; Friedman et al., 1997) trained using the name and gender of household members from the beneficiary dataset.[9]

[6] In most states, the priority households are entitled up to 5 kg of rice per person per month at the issue prices of INR 1. The AAY households can claim up to 35 kg of food grains per household per month at the same price.

[7] The latitude and longitude data could not be identified for 7K shops. For these shops, we use the FPS address to identify the most granular geographic location (village/colony/street number) and populate the corresponding coordinates extracted using Google API. Further, we test the accuracy of the extracted co-ordinates by triangulating them with village level coordinates populated in national village census data 2011.

[8] The exhaustive list of key words searched to identity FPS not manged by private dealers is obtained by a combination of substring analysis in SAS and manual inspection. The list includes various combinations of the key words mentioned below – SHG, Co-op, MSS, Society, PACS, VRA, GPMC, Mahila, Sangham, Group, DWARCA, Podupu, Cooperative.

[9] We test the algorithm's accuracy by administering it on a sample data drawn from the household dataset, for which the gender is already known. Gender predicted by the algorithm matches with the actual gender in 96% of the cases.

We categorize the location of a FPS (urban/town/rural) by extracting the specific village or city name from its address and mapping it to the village amenities, towns amenities and urban agglomerations datasets of the 2011 Census of India.[10] [11]

Transaction Dataset

Each transaction in the *transaction dataset* is identified by a unique transaction ID and contains the beneficiary household and shop unique identification codes, date of the transaction, and quantities of each commodity purchased.

3.3.2 Measures

Outcomes

As mentioned in Section 3.1, we consider three outcome variables, whether portability was utilized in a given month, number of shops used for portability transactions, and the number of months during which portability was utilized by a household. We capture the utilization of portability by household i (hereafter we use the terms beneficiaries, households, and beneficiary households interchangeably) in month t using an indicator variable, $Portability\ Usage_{it}$. It takes a value 1 if household i in a given month t transacts at a shop different from its home shop, and 0 otherwise. We define '$FPS\ Count_i$' as the number of unique shops other than the home shop (hereafter 'alternate shops') where at least one transaction was made by household i during the study period. Similarly, we define '$Month\ Count_i$' as the number of months in which at least one transaction was made by household i at an alternate

[10] Our search is based on a fuzzy match using Levenstein distance. If the village/city name extracted from the FPS address is more than 80% similar to the ones populated in census, we consider it a match. In cases where the FPS city/village name finds more than one match in census data, we choose the census village/city name with the highest percentage match.

[11] Census is an enumeration exercise carried out by the Ministry of Home Affairs, Government of India. This exercise is carried out once in every 10 years. This exercise was last conducted in the year 2011 and the data generated from it is shared in the link below – `http://censusindia.gov.in/2011-Common/CensusData2011.html`

shop.

Determinants

Availability of Choice - Our main predictor is the availability of alternate shops for each household, which we capture using an integer variable, '$FPS\ density\ (x\ km_i$'. This variable denotes the number of shops within x km from household i's home shop (including the home shop).[12] The home shop of a household is typically the one closest to it and the additional (direct and indirect) cost of using an alternate shop is likely to be lower when there are more shops in close proximity of the home shop. Further, given that shops are not likely to be very different from each other in terms of commodities sold, households may be indifferent to using a specific alternate shop and might utilize the functionality more frequently when more shops are present. We therefore hypothesize that all else being equal, odds of utilizing portability, the number of shops and the number of months in which portability is utilized by a household will be positively associated with the number of alternative shops (*FPS density*) in its proximity.

Other key determinants - Besides shop density, we posit that a household's utilization of portability, number of shops used and number of months in which portability is utilized are influenced by two categories of factors: (i) those related to the quality of the service and the operation of the home shop, which we call FPS characteristics (FPSC), and (ii) those related to the sociodemographic characteristics of the households themselves, which we term household characteristics (HHC).

FPS Characteristics - We control for two FPS characteristics relevant to a household's utilization of portability: the number of days for which the home shop is open in a month, and the type of dealership of the home shop. FPSs in Andhra Pradesh are expected to be open during the first 15 days of the month. Several studies

[12] We consider different values of x (0.5 Km, 1 Km, and 2 Km), which we calculate as the Haversine distance based on the latitude and longitude of the shops from the FPS dataset.

report that one of the major concerns among households is that their home shop is not open for the stipulated duration (Sharma and Gupta, 2019; Vaidya et al., 2014). Households whose home shops are closed more often may have to make open market purchases more frequently or make multiple visits to their home shop, both of which entail significant additional costs. We therefore hypothesize that households whose home shops are open for a fewer number of days in a given month would be more likely to use alternate shops. We define '$FPS\ open\ days_{it}$' as the number of days for which we observe at least one transaction associated with the home shop of household i by any beneficiary household in month t.[13]

Mistreatment and gender-based harassment and discrimination by shop dealers is known to be a major concern among beneficiary households (Vaidya et al., 2014). Given that women take the primary responsibility of collecting grains in most households (Sharma and Gupta, 2019; Pradhan and Rao, 2018), we hypothesize that households with home shops managed by men are more likely to utilize portability.[14] Further, several studies highlight that shops operated by cooperatives/self-help groups are likely to be more beneficiary centric than those managed by private agents (Desai and Olofsgård, 2019; Kumar et al., 2019; Nagavarapu and Sekhri, 2016). To capture this, we include a categorical variable, '$dealership_i$', which takes one of three values—co-operative managed, privately managed by a male or privately managed by a female dealer.

Household Characteristics - A household's economic, physiological, geographical and social characteristics are likely to affect its need for, and actual utilization of portability. Economically weaker households, controlling for other characteristics,

[13]There is a likelihood that a shop is kept open but there are no transactions registered. However, based on our semi-structured interviews, such likelihood is very less. We gathered that most FPS dealers have secondary occupations such as farming and small-scale businesses. They typically open the FPS at the start of the month, keep it open for a consecutive streak of days and move on to other occupations during the rest of the month. Therefore, it is reasonable to believe that our definition, although not perfect, is a close proxy to the actual number of days a shop is kept open.

[14]Our data does not capture the details of the individual within a household who transacts at the FPS.

are more likely to need portability because of their higher dependence on PDS for food security. However, they are also less likely to be able to afford the additional costs associated with using portability, especially if the alternate shops are farther than the home shop. Thus, the net impact of economic vulnerability on utilization of portability is unclear. We use the household's PDS category to capture economic vulnerability and define a binary variable 'AAY_i' that takes a value 1 if it belongs to the AAY category, and 0 if it belongs to PHH category.

A household consisting of elderly members is less likely to utilize portability because of their inability to physically travel and carry grains over longer distances, especially if mechanized means of transport are not readily available or expensive. To capture this physiological vulnerability, we construct an indicator variable, '$Elderly_i$', which takes the value of 1 if all members of the household are above the age of 60 years and 0 if at least one member is below the age of 60 years.

Households belonging to socially vulnerable groups such as Scheduled Castes (SC), Scheduled Tribes (ST) and other Primitive Tribal Groups (PTG) face discrimination in accessing government schemes including PDS entitlements (Sahas, 2009; Newman and Thorat, 2010; Sabharwal, 2011; Nagavarapu and Sekhri, 2016; Pradhan and Rao, 2018). These households are typically not well versed with digital media and hence are less likely to access information posted by the government that may be useful for choosing an alternate shop, e.g., stock availability, opening times, and addresses of shops in their proximity (Kumar and Best, 2006; Ali and Kumar, 2011; OFT, 2010).[15] We do not have access to data on caste at the household level. Instead, we use Socio-Economic Caste Census (SECC) data published in the year 2014 to define socially vulnerable sub-districts as those with proportion of population belonging to socially vulnerable groups above the 75^{th} percentile across all sub-districts in the

[15]The governments of most states share this information on their webpages. The source data used in this analysis is also scraped from one such webpage. In addition, in some states such as Chhattisgarh, we also observe such information being shared with beneficiaries through SMSs on mobile phone.

state.[16] We then define a binary variable '$SC/ST/PTG_i$' which takes a value 1 if household i resides in a socially vulnerable sub-district and 0 otherwise.

Households in urban areas are more likely to have easier access to cheaper modes of transportation compared to those in smaller towns and rural locations and therefore more likely to utilize portability. Since we do not observe the address of the beneficiary household, we use the location of the household's home shop as a proxy to define a categorical variable, '$Location_i$' which takes one of the following three values: urban, semi-urban, or rural, if the location of a FPS is urban, town or rural, as described in section 3.3.1.

3.4 Descriptive Statistics

In this section, we provide descriptive statistics on the portability utilization outcome measures and their determinants.

3.4.1 Portability Utilization

We find that 27.5% beneficiary households (about 3.8 million) made at least one portability transaction during our study period of six months (henceforth "portability users"). However, in any given month, only around 18% of the households utilized portability in any given month on average.

Table 3.1 shows the break-up of the number of months in which the portability users collected their entitlements from an alternate shop. Almost 35% (of the 3.8 million households who utilize portability) utilized it in all 6 months whereas just under 19% of them utilized portability only in one month. In Table 3.2, we show a cross-tabulation of the number of months in which portability users collected

[16] Identification of sub-district as vulnerable is based on the percentage of its households belonging to primitive tribal groups, SC and ST categories. We compute these percentages across all 664 sub-districts in the state. We categorise a sub-district as vulnerable if the percentage of households belonging to any of these categories is higher than the 75^{th} percentile. The 75^{th} percentile for primitive tribal groups, SC and ST are 2%, 30% and 31% respectively.

their entitlements from the PDS and the number of months in which they utilized portability. More than half (52.6%) of the portability users used an alternate shop in each month that they collected their PDS entitlements (sum of the diagonal elements in Table 3.2).

Next, in Figure 3.1, we show the distribution of distance between the home shop and the transacted shop for all portability users, which provides an understanding of the primary drivers for utilizing portability. Close to 50% of portability transactions occurred at a shop within 1 km and about 75% of portability transactions occurred within 7 km from the home shop of the household. This suggests that most beneficiaries utilize portability for reasons related to service quality/convenience and therefore availability of alternatives in the vicinity of a household's residence is likely to be an important determinant of portability utilization. In contrast, only 3.4% of transactions occurred at an alternate shop greater than 90km from the home shop, which could be indicative of portability utilization due to longer-term changes in the household characteristics, e.g., migration.

Table 3.3 shows the distribution of portability users by the number of alternate shops (shops other than their home shop) used during the study period. Almost three-quarters (73%) of households used only one alternate shop for their portability transactions. This could indicate a strong preference for a specific alternate shop that best suits their needs or a lack of availability of an effective choice due to reasons such as prevailing caste dynamics or collusion among shop dealers.[17]

3.4.2 Determinants

Table 3.4 shows the descriptive statistics for shop density and other control variables used in our analysis. On average, each household had about 3 other shops within 0.5 km from its home shop whereas 99% of households had less than 10 shops within 0.5

[17] In our field visits, we observed that SC/ST households were reluctant to use a shop located in a street/neighborhood dominated by upper castes households.

km radius from the home shop. We find that shops were open for an average of 12.23 days per month compared to the specified operational guideline of 15 days. Home shop of 5.24% households was managed by cooperatives or self-help groups, those of 48.29% households was managed by private male dealers whereas that of the rest was managed by private female dealers. Of all the households, 79.27% were located in rural areas, nearly 12% of them were in sub-districts with large SC/ST/PTG population, 6.39% had only elderly beneficiaries, and 6.24% households belonged to AAY category.

3.5 Model specification and main results

In this section, we quantify the association of outcome variables of interest, described in Section 3.4.1, with the main predictor variable, shop density. For computational ease, we draw a simple random sample of 500,000 households from the population of 13.92 MN in the beneficiary dataset and obtain their monthly transaction activity for 6 months from the transaction dataset. To assess the representativeness of our sample, we check the difference between our sample and the population on all predictors described in Section 4 and do not find any statistically significant difference.[18]

We exclude around 2.7% of the observations in the sample (64,899 unique household-month combinations) that belong to households whose home shop did not register any transactions in that calendar month, i.e., $FPS\ open\ days_{it} = 0$. These observations are likely to correspond to shop closures due to administrative reasons such as shop dealership changes and audits and hence utilization of portability in these observations is unlikely to be driven by factors of our primary interest, which are discussed in Section 3.3.[19]

[18] Table 3.9 in the appendix shows the comparison of descriptive statistics of all predictors between the sample and the population. We also repeat our analysis on other independently drawn random samples and all our results continue to hold.

[19] We check the robustness of our model by estimating our model for the entire sample. These estimation results are shown in Table 3.10 in appendix section. Coefficients on variables of interest

3.5.1 Odds of portability Utilization

We estimate the following logistic regression model to examine the association between households' odds of using portability and various PDS and household characteristics:

$$ln(\frac{p_{it}}{1-p_{it}}) = \alpha_d + \beta\ (FPS\ density)_i + \gamma_1\ (FPSC)_{it} + \gamma_2\ (HHC)_i + \epsilon_{it} \qquad (3.1)$$

where $p_{it} = Prob\ (Portability\ Usage_{it} = 1)$ is the probability of household i using an alternate shop in month t. The vector β is the set of coefficients of interest, corresponding to three variables – FPS density (0.5 km), FPS density (0.5km - 1.0 km), FPS density (1.0km - 2.0 km). The vectors γ_1 and γ_2 are the coefficients that capture the association of portability utilization to FPS characteristics (**FPSC**) and household characteristics (**HHC**), respectively, and α_d represents district fixed effects, i.e., factors that are common to all households in a district such as those related to private markets and government administration. We estimate clustered robust standard errors with two-way clustering at the sub-district and month level. [20]

The marginal effect of the various determinants on the odds ratio $\frac{p_{it}}{1-p_{it}}$ of utilization of portability are shown in Column (1) of Table 3.5. We find that, in accordance with our initial hypothesis, portability utilization increases with the number of alternate shops available for transaction in the vicinity of the household. Every additional shop available within 0.5 km radius from the household's home shop is associated with increase in the odds of utilizing portability by 6.8%. As expected, this impact is lower for alternate shops that are farther away: 1.9% for

do not change significantly.

[20]Error terms are likely to be correlated within time period (months) across sub-districts . The correlation could arise because of unobserved household behaviour and characteristics that are persistent over time. Further, Households in the same sub district are likely to be similarly impacted by factors such as change in administration, focus on PDS related developmental initiatives. As a robustness check, we also estimate our model by clustering the errors at a district level and our results continue to hold.

every additional shop within 0.5 km to 1 km and 0.9% for every additional shop within 1km to 2km.

We also find that a household's odds of utilizing portability decreases by 6.7% with every additional day its home shop is open. This suggests that beneficiaries utilize portability to overcome lack of availability of the home shop. It is important to note that utilization of portability was significantly lower among vulnerable populations (e.g., poor, socially disadvantaged, elderly, rural). The odds of utilization of portability by AAY households are 29.2% lower compared to PHH households, while the odds of utilization of portability by households residing in socially backward regions, i.e., villages that have a larger proportion of SC, ST, and Primitive Tribal Groups are 16.3% lower compared to households in other regions. Similarly, the odds of utilization of portability by households consisting entirely of elderly beneficiaries are 24.2% lower and the odds of utilization of portability by households in urban and semi-urban areas compared to those in rural areas are higher by 37.2% and 24%, respectively. This supports the argument that non-technology features of the program as well as social and economic barriers faced by households can make technology-driven enhancements in welfare programs less accessible to the more vulnerable households.[21]

Further, we use the results from estimation of Equation 3.1 to compute the likelihood of utilizing portability (p_{it}) at different values of *FPS density (0.5 km)* for a typical household in AP which takes average values of all predictors described in 3.4.2 (Cameron & Trivedi, 2009). Figure 3.2 shows the likelihood of utilizing portability as a function of *FPS density (0.5 km)* and suggests that utilization of probability increases almost linearly with *FPS density (0.5 km).* A typical household's probability of using portability is 11% when there are no alternate shops within 0.5 km radius *(FPS density (0.5 km) = 1)* and it nearly doubles to 20% when there

[21] From estimation results of Equation (1), we also calculate the marginal effect at means (MEMS) of the predictors on the probability of utilizing portability (See appendix in Table 3.11)

are 10 alternate shops in the vicinity. This significant jump is most likely due to a decrease in the cost of accessing an alternate shop with increasing number of shops in the neighborhood.

We conducted variable decomposition analysis (Grömping, 2007; Azen and Traxel, 2009; Tonidandel and LeBreton, 2010) for the logistic regression model to establish the relative importance of different factors in explaining the variation in utilization of portability. Column (1) of Table 3.6 shows that availability of choice as measured by number of shops around the household was the most important determinant of portability utilization and explained 67% of the variation. Location of the household (urban / town / village) and number of days the home shop is open explained 17% and 8% of the variation in portability utilization, respectively. In light of these findings, it is worth noting that close to 27% of households did not have an alternate shop within 1 km of their home shop. Arguably, the cost of accessing portability is high for these beneficiaries and the monopoly of shop dealers continues to persist in these regions despite the implementation of portability. Furthermore, 50% of the households without an alternate shop within a 1 km reside, on average, more than 80 km from the nearest administrative headquarters, which can lead to poor program implementation (Krishna and Schober, 2014).

3.5.2 Number of shops used by portability users

In this section, we quantify the association between the number of shops used by portability users ('FPS $Count$') during the six months of our study period and the number of alternatives available in the vicinity (FPS $density$ (x km)) while controlling for home shop and household characteristics (HHC). Given that FPS $Count$ is a count variable, we estimate a Poisson regression model (Cameron and Trivedi, 2013) as shown below on a subsample of portability users (households which did not use portability even once in 6 months use only one shop, their home shop) and whose home shop was open in all the six months.

$$ln(FPS\ Count_i) = \alpha_d + \beta\ (FPS\ density)_i + \gamma_1\ (FPSC)_{it} + \gamma_2\ (HHC)_i + \epsilon_i \quad (3.2)$$

While all other FPSCs and HHCs are time invariant and can be directly used in this cross-sectional model, we control for the effect of the number of open days of the home shop by calculating the average of '$Shop\ open\ days_{it}$' over six months.

Column 2 of Table 3.5 shows the estimated incidence rates of the predictors in the above model 3.2. All else being equal, portability users with more shops within 0.5 km from their home shop use more shops for making their transactions but the magnitude of the marginal effect is negligible. One additional shop within 0.5 km is associated with an increase of 0.4% in the incidence rate of the number of shops used. We also find that AAY and elderly households, and households whose home shops are open for a greater number of days use lesser number of shops.

Results of variable decomposition analysis for this model are shown in Column (2) of Table 3.6. We find that availability of choice as measured by number of shops around the household explains 63% of the variation and is the most important determinant of number of shops used. This suggests that shop density plays a critical role in the government's objective of stimulating competition among shop owners. It is plausible that when a household is surrounded by more shops it has greater number of economically viable options to choose from, which may incentivize shop owners to attract them through improved service quality.

3.5.3 Number of months in which portability is utilized

Here, we quantify the association between frequency of utilizing portability, i.e., number of months in which portability is used by a household ($Month\ Count$) and the number of alternatives available in its vicinity ($FPS\ density$ (x km)) on a subsample of portability users and whose home shops were open in all six months using the following Poisson regression model (Cameron and Trivedi, 2013):

$$ln(Month\ Count_i) = \alpha_d + \beta\ (FPS\ density)_i + \gamma_1\ (FPSC)_{it} + \gamma_2\ (HHC)_i + \epsilon_i \tag{3.3}$$

Similar to model 3.2 in Section 3.5.2, we control for the effect of the number of open days of the home shop by using the average of '$Shop\ open\ days_{it}$' over six months.

Column (3) of Table 3.5 shows the incidence rates of the predictors of the above model. We find a positive association between the shop density in 0.5 to 1 km radius and frequency of utilizing portability, but its magnitude is not practically significant (0.24%). Interestingly, we do not find a significant association with shop density within 0.5 km radius. These findings, along with results in Column (1) of Table 3.5, suggest that households' decision of whether to use portability at all and how frequently to use it may be driven by different factors. This is partly validated by the decomposition analysis shown in Column (3) of Table 3.6, which shows that FPS characteristics, other than shop density, explain 40% of the variation in the number of months in which portability is utilized as against less than 9% Table 3.6 in Column (1).

3.6 Robustness Checks and Model Extensions

In this section, we estimate two additional models to: i) rule out alternate explanations for our findings that are not explicitly accounted for by our predictor variables, and ii) include non-linear effects of shop density on the outcomes.[22]

[22] In addition, we also test the possibility of quality of road network in a region acting as a confounding variable using three variables - (i) Availability of public buses, (ii) Availability of private buses, and (iii) Availability of Auto rikshaws. We find that household with access to public buses, private buses and auto rickshaws are 11%, 14% and 9% more likely to use portability respectively. However, addition of these controls does not significantly change the coefficient of interest. Therefore, we omit discussion of these results for brevity.

3.6.1 Migration related utilization of portability

As discussed in Section 3.4.1, patterns of portability utilization suggest that some households transact at alternate shops that are very far from their home shops. Such transactions are possibly driven by factors such as migration and not due to poor quality of service or inconvenience at the home shop. As a result, the impact of shop density on utilization of alternate shops may be lower for such households. As we do not have access to household level migration data, we construct a distance-based proxy measure and re-estimate our model by excluding such observations from our sample.

For this analysis, we use a subsample of 119,155 transactions (4.97% of all observations), where the distance between a household's alternate shop (where the transaction occurred) and its home shop is greater than 40 kms (95^{th}) percentile of the distribution presented in Figure 3.1) as these transactions are more likely to be driven by migration.[23] Estimation results for this model are shown in Column (1) of Table 3.7. As expected, the association of home shop and administrative characteristics with odds of utilizing portability is higher in this subsample. Specifically, the odds of utilizing portability associated with '*FPS density* (0.5 km)' increases from 6.7% (see Column (1) of Table 3.5) to 8.1%. Column (2) and Column (3) of Table 3.7 contain corresponding results for the remaining two outcome variables: number of shops used for portability and number of months of portability utilization. We find that the impact of '*FPS density*(0.5 km)' on odds of utilizing portability increases from 0.39% to 0.50% for the number of shops. Marginal impact of '*FPS density*(0.5 km)' continues to be statistically insignificant for the number of months in which portability is utilized.

[23] We re-estimate the model considering another proxy measure as all transactions where a household's alternate shop (where the transaction occurred) and its home shop are in different sub-districts. There are 159,940 such transactions (6.68% of the sample with all observations), median distance between the home shop and alternate shop for these observations is 22.81 km. Estimation results are shown for in Table 3.12.

3.6.2 Non-Linear Effects of Shop Density on Portability Utilization

Prior research suggests that the marginal value of each additional alternative may not be constant, especially when the alternatives are not adequately differentiated (Ackerman and Gross, 2006). In our context, where all shops distribute the same commodities, we contend that a major source of differentiation is likely to be the (direct and indirect) transaction cost incurred by the households for accessing their entitlements at a shop and hauling it back, which is a function of the distance to the shop. We posit that the reduction in this transaction cost due to an additional shop is likely to be lower for households that have a greater number of shops in their proximity compared to those that have fewer shops. As a result, we expect shop density to have a non-linear (diminishing marginal) effect on portability utilization.[24]

To estimate this non-linear effect, we include squared terms of the variable 'FPS *density*(x km)' to model 3.1. Table 3.8 shows that, in line with our expectations, the coefficients of square terms of 'FPS *density*(x km)' are negative and statistically significant while the linear terms continue to be statistically significant. Similar to Section 3.5.1, we compute probability of utilizing portability as a function of shops in the vicinity for a typical household. Figure 3.3, plots the likelihood of using portability as a function of number of shops in the vicinity within 0.5 km radius. In contrast to Figure 3.2, it shows a non-linear increase in the likelihood of utilizing portability with the number of options (shops) available in the vicinity of a household. However. the magnitude of this increase (slope of the curve) decreases with the number of options. For instance, addition of the first shop in 0 to 0.5 km radius increases portability utilization by ≈ 2 percentage points (11.01% to 13.04%)

[24] Arguably, similar rationale does not apply to the other two outcome variables, frequency of portability use and number of unique shops used for portability. Nonetheless, for completeness, we estimate the non-linear variants of those models and find mixed support for non-linear effects. The results are available in the Appendix in Table 3.13.

whereas addition of the 7th shop increases the magnitude only by 0.6 (21.07% to 21.71%) percentage points.

3.7 Conclusion and future work

Poor quality of service delivery by monopolistic private agents in food security programs is a major impediment to achieving the goal of 'Zero hunger' in developing countries. Policy makers in various countries have attempted to address this issue through different interventions that aim to induce competition by increasing the choices available to beneficiaries. In this paper, we provide evidence on the utilization of one such large-scale intervention in India's PDS, termed *portability*, wherein beneficiaries were provided the choice of availing their food grain entitlements from any of the licensed shops in the state. Using program data from one of the earliest adopters of this initiative (the state of Andhra Pradesh), we find that the number of shops in a household's vicinity is a much more important determinant of utilization of portability and the number of alternate shops used in comparison to other program and household characteristics. These findings imply that the potential benefit of the *portability* initiative is less likely to be realized unless fundamental changes are made to the shop network, with special attention being paid to regions with lower shop density.

Although the results of our study provide valuable insights into the utilization of portability by beneficiaries and its association with shop density, its design has certain limitations. First, our outcome variable, utilization of portability, may not have a simple and direct relationship with beneficiary welfare due to the complex dynamics involved in the response of shop dealers and beneficiaries. For instance, on the one hand, if all shops improved the quality of their service due to the competition induced by portability, households may experience a welfare improvement without necessarily utilizing portability. On the other hand, if shop dealers collude to not serve each other's beneficiaries, utilization of portability and beneficiary welfare may be low. Second, any evaluation of welfare implications of providing portability must also account for the potentially higher costs incurred by both beneficiaries and the program. Beneficiaries are likely to incur higher transportation and time costs to avail their entitlements from an alternate shop. Similarly, utilization of portability can increase variability in demand at each shop and additional inventory may be required to reduce the chance of stock outs. Indeed, we find that the average inventory of grains carried by all shops together was 22% more than the total monthly entitlements of beneficiaries in the state. Consequently, operational policies will need to be modified to balance the trade-off between higher beneficiary welfare and program costs.

Future research that addresses these limitations is needed to inform and improve the design of *intra-state portability* as well as *inter-state portability* as currently envisioned by the Government of India (Today, 2019). Findings from future studies can also help several other developing countries such as Indonesia, Bangladesh, Ethiopia and Malawi, which are currently revamping their food security programs, to explore options beyond the apparent dichotomy embodied in the cash vs. in-kind as well as privatization vs. government monopoly debates (Ahmed et al., 2009; Bailey and Hedlund, 2012; Fernández, 2015; Banerjee et al., 2018).

Tables and Figures

Table 3.1: Number of months in which portability was utilized

	% portability users
One month	18.90%
Two months	11.40%
Three months	9.90%
Four months	11.70%
Five months	14.20%
Six months	34.20%

Notes: Percentages are computed with base as portability users which constitute 27.5% of 13.92 MN beneficiary households in AP PDS

Table 3.2: Number of months in which portability was utilized

		Number of months in which portability was used						
		(1)	(2)	(3)	(4)	(5)	(6)	Total
Number of months in which grains were collected from PDS	(1)	1.5%						1.5%
	(2)	0.7%	1.7%					2.4%
	(3)	0.8%	0.8%	2.2%				3.8%
	(4)	1.3%	0.9%	0.9%	3.9%			7.0%
	(5)	3.1%	1.9%	1.8%	1.9%	9.1%		17.8%
	(6)	11.6%	6.1%	4.9%	5.9%	5.1%	34.2%	67.7%
	Total	18.9%	11.4%	9.9%	11.7%	14.2%	34.2%	100%

Notes: Percentages are computed with base as portability users who constitute 27.5% of 13.92 MN beneficiary households in AP PDS

Figure 3.1: Distribution of portability transactions by distance from home shop

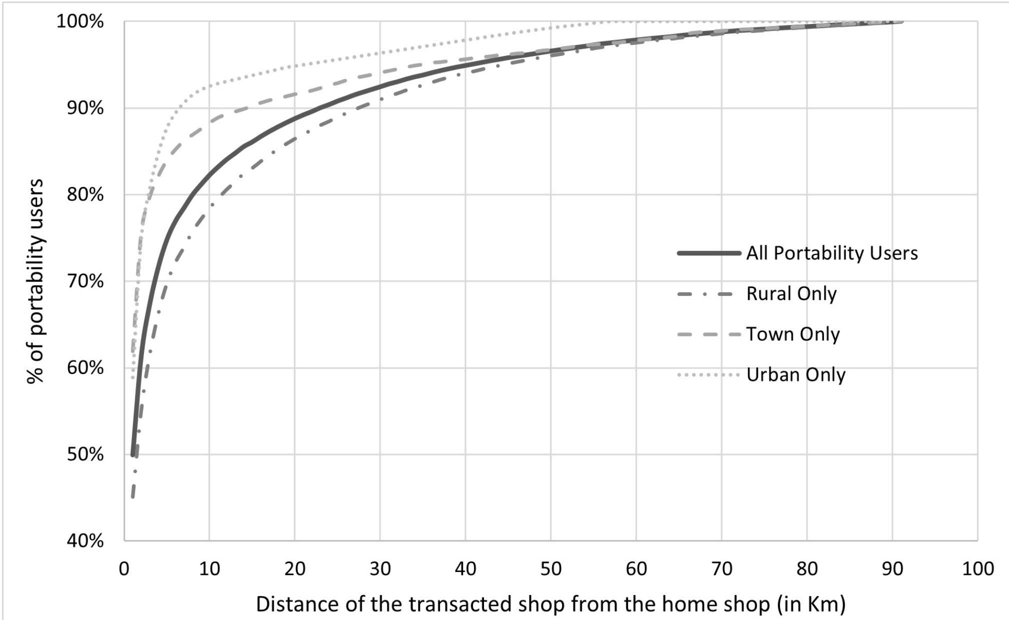


Table 3.3: Distribution of portability users by the number of alternate shops used

	% portability users
One shop	73.15%
Two shops	21.60%
Three shops	4.45%
Four or more shops	0.80%

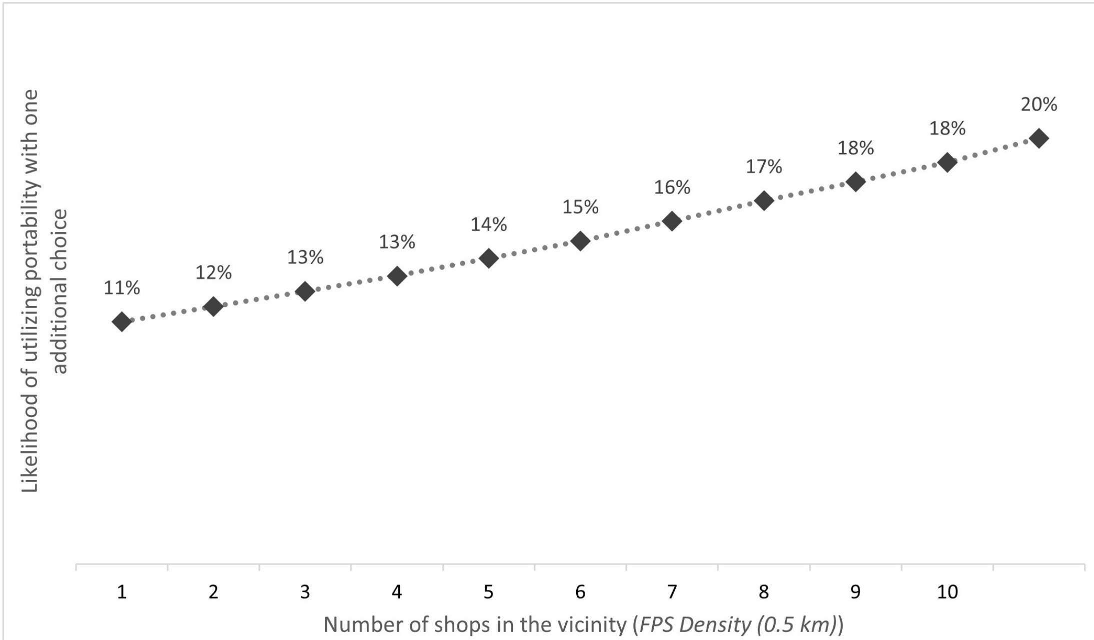
Notes: Percentages are computed with base as portability users who constitute 27.5% of 13.92 MN beneficiary households in AP PDS

Table 3.4: Descriptive statistics of factors affecting portability utilization

	Mean (SD)	10^{th} percentile	25^{th} percentile	Median	75^{th} percentile	90^{th} percentile
FPS density (0.5 km)	3.04 (4.26)	1	1	2	4	7
FPS density (0.5km - 1.0 km)	2.36 (4.59)	0	0	0	2	8
FPS density (1.0km - 2.0 km)	5.55 (9.64)	0	0	2	5	17
FPS characteristics (FPSC)						
FPS open days	12.90 (2.67)	9	11	14	15	16
Dealership *						
Co-operative	5.22 %					
Private male	48.29 %					
Private female	46.53 %					
Household characteristics (HHC)						
AAY *	6.24 %					
SC/ST/PTG *	11.15 %					
Elderly *	6.04 %					
Location *						
Rural	78.61 %					
Town	11.32 %					
Urban	9.90 %					

Notes: Variables in * are categorical variables. Summary generated from the randomly selected sample, panel data of 0.5 MN beneficiaries observed over 6 months. The numbers represent the percentage of households. FPS density (x km) - number of shops within x km from the home shop, FPS open days - number of days the home shop is open in a month, Dealership – the characteristics of the agent (male, female or an SHG), AAY (=1) – if household is economically vulnerable, SC/ST (=1) – if households is in socially vulnerable region, Elderly (= 1) – if the households has a member over 60 years in age. Location – urban, town or rural based on household's address

Figure 3.2: Probability of utilizing portability as a function of the number of shops within 0.5km radius



Notes:50% of the beneficiary households have 2 or less shops, 75% have 4 or less shops and 95% have 10 or less shops within 0.5 km radius from their home shop.

Table 3.5: Estimated effect of FPS Density (x km) on odds of utilizing portability, incidence rates on number of shops used and the number months in which portability was utilized

	(1)	(2)	(3)
	DV = Portability Utilization	DV = Number of shops used	DV = Number of months used
FPS density (0.5 km)	0.068***	0.0039 ***	0.0014
	(0.0059)	(0.0011)	(0.0013)
FPS density (0.5 km – 1 km)	0.019***	0.0018*	0.0024**
	(0.0041)	(0.0010)	(0.0009)
FPS density (1 km – 2km)	0.009***	0.0009	-0.0007
	(0.0019)	(0.0006)	(0.0004)
FPS Characteristics			
FPS open days	-0.067***	-0.0126 ***	0.0119***
	(0.0044)	(0.0021)	(0.0016)
Dealership (Base = Co-operative)			
Private male	-0.175***	-0.0079	0.0089
	(0.0536)	(0.0041)	(0.0062)
Private female	-0.168***	-0.0008	-0.0270
	(0.0533)	(0.0159)	(0.0176)
Household Characteristics			
AAY	-0.293***	-0.0234***	-0.0529***
	(0.0204)	(0.0075)	(0.0094)
SC/ST	-0.162***	-0.0219	-0.0265***
	(0.0520)	(0.0138)	(0.0086)
Elderly	-0.243***	-0.0558***	-0.0280 *
	(0.0190)	(0.0072)	(0.0152)
Location (Base = Rural)			
Semi-Urban	0.371***	0.0206	0.0284**
	(0.0533)	(0.0153)	(0.0126)
Urban	0.238***	0.0402	-0.0296
	(0.0754)	(0.0410)	(0.0217)
Number of observations	2,328,981	78,801	78,801
Pseudo R2	0.049	0.0030	0.0035
Wald Chi2 (23)	8,585.38	328.34	311.6
Prob >Chi2	0.000	0.000	0.000

Notes: All models are estimated with district fixed effects and errors clustered at sub-district month level on the randomly selected sample. FPS density (x km) - number of shops within x km from the home shop, FPS open days - number of days the home shop is open in a month, Dealership – the characteristics of the agent (male, female or an SHG), AAY (=1) – if household is economically vulnerable, SC/ST (=1) – if households is in socially vulnerable region, Elderly (= 1) – if the households has a member over 60 years in age. Location – urban, town or rural based on household's address

Table 3.6: Variance decomposition of odds of utilizing portability, incidence rates on number of shops used and the number months in which portability was utilized

Coefficient	(1) DV=Portability utilization	(2) DV = Number of shops used	(3) DV = Number of months used
FPS density (0.5 km)	37.02%	16.67 %	8.21 %
FPS density (0.5 km – 1 km)	19.23%	22.59 %	18.53 %
FPS density (1 km – 2km)	11.26%	25.87 %	14.12 %
FPS Characteristics			
FPS open days	7.72%	12.28 %	31.91 %
Dealership	0.97%	0.94 %	8.23 %
Household Characteristics			
AAY	3.68%	1.93 %	10.87 %
SC/ST	0.74%	0.91 %	1.74 %
Elderly	1.90%	5.98 %	1.82 %
Location	17.37%	12.83 %	4.58 %

Notes: FPS density (x km) - number of shops within x km from the home shop, FPS open days - number of days the home shop is open in a month, Dealership – the characteristics of the agent (male, female or an SHG), AAY (=1) – if household is economically vulnerable, SC/ST (=1) – if households is in socially vulnerable region, Elderly (= 1) – if the households has a member over 60 years in age. Location – urban, town or rural based on household's address

Table 3.7: Estimated effect of FPS Density (x km) on odds of utilizing portability, incidence rates on number of shops used and the number months in which portability is utilized on a subsample excluding households transacting at a shop beyond 40 kms from home shop

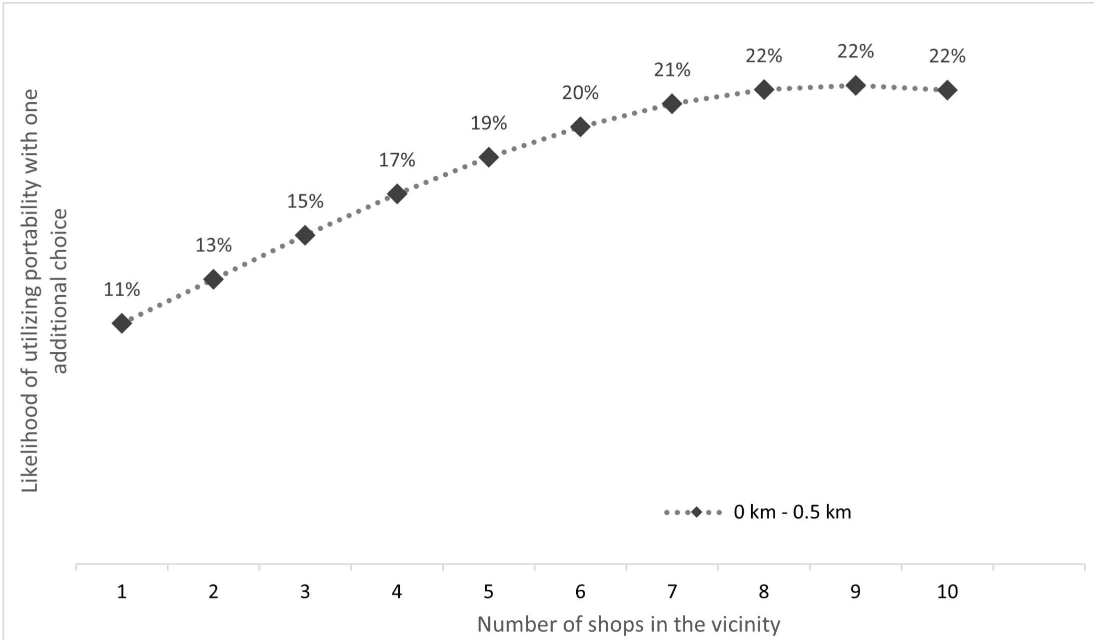
	(1)	(2)	(3)
	DV = Portability Utilization	DV = Number of shops used	DV = Number of months used
FPS density (0.5 km)	0.081***	0.0050 ***	0.0018
	(0.0065)	(0.0012)	(0.0014)
FPS density (0.5 km – 1 km)	0.197***	0.0018	0.0023**
	(0.0044)	(0.0011)	(0.0010)
FPS density (1 km – 2km)	0.009***	0.0010	0.0008*
	(0.0019)	(0.0007)	(0.0004)
FPS Characteristics			
FPS open days	-0.071***	-0.0147 ***	0.0130***
	(0.0050)	(0.0023)	(0.0019)
Dealership (Base = Co-operative)			
Private male	-0.200***	-0.0126	0.0094
	(0.0587)	(0.0045)	(0.0070)
Private female	-0.191***	0.0008	-0.0251
	(0.0583)	(0.0177)	(0.0195)
Household Characteristics			
AAY	-0.282***	-0.024***	-0.0546***
	(0.0236)	(0.0076)	(0.0105)
SC/ST	-0.155***	-0.0251*	-0.0267
	(0.0668)	(0.0145)	(0.0160)
Elderly	-0.189***	-0.0542***	-0.0181**
	(0.0210)	(0.0069)	(0.0092)
Location (Base = Rural)			
Semi-Urban	0.364***	0.0335	0.0292**
	(0.0570)	(0.0179)	(0.0132)
Urban	0.278***	0.0568	-0.0322
	(0.0756)	(0.0416)	(0.0199)
Number of observations	2,274,725	68,486	68,486
Pseudo R2	0.060	0.0041	0.0047
Wald Chi2 (23)	9,016.94	463.58	329.62
Prob >Chi2	0.000	0.0000	0.000

Notes: All models are estimated with district fixed effects and errors clustered at sub-district month level on the randomly selected sample. FPS density (x km) - number of shops within x km from the home shop, FPS open days - number of days the home shop is open in a month, Dealership – the characteristics of the agent (male, female or an SHG), AAY (=1) – if household is economically vulnerable, SC/ST (=1) – if households is in socially vulnerable region, Elderly (= 1) – if the households has a member over 60 years in age. Location – urban, town or rural based on household's address

Table 3.8: Estimated effect of FPS Density (x km) on odds of utilizing portability with quadratic terms of FPS Density included

	(1)
	DV = Portability Utilization
FPS density (0.5 km)	0.259***
	(0.0155)
FPS density (0.5 km – 1 km)	0.060***
	(0.0081)
FPS density (1 km – 2km)	0.014***
	(0.0030)
FPS density (0.5 km)2	-0.013***
	(0.0008)
FPS density (0.5 km – 1 km)2	-0.002***
	(0.0003)
FPS density (1 km – 2km)2	-0.000***
	(0.0000)
FPS Characteristics	
FPS open days	-0.066***
	(0.0040)
Dealership (Base = Co-operative)	
Private male	-0.169***
	(0.044)
Private female	-0.164***
	(0.044)
Household Characteristics	
AAY	-0.286***
	(0.014)
SC/ST	-0.126***
	(0.042)
Elderly	-0.239***
	(0.014)
Location (Base = Rural)	
Semi-Urban	0.152***
	(0.062)
Urban	0.168**
	(0.093)
Number of observations	2,328,981
Pseudo R2	0.055
Wald Chi2 (23)	14,041.28
Prob >Chi2	0.000

Figure 3.3: Non-Linear effect of number of shops within 0.5km radius on portability utilization



Notes:50% of the beneficiary households have 2 or less shops, 75% have 4 or less shops and 95% have 10 or less shops within 0.5 km radius from their home shop.

Appendix

Table 3.9: Comparison of descriptive statistics of all factors affecting portability between the sample and the population

	Mean (SD)	
	Population	Sample
FPS density (0.5 km)	3.04 (4.26)	3.02 (2.97)
FPS density (0.5km - 1.0 km)	2.36 (4.59)	2.26 (4.43)
FPS density (1.0km - 2.0 km)	5.55 (9.64)	5.15 (9.41)
FPS characteristics (FPSC)		
FPS open days	12.9 (2.67)	12.38 (3.78)
Dealership *		
- Co-operative	5.22%	5.24%
- Private male	48.25%	48.29%
- Private female	46.53%	46.47%
Household characteristics (HHC)		
AAY *	6.24%	6.24%
SC/ST/PTG *	11.15%	11.83%
Elderly *	6.04%	6.40%
Location *		
- Rural	78.78%	79.27%
- Town	11.32%	11.34%
- Urban	9.90%	9.39%

Table 3.10: Estimated effect of FPS Density (x km) on odds of utilizing portability, incidence rates on number of shops used and the number months in which portability is utilized on all households (including those households for which home shop did not see any transaction in a given month)

	(1)	(2)	(3)
	DV = Portability Utilization	DV = Number of shops used	DV = Number of months used
FPS density (0.5 km)	0.064***	0.0055***	0.0019
	(0.0056)	(0.0011)	(0.0161)
FPS density (0.5 km – 1 km)	0.014***	0.0021**	-0.0023**
	(0.0040)	(0.0009)	(0.0131)
FPS density (1 km – 2km)	0.008***	0.0006	-0.0005
	(0.0019)	(0.0007)	(0.0109)
FPS Characteristics			
FPS open days	-0.129***	-0.0113 ***	-0.0032**
	(0.0120)	(0.0011)	(0.0150)
Dealership (Base = Co-operative)			
Private male	-0.137***	-0.0049	-0.011*
	(0.0550)	(0.0045)	(0.0348)
Private female	-0.186***	-0.0011	0.0610***
	(0.0545)	(0.0132)	(0.0529)
Household Characteristics			
AAY	-0.263***	-0.0125***	-0.047***
	(0.0194)	(0.0066)	(0.0084)
SC/ST	-0.137***	-0.0231	-0.022
	(0.0537)	(0.0143)	(0.0166)
Elderly	-0.245***	-0.0534***	-0.027***
	(0.0178)	(0.0065)	(0.0079)
Location (Base = Rural)			
Town	0.3796***	0.0104	0.043***
	(0.0547)	(0.0152)	(0.0617)
Urban	0.3403***	0.0465	-0.0034
	(0.0794)	(0.0376)	(0.0707)
Number of observations	2,393,880	97,018	97,018
Pseudo R2	0.082	0.0036	0.0033
Wald Chi2 (23)	8,727.59	442.77	282.99
Prob >Chi2	0.000	0.000	0.000

Table 3.11: Estimated effect of FPS Density (x km) on likelihood of utilizing portability, predicted counts for number of shops used and number of months used

	(1)	(2)	(3)
	DV = Portability Utilization	DV = Number of shops used	DV = Number of months used
FPS density (0.5 km)	0.0075***	0.0051***	0.0055
	(0.0007)	(0.0014)	(0.0052)
FPS density (0.5 km – 1 km)	0.0021***	0.0023*	0.0095**
	(0.0005)	(0.0013)	(0.0038)
FPS density (1 km – 2km)	0.0010***	0.0012	0.0029
	(0.0002)	(0.0008)	(0.0018)
FPS Characteristics			
FPS open days	-0.0079***	-0.0163 ***	0.0464***
	(0.0004)	(0.0027)	(0.0243)
Dealership (Base = Co-operative)			
Private male	-0.0234***	-0.0102	0.0348
	(0.0075)	(0.0054)	(0.0243)
Private female	-0.0224***	-0.0010	-0.105
	(0.0075)	(0.0207)	(0.0694)
Household Characteristics			
AAY	-0.035***	-0.0302***	-0.208***
	(0.0019)	(0.0096)	(0.0373)
SC/ST	-0.0191***	-0.0283	-0.110*
	(0.0058)	(0.0178)	(0.0601)
Elderly	-0.0290***	-0.0722***	-0.104***
	(0.0019)	(0.0095)	(0.0339)
Location (Base = Rural)			
Town	0.0393***	0.0265	0.111**
	(0.0080)	(0.0195)	(0.0493)
Urban	0.0256**	0.0516	-0.116
	(0.0106)	(0.0524)	(0.0853)
Number of observations	2,328,981	78,801	78,801

Table 3.12: Estimated effect of FPS Density (x km) on odds of utilizing portability, incidence rates on number of shops used and the number months in which portability is utilized on a subsample excluding households transacting in a sub-district different from its home shop

	(1)	(2)	(3)
	DV = Portability Utilization	DV = Number of shops used	DV = Number of months used
FPS density (0.5 km)	0.091***	0.0067***	0.0036**
	(0.0073)	(0.0014)	(0.00165)
FPS density (0.5 km – 1 km)	0.020***	0.0021*	0.0026
	(0.0047)	(0.0012)	(0.0011)
FPS density (1 km – 2km)	0.009***	0.0010	0.0011
	(0.0022)	(0.0007)	(0.0006)
FPS Characteristics			
FPS open days	-0.081***	-0.0172 ***	0.0138***
	(0.0051)	(0.0025)	(0.0022)
Dealership (Base = Co-operative)			
Private male	-0.215***	-0.0134	0.0111
	(0.0665)	(0.0052)	(0.0084)
Private female	-0.207***	0.0001	-0.0245
	(0.0662)	(0.0204)	(0.0229)
Household Characteristics			
AAY	-0.277***	-0.0283***	0.0571***
	(0.0236)	(0.0074)	(0.0117)
SC/ST	-0.179***	-0.0298*	-0.0304
	(0.0668)	(0.0173)	(0.0209)
Elderly	-0.120***	-0.0448 ***	0.0106
	(0.0210)	(0.0075)	(0.0101)
Location (Base = Rural)			
Town	0.610***	0.0390	0.0597***
	(0.0647)	(0.0189)	(0.0170)
Urban	0.445***	0.0767	-0.0193
	(0.0878)	(0.0425)	(0.0235)
Number of observations	2,233,940	56,774	56,774
Pseudo R2	0.075	0.0053	0.0066
Wald Chi2 (23)	8,098.26	664.10	310.10
Prob >Chi2	0.000	0.000	0.000

Table 3.13: Estimated effect of FPS Density (x km) on incidence rates on number of shops used and the number of months in which portability was utilized with quadratic terms of FPS Density included

	(1)	(2)
	DV = Number of shops used	DV = Number of months used
FPS density (0.5 km)	0.0043	0.0008
	(0.00315)	(0.0004)
FPS density (0.5 km – 1 km)	0.0062***	0.0063***
	(0.0018)	(0.0004)
FPS density (1 km – 2km)	0.0028**	0.0020**
	(0.0013)	(0.0007)
FPS density (0.5 km)2	0.0000	0.0000***
	(0.0002)	(0.0000)
FPS density (0.5 km – 1 km)2	-0.0002***	-0.0002
	(0.0000)	(0.0000)
FPS density (1 km – 2km)2	0.0000*	0.0000**
	(0.0000)	(0.0000)
FPS Characteristics		
FPS open days	-0.0124***	0.0120***
	(0.0021)	(0.0016)
Dealership (Base = Co-operative)		
Private male	-0.008	0.009
	(0.0042)	(0.0062)
Private female	-0.001	-0.0270
	(0.0159)	(0.01778)
Household Characteristics		
AAY	-0.0225***	-0.0523***
	(0.0075)	(0.0095)
SC/ST	-0.0214	-0.0279
	(0.0137)	(0.0148)
Elderly	-0.0557***	-0.0264***
	(0.0072)	(0.0086)
Location (Base = Rural)		
Semi-Urban	0.0035	-0.0154
	(0.0176)	(0.0134)
Urban	0.0333	-0.0343
	(0.0419)	(0.0218)
Number of observations	78,801	78,801
Pseudo R2	0.0032	0.0036
Wald Chi2 (23)	358.25	343.22
Prob >Chi2	0.0000	0.000

Chapter 4

Impact of store choice on demand distribution and welfare in public sector supply chains - Evidence from India's food security program

4.1 Introduction

Most public welfare schemes are designed such that individuals or agencies are given licenses to operate as a channel for disbursement of benefits. For instance, in India, pensions under the pension scheme, work and wages under employment guarantee schemes, nutritious food items under maternity benefit programmes, fertilisers for farmers under fertiliser subsidy schemes, among others are all disbursed by licensed agencies. Typically, a certain number of beneficiaries are affiliated to each of the licensed parties and served by them. Therefore, the beneficiaries depend on these service providers to receive their entitlement (Davis, 2004; CMS, 2005; Olken, 2006; Commission et al., 2015; Salunke, 2015; Jebaraj, 2019). This results in restricted access to, and / or poor quality of services (Kanda, 2018; Express, 2018).

Policy makers of certain welfare programs have introduced a choice feature to empower beneficiaries by reducing the dominance of service providers (Fotaki et al., 2005; Library Congress, 2016; CCS, 2018). This choice feature gives beneficiaries the prerogative to choose the provider(s) from whom they wish to receive the service. If beneficiaries exercise their choice, there will be a change in the demand pattern and an increase in temporal demand variability among service providers. To account for this demand variability, associated supply side modifications in replenishment and inventory policies will be required. the absence of associated supply side modifications can lead to a failed attempt at offering the choice efficiently (Appleby and Dixon, 2004; Ferlie et al., 2006; Damera, 2017).

Our context pertains to India's food security program, known as the Public Distribution System (PDS) in which several states have introduced a store choice feature called *portability* with the goal of breaking the monopoly of service providers (Special correspondent, 2014; Correspondent, 2016; Ramakrishnan, 2017). Portability enables beneficiaries to collect their entitlements from any licensed store within the state. In this paper, we evaluate the impact of the introduction of store choice, under circumstances in which no modifications were made to the store replenishment policy. We use large-scale PDS program data from the Indian state of Andhra Pradesh, which was one of the earliest ones to introduce portability in 2015. We first develop a model of beneficiaries' store choice using a multinomial logit model and estimate the drivers of choice. We find that the effect of distance is negative and that one standard deviation (0.43 km) increase in distance to a store leads to a 58.4% decrease in the likelihood of its usage. Furthermore, we find that the effect of number of days that the store is open is positive and for every extra day that a store is kept open, the likelihood of its usage increases by 1.1%. We also find that beneficiaries are twice more likely to purchase from the PDS than from the open market. Furthermore, conditional on buying from the PDS, we find that beneficiaries are 11 times more likely to purchase grains from the store that they were assigned to before the introduction of portability, indicating some presence of stickiness.

Our goal is to evaluate the impact of choice on program performance in terms of the welfare of beneficiaries. However, the absence of data before implementation makes it difficult to evaluate this impact using traditional impact evaluation techniques. Therefore, we use the estimates from the choice model to construct the baseline scenario (no choice) and evaluate the impact of choice. Using a simulation analysis, we compare the pre-portability performance with the post-portability performance where the pre-portability replenishment policy was to continue in the post-portability scenario. We find that the provision of choice increases the proportion of beneficiaries purchasing from the PDS by 5.4% and the average beneficiary utility by 12.04%. We find that the maximum attainable utility increase is 25.4%, which is not realised due to stock outs at the beneficiaries' most preferred FPS, which are to the tune of 5.96%. These stock outs result from continuing the replenishment policy defined during the pre-portability period, which does not account for increased demand variability experienced at the FPS after the provision of choice. Our findings imply that a large portion of potential welfare gain is not realised due to the absence of complementary modifications in the replenishment policy which can account for the demand variation resulting from the provision of choice.

The rest of the paper is structured as follows. In §2, we discuss our contribution to relevant streams of prior literature and in §3, we contextualise our study accompanied by data and measures. The beneficiary choice process, empirical methodology, model estimation and results are presented in §4. In §5, we describe the simulation analysis to evaluate the impact of choice. Finally, §6 presents the concluding remarks.

4.2 Literature review

In this study, we examine the impact of providing variety in terms of store choice on demand, operational outcomes measured as stock outs and welfare outcomes measured as sales. Our paper is therefore related to research within two streams of literature - (i) product variety and its impact on sales, and (ii) store choice models.

4.2.1 Impact of product variety on sales

Firms have offered product variety as a way to improve performance. The impact of product variety on sales has been studied extensively. Studies have found that high product variety can increase sales by allowing firms to satisfy the needs of heterogeneous consumers (Bayus and Putsis Jr, 1999; Xia and Rajagopalan, 2009).[1] Additionally, studies have highlighted that increasing variety increases the difficulty of managing inventory and can instead undermine sales due to stock outs (Fisher and Ittner, 1999; Ton and Raman, 2010), owing to the fact that product variety makes it harder to accurately forecast demand, thereby resulting in mismatches between supply and demand. While our study is in the context of store choice, all the issues arising from this choice provision are similar to the ones originating in the context of offering product variety. Other studies have focused on the impact of lower search costs, enabled by new information technology, on demand concentration (Cachon et al., 2008; Brynjolfsson et al., 2011; Zentner et al., 2013). For instance, Brynjolfsson et al. (2011) empirically analyze a retailer that offers the same product assortment online and offline and find that the online store exhibits less concentrated demand because of its lower search costs. Our paper is closely related to the following three papers which examine the impact of product variety decisions on sales and demand concentration.

Wan et al. (2012) study the impact of product variety decisions on an operational outcome and on sales. They find that the effect of product variety on sales would be overestimated if the indirect effect of operational outcome is overlooked. While the aforementioned study measures the indirect effect of variety on sales using fill rate (fraction of demand that is met) as a proxy for operational outcome, our paper measures the impact of variety with respect to store choice on both demand and sales

[1]Research has also suggested that excess product variety can result in negative consequences due to selection confusion for customers, thus reducing the marginal benefits from variety (Thompson et al., 2005). However, this point is not very pertinent to our setting given the manageable number of alternatives in the consideration set of beneficiaries.

using the estimates from the structural model to create a simulation model. Due to the design of our simulation analysis, we can observe both demand and sales and measure stock outs (without resorting to indirect measures of operational outcomes). Therefore, we can measure the welfare loss from lost demand at a store, which results from beneficiaries being unable to transact at their most preferred store.

Tan et al. (2017) find empirically that product variety is likely to increase demand concentration. This goes against the long tail effect which predicts that demand will become less concentrated on hit products because of expanded product variety. Similar to Tan et al. (2017), we measure the impact of variety on demand. Tan et al. (2017) aims to decrease supply-demand mismatch by studying the impact on demand concentration, which measures how the mean demand has changed. In our context, this mismatch which is measured as stock outs and is driven both by the change in mean demand as well as the temporal variability in demand within FPSs. Therefore, the measure of demand concentration alone is insufficient to quantify the loss due to the supply-demand mismatch in our context. Therefore, we measure the impact of variety on stock outs, which is a manifestation of both change in mean demand at the FPS as well as temporal variability within FPSs.

Jain and Tan (2021) study a setting in which effective product variety seen by customers is influenced by the sales channel. Using a natural experiment at a leading e-retailer that discontinued the PC sales channel, they find that the mobile channel increases the share of sales of popular products compared to the PC channel. They further examine the consequence of ignoring the differences between sales channels in terms of how it shapes sales concentration across products, on inventory management. Similarly, we examine the cost (in terms of welfare loss) of ignoring the effect of implementing store choice in changing the demand distribution as well. Unlike Jain and Tan (2021) who infer how sales concentration can affect operational decisions, we estimate how demand variability drives sales through stock outs and the resulting substitution by virtue of the nature of our simulation design. Therefore, our model allows us to study the implications on operational decisions by directly varying the degree of stock outs through alternate replenishment policy designs.

4.2.2 Store choice models

There is extant literature on store choice models (Bell and Lattin, 1998; Davis, 2006; Chernev and Hamilton, 2009; Briesch et al., 2009, 2013; Wang and Bell, 2015). Most papers that model store choice consider the effect of marketing variables such as pricing, promotions, bundling, assortment, retail price format, travel distance and category positioning. However, they do not focus on how choice affects demand variability, and this presents a gap. Therefore, we add to the literature on store choice by quantifying the effect of choice on demand variability, stock outs and welfare loss. Our paper is closely related to Kabra et al. (2020), who model station choice and empirically estimate the effect of station accessibility and bike availability in the context of bike-share systems. They use structural estimation to model customer choice and further use the estimates to evaluate operational decisions. Similar to Kabra et al. (2020), we too measure the intermediate operational outcome of resulting stock outs. Further, the availability of granular data (household level transaction) combined with the design of our simulation analysis enables us to observe the substitutions made and complements Kabra et al. (2020) by measuring welfare losses resulting from the underlying operational decisions.

4.3 Study setting and Data

4.3.1 Study setting: Indian Public Distribution System

India's food security program, known as the public distribution system (PDS) delivers subsidised food grains to nearly 160 million economically weaker households through a supply chain comprising of more than half a million fair price shops (FPS). The FPS constitutes the last mile delivery of the PDS supply chain. Each FPS has a set of beneficiary households affiliated to it and issues grains to only these households. Each beneficiary household, based on its economic status, is entitled to receive a defined quantity of food grains every month. Until recently, beneficiaries could collect their entitlement only from the FPS to which they were affiliated (referred to as the home FPS). This restriction to the home FPS is cited as one of the main causes of poor quality of service in PDS resulting fro concerns which include FPS not being open or being open at irregular times, mistreatment by the dealer, long queues, adulteration of food grains and non-availability of commodities (Sati, 2015; Sargar et al., 2014; Vaidya et al., 2014; Dreze and Khera, 2015; Dhanaraj and Gade, 2012; Sharma and Gupta, 2019). However, despite these issues, beneficiaries are constrained to use their home FPS due to the operational design of PDS.

To address the beneficiaries' inconveniences with the service quality of home FPS, several state governments have lately introduced the feature of *'portability'* (Business-Standard, 2018; NIC, 2018). This feature allows beneficiaries to buy grains from any FPS within the state by digitally authenticating their identity. It is expected that this choice would reduce the beneficiaries' dependence on a single FPS and hence lead to welfare improvements. However, there have been no supply side changes to accompany the introduction of portability. Based on our field visits and interactions with the officials of the food and civil supplies ministry, we noted no modifications were made to the replenishment policy post the introduction of portability.

4.3.2 Data

Andhra Pradesh (AP) was one of the first Indian states to introduce state-wide portability of PDS in India in 2015. We use publicly available program data related to PDS operations from the 13 districts of AP from April 2018 to August 2018.[2] We scraped the publicly available data related to 75.57 million PDS transactions made by 13.92 million beneficiaries at 29,212 FPSs during these five months. We use three primary datasets – beneficiary dataset, FPS dataset and transaction dataset. The beneficiary dataset contains demographic information about the beneficiary household. Each beneficiary household has a unique identifier called the Ration Card (RC) number. We have information about the economic status, gender and age profile of the household members, and home FPS of the household. The FPS dataset has information about the location (geographical coordinates) of the FPS, the number of days the FPS is open in a month and the average daily closing inventory details for the month. The transaction dataset has information on every transaction made by the beneficiary in these five months. We have details regarding the FPS where the beneficiary chose to transact and the day of transaction in the month.

4.4 Beneficiary choice model and drivers of choice

4.4.1 Beneficiary choice model

We model the beneficiary household's decision to buy grains from a particular FPS as a function of distance and FPS characteristics described in the previous section. We use a multinomial logit model, which is widely used to model product choice in the economics and marketing literature (Train, 2009; Ben-Akiva et al., 1985). We specify the indirect utility of household h from choosing to purchase from FPS f in month t as a function of several covariates, $\mathbf{Z}$, which are described in the following subsection.

$$U_{h,f,t} = \beta\,\mathbf{Z} + \epsilon_{h,f,t} \tag{4.1}$$

[2]Uploaded by the Department of Consumer Affairs, Food and Civil Supplies; `https://aepos.ap.gov.in/ePos/`

4.4.2 Drivers of choice

Distance travelled

We model the negative utility arising from beneficiaries' travel to the FPS using the variable, $Distance_{h,f}$, which denotes the distance travelled by household h to FPS f. Empirical studies have shown that distance is a vital factor influencing store choice for Indian retail shoppers (Sinha et al., 2002)). Koul and Mishra (2013) find that customers do not want to travel long distances to buy goods of daily usage. Jayasankara Prasad and Ramachandra Aryasri (2011), in the context of Indian consumers, find that the distance travelled to reach the store is significantly associated with retail format choice decisions. Furthermore, Dennis et al. (1999) find that shoppers often tend to patronise their nearest shopping malls more; hence, distance influences shopping mall attractiveness. In the context of grocery shopping, Prasad (2010) use primary data to demonstrate that the distance travelled to a store is a significant predictor of store choice behaviour. Due to unavailability of information on the location of beneficiaries, we use an indirect approach of arriving at a proxy for distance. We acquired information about the addresses (geographic coordinates) of the FPSs and used them to calculate the distance between different FPSs. We calculate the distance to different alternatives in the choice set, relative to the beneficiary's home FPS location.[3]

FPS characteristics

We define two FPS characteristics relevant to a household's utility: availability of grains in the FPS and the number of days for which the FPS is open in a month.

Availability of grains in the FPS. Several papers have shown that the

[3]Based on our field visits and discussion with officials at the food and civil supplies ministry, we understand that the allocation of beneficiaries to FPSs (home FPS) in the pre-portability period was based on the FPS closest to their place of residence. Hence we calculate distance to different alternatives in the choice set based on the beneficiary's home FPS location.

availability of products or services is a crucial factor in many consumer choice decisions (Mahajan and Van Ryzin, 2001). Availability can take two forms – expectation about the level of availability which appears in the decision making function and stock outs in which case the beneficiaries cannot choose a particular alternative. Marketing researchers who control for product availability in studying consumer choices have typically removed the unavailable SKUs from the choice set for all consumers. Accordingly, we allow the beneficiaries' choice set to vary across different purchase occasions (months) based on availability. Further, Swait and Erdem (2002) show in the context of brand choice that availability is an important component in the utility function of the consumer and that consumers opt to substitute a competing good if their preferred good is unavailable rather than searching for other locations or delaying their purchase. Since substitution of grains is not feasible in our context, it is more likely that beneficiaries opt to buy from a substitute FPS. We use two variables to model availability from the perspective of beneficiaries - $Average\ stock_{f,t}$ which denotes the average daily rice stock available at FPS f during month, t and $Opening\ balance_{f,t}$ which denotes the quantity of grains available at FPS f during month t on the beginning of the day when the beneficiary makes the purchase.[4]

Number of days for which the FPS is open in a month. FPSs are generally required to be open during the first 15 - 20 days of the month. If the FPSs are closed during the stipulated duration, it might result in the beneficiary's failed attempt to purchase grains and households may have to make multiple visits to the same FPS or a visit a different one. Both these options entail significant additional costs in terms of travel and time. We therefore hypothesise that the expectation about the FPS being open enters the utility function of beneficiaries. In fact, studies have shown that one of the major concerns among beneficiaries is that FPSs are not open for

[4] We estimate a model variant by using the historical number of instances that the FPS was stocked out on the day of beneficiary's purchase as a co-variate and find that the model results do not change.

the stipulated duration (Sharma and Gupta, 2019; Vaidya et al., 2014), resulting in beneficiaries having to make multiple trips or having to forgo buying from the PDS. Both these situations have a negative impact on their utility. We use the variable, $FPS\ open\ days_{f,t}$ which denotes the number of days FPS f was open in month t, to indicate beneficiaries' expectation about the FPS being open.[5]

Stickiness to home FPS

Papers which model customer choice also include the notion of loyalty in the model, which describes the customer's tendency to repurchase the same brand (Guadagni and Little, 1983). Existing literature has operationalised loyalty by introducing some measure of past customer purchase behavior as an explanatory variable. In our context, before the introduction of portability, beneficiaries purchased from the home FPS for the past several years. Therefore, we hypothesise that there would exist some sort of preference for the home FPS. We define an indicator variable, $Home\ FPS_{h,f}$ which takes the value one if the FPS f in consideration set of the household h represents a home FPS and zero otherwise.

4.4.3 Summary Statistics

We examine the evidence of households transacting at non-home FPS and find that on an average 18% of households use a non-home FPS in any given month. Furthermore, 28% of households use a non-home FPS at least once in the five month period (analysis period). In terms of number of alternative FPSs available for a household, on an average households have 3.2 FPSs within 0.5 km, 5.8 FPSs within 1 km and 12.2 FPSs within 2 km of their home FPS. The average distance travelled by beneficiaries as measured from the home FPS is 0.65 km. We also find that the average number of days an FPS is kept open in a month is 12.2 days. Table 4.1 provides details about

[5]We derive this variable by assuming that the FPS is deemed open on a particular day if at least one transaction is recorded against it on that day.

summary statistics for these variables.

4.5 Model specification and Results

4.5.1 Model specification

We specify the indirect utility of household h from choosing to purchase from FPS f in month t, as a function of several covariates including distance, FPS characteristics and an idiosyncratic household–month–FPS specific error term, using a multinomial logit model as shown below:

$$\begin{aligned} U_{h,f,t} &= \beta_1\ Distance_{h,f} + \beta_2\ FPS\ open\ days_{f,t-1} + \beta_3\ Average\ stock_{f,t-1} \\ &+\beta_4\ Opening\ balance_{f,t} + \beta_5\ Home\ FPS_{h,f} + \beta_6\ Open\ market_{h,f} + \epsilon_{h,f,t} \end{aligned} \tag{4.2}$$

where, $Distance_{h,f}$ = Distance from home FPS of household h to FPS f, $FPS\ open\ days_{f,t-1}$ = Number of days FPS f was open for business in month $t-1$, $Average\ stock_{f,t-1}$ = Average daily inventory at FPS f during month $t-1$, $Opening\ balance_{f,t}$ = Opening inventory on the day of purchase at FPS f during month t, $Home\ FPS\ indicator_{h,f}$ is an indicator variable which takes the value 1 if alternative f is a home FPS and 0 otherwise & $Open\ market\ indicator_{h,f}$ is an indicator variable which takes the value 1 if the alternative f is an open market (outside option) and 0 otherwise.

Although the beneficiary household can buy grains from any FPS in the state, we limit the household's choice set to its nearest FPSs within a distance threshold.[6] As mentioned in the previous section, we account for stock outs by allowing beneficiaries' choice set to vary across different months. We denote the set of nearest FPSs after removing the stocked out FPSs as $\mathcal{N}_h$. We also include the 'open market' in the choice

[6]threshold distance = 1.5 km is defined based on the 90th percentile of the distance travelled variable.

set to incorporate the no-choice option into the model. We characterise the no-choice option by assigning it a utility value of zero, as is the standard practice in literature (Chandukala et al., 2008). Therefore, the probability of household h choosing to make a purchase from FPS f in month t defined as a function of the household's indirect utility function is given by, $P_{h,f,t} = \frac{e^{E[U_{h,f,t}]}}{\sum\limits_{g \in N_h} e^{E[U_{h,g,t}]}}$.

4.5.2 Results

Table 4.2 reports the estimation results of the multinomial logit model described in equation (4.2).[7] Overall, we find that the effect of *Distance* is negative and significant and that the effect of *FPS open days* is positive and significant. This suggests that beneficiaries obtain a negative utility from farther FPSs and receive a positive utility when FPSs are more likely to be open. We also find that beneficiaries prefer their home FPS and derive a net negative utility from buying from the open market instead of PDS. Furthermore, if the distance to an FPS increases by 0.1 kilometre, the percentage of using that FPS decreases by 18.2%. The seemingly high magnitude of the coefficient can be explained by the fact that beneficiaries need to haul close to 20-35 kilograms of grain during the return journey after purchase. In terms of standard deviation, we find that if the distance to an FPS increases by one SD (0.43 km), the percentage of using that FPS decreases by 58.4%. We also find that if the number of days on which the FPS was open in the previous month is increased by one day, the percentage of using the said FPS increases by 1.13%. We find that beneficiaries are twice more likely to purchase from the PDS than from the open market. We also find that conditional on buying from the PDS, beneficiaries are 11 times more likely to purchase grains from the home FPS than from non-home FPSs.

Interestingly, we find that both the variables in the utility function which represent

[7] Processing this magnitude of data is a computational challenge. Therefore, we estimate the model using a sample chosen from the 72 million transactions.

the notion of stock availability at the FPS (*Average stock* and *Opening balance*) have no impact on the beneficiary's choice of FPS. Based on our field visits, it appears that the FPSs do not open for business if grains are not available at the FPS. If this is true, then beneficiaries do not really need to consider availability in their decision making as they can deduce that, if the FPS is open, availability of grains is ensured. It could be argued that, there are other items in the FPS and the FPS owner could find value in opening the shop to sell these other items despite not having grains. But, based on the understanding from our field visits, such a scenario seems unlikely in our context. Finally, we conduct validation checks to verify that our structural model is a close representation of the empirical data. (See section 4.8.2 in the Appendix for details.)

4.5.3 Estimation of heterogeneous preferences

Households differ in how they respond to choice, based on several dimensions related to their characteristics. In this section, we estimate heterogeneous preferences based on multiple dimensions. We extend our multinomial logit model to account for preference heterogeneity among decision makers based on observable characteristics (Ai and Norton, 2003; Vij and Krueger, 2017).

The effect of the drivers of choice in our setting is likely to be moderated by the economic status, gender composition and age composition of the households. We estimate heterogeneous estimates of drivers based on the following household characteristics: i) Economic status of the household, ii) Gender composition of the household and iii) Age composition of the household. Each household is classified into one of the two categories – Antyodaya Anna Yojana (AAY) and Priority Households (PHH). AAY comprises households that face extreme economic vulnerability and PHH comprises other low-income households. We use household category as a proxy to indicate economic status, because AAY households are amongst the poorest of the poor. We use number of female members in the household and number of elderly members in the household to indicate gender and age composition, respectively. We separately estimate three models for all three dimensions of household characteristics (HHC) using the following equation:

$$\begin{aligned} U_{h,f,t} = \beta_1\ Distance_{h,f} + \beta_2\ FPS\ days\ open_{f,t-1} + \beta_3\ Average\ stock_{f,t-1} \\ +\beta_4\ Opening\ balance_{f,t} + \beta_5\ Home\ FPS_{h,f} + \beta_6\ Open\ market_{h,f} \\ +\beta_7\ Distance_{h,f} \times HHC_h + \beta_8\ FPS\ open\ days_{f,t-1} \times HHC_h \\ +\beta_9\ Home\ FPS_{h,f} \times HHC_h + \beta_{10}\ Open\ market_{h,f} \times HHC_h + \epsilon_{h,f,t} \end{aligned} \tag{4.3}$$

where HHC_h is an indicator variable which represents the three dimensions of household characteristics described above and takes the value 1 if household h belongs to AAY category, more than half of members of the household h are women or if all members of the household h are above the age of 60 years respectively in each of the models separately and 0 otherwise.

Columns (1) to (3) of Table 4.3 show the coefficient estimates for the heterogeneous models estimated using the above equation. We find that the negative relationship between distance and choice is higher in female-dominated households (0.1 km increase in distance to an FPS leads to 19% decrease in likelihood of its usage as compared to 18% for households which are not female-dominated) and lower in economically vulnerable (AAY) households (9.8% for AAY households and 18.5% for non-AAY households). This indicates higher sensitivity of female-dominated households to travel longer distances and lower sensitivity of AAY households to travel longer distances. Therefore, economically vulnerable households are willing to travel longer distances to get their entitlements. We also find that economically vulnerable households and elderly households are more sticky to their home FPSs (28 times for AAY households compared to 10 times for non-AAY households and 13 times for elderly households compared to 11 times for non-elderly households). Finally, AAY, elderly and female-dominated households have a slightly higher preference to buy from the PDS than the open market compared to their respective counterparts.

4.6 Impact of choice on beneficiary welfare and demand variability

In this section, we evaluate the impact of portability on beneficiary welfare. As mentioned earlier, absence of data before portability makes it difficult to conduct this evaluation using traditional impact evaluation techniques like difference-in difference methods, regression discontinuity design or any other causal inference methods. Therefore, we use the coefficients that we estimated using our multinomial logit model in a simulation design framework to construct the baseline scenario of no-choice (where the beneficiary can purchase either from the home FPS or open market) and subsequently evaluate the impact of choice. Many papers have applied the technique of using the coefficients estimated from a discrete choice model to predict choices in either a hold out sample for model validation or a new sample for forecasting purposes (Keane and Wolpin, 2007; El Zarwi et al., 2017). Moreover, using a simulation design framework allows us to observe not only the revealed preference of the beneficiaries but also the first preferred choice and the ensuing substitution patterns in the instance of a stock out.

In the following subsections, we outline the simulation design and discuss the results of the simulation analysis. We use the coefficients from the choice model and the replenishment policy defined during the pre-portability period to simulate the pre- and post-portability scenarios and calculate the utility values under these scenarios. Thereafter, we compare the proportion of purchases from the PDS, welfare (utility) values, instances of stock outs and coefficient of variation of demand between pre- and post-portability scenarios. We finally compare the magnitude of the realised increase in utility relative to the maximum attainable increase in utility and compute the various sources of utility loss.

4.6.1 Simulation design

We use the estimates from the beneficiary choice model in a simulation framework to generate both pre- and post-portability scenarios.[8] In the pre-portability scenario, every household has only two choices: their home FPS and the open market. We use the coefficient estimates from the choice model and the value of covariates from the original data and predict the choice for every household as either buying from their home FPS or the open market. We calculate the replenishment quantity for each FPS by using the replenishment policy defined during the pre-portability period. The policy is similar to a base stock policy in which the order-up to level is equal to the gross requirement (GR) of the FPS. Gross requirement is calculated as the total entitlement quantity of all the households affiliated to the FPS, as defined before portability was introduced. We populate the values of all variables, except for $Average\ stock_{f,t-1}$ and $Opening\ balance_{f,t}$ from our data. We cannot use the historical data for $Average\ stock_{f,t-1}$ and $Opening\ balance_{f,t}$, because the value of these variables will depend on demand realisations. Therefore, we use their values for the first month from the original data and predict the choice for every household for the second month. Then, for every beneficiary who has bought from the PDS in that month, we use the same day of purchase as recorded in the original data. We combine both the choice decision and the day of purchase to arrive at the daily sales of each FPS. Further we calculate the *Opening balance* and *Average stock* for each

[8]The IIA assumption is important when the estimates from the multinomial logit model are used to predict choices on a subset of alternatives in the choice set. (Train, 2009). In our context, the IIA is to be verified between the FPSs in the PDS, and not as compared to the outside option. Hausman – McFadden test of independence from irrelevant alternatives (test of hypothesis that the parameters on the subset are the same as the parameters on the full set constitutes a test of IIA) has been verified by randomly dropping some alternative FPSs in the consideration set. Because the no purchase option being an outside option (open market) and a covariate in the model, is included in both scenarios (pre and post) of the simulation analysis, it has not been dropped while verifying the IIA assumption.

FPS and use these as the values of covariates for the third month. We continue this procedure for the remaining months.

We subsequently recreate the post-portability scenario where the pre-portability replenishment policy is to be continued. To do so, we use the same procedure as described above for the pre-portability scenario, but with the entire choice set, $\mathcal{N}_h$. We modify the choice sets dynamically based on stock out information derived from the replenishment quantity, availability and sales information. We cannot use our empirical data for the post-portability scenario because it only indicates the final choice made by the beneficiaries. In contrast, the simulated data allows us to observe the intermediate steps of most preferred choice and substitutions made by the beneficiaries in the instance of stock outs. We finally calculate four metrics for both pre- and post-portability scenarios: the coefficient of variation of demand, percentage instances of stock outs, percentage instances of purchases from the PDS and the average utility received by the beneficiaries.[9]

4.6.2 Results

Table 4.4 presents the results of the simulation analysis. As defined previously, the operational outcome can be measured in terms of variation and stock outs while welfare enhancement can be measured in terms of change in proportion of purchases made from the PDS, and change in utility value after the introduction of choice.

Operational outcome. We find that the average value of coefficient of variation of demand at the FPS has increased by 62.1%, from 0.037 to 0.06. Moreover, instances of stock outs at beneficiaries' most preferred FPS in the post-portability scenario is 5.96%. There are no stock outs in the pre-portability scenario because of the replenishment policy, which was an order upto level policy. The order upto level is the gross requirement calculated as the total entitlement quantity of all the

[9]See §4.8.1 in the Appendix for details of the steps involved in the construction of the pre- and post-portability scenarios.

households affiliated to the FPS. Considering the total instances where beneficiaries' most preferred FPS encountered stock outs, 85.4% could purchase grains at their second preferred FPS, while the remaining 14.6% had to resort to purchasing from their third preferred FPS. We find the following results: (i) 5.4% increase in purchases from the PDS, (ii) 12.04% increase in utility, (iii) 5.96% stock outs at beneficiaries' most preferred FPS and (iv) 62.1% increase in coefficient of variation of demand within FPSs over time.

Welfare outcome. We find a 5.4% increase in purchases from the PDS with 93% purchases from the PDS in the pre-portability scenario and 98% in the post-portability scenario. The estimated increase in utility is 12.04%. This includes instances where the households had to transact at the second or third preferred FPSs because of stock outs. If the stock outs could have been avoided, the increase in utility would have been 25.4%. Therefore, the households unable to buy from their most preferred FPS are forced to use their less preferred alternatives. This indicates that a large portion of potential welfare gain can be lost due to stock outs resulting from continuing the pre-portability replenishment policy.

Analysis of portability transactions

We further quantify and analyse the portability transactions resulting after provision of choice, as shown in Table 4.5. We find that in the absence of stock outs, the percentage of portability transactions would have been 24.8. However, in the presence of stock outs, 25.1% transactions are found to be portability transactions. While the two numbers are comparable, it is noteworthy that some households whose most preferred FPS is the home FPS were forced to buy from a non-home FPS due to stock outs in the former. We also find that there are 2.94% of such instances out of the total PDS transactions (4,044 transactions out of a total of 137,315 transactions) and we call this '*involuntary-portability*' transactions. On the other hand, again due to stock outs, some households were unable to buy grains from their most preferred FPS and had to buy from their home FPS instead. We find 2.62% of such instances out of the total PDS transactions (3,606 transactions out of a total of 137,315 transactions) and we call theis '*unsuccessful-portability*' transactions. Table 4.6 presents the utility loss associated with all possibilities resulting from the inability to transact at the most preferred FPS. It is evident that '*involuntary-portability*' and '*unsuccessful-portability*' transactions constitute 6.4% ($5.51 + 0.89$) and 5.83% of the total utility loss of 13.3%, respectively.[10]

To summarise, we find that after the introduction of portability, all households which prefer to transact within the PDS are able to buy their grains from FPSs. Therefore, portability seems to have been successful in eliminating all non-preferred open market transactions. The presence of *involuntary-portability* and *unsuccessful-portability* transactions indicate that tracking and measuring portability transactions alone can be misleading. A high proportion of portability transactions need not indicate proportional welfare enhancement, especially when complementary supply-side (replenishment policy) changes are not undertaken. As evidenced by our simulation results, 11.6% of the observed portability transactions (4,044 transactions out of of 34,581 portability transactions) are *involuntary-portability* transactions which does not denote true choice and consequently does not translate to welfare improvement.

4.7 Conclusion

Our paper employs a combination of structural estimation modelling and simulated counterfactual analysis to evaluate choice provision in a public welfare scheme. The estimates from the structural model of beneficiaries' store choice quantify the drivers

[10]We conduct validation checks to verify that the simulation model is a close representation of the empirical data. See §4.8.2 in the Appendix for details.

of choice and provide guidance to policy makers on future decisions related to portability. The results from our simulation analysis demonstrate that the extent of welfare enhancement due to provision of choice is driven by the demand variance and resulting stock outs. We find that a large portion of potential welfare gain is not realised due to the absence of associated supply side modifications.

Our paper provides early empirical evidence on the impact of choice provision in public sector supply chains. Introduction of portability in India has just begun and very soon other states will follow suit. In fact, portability is planned to be scaled, first allowing inter-state portability within a cluster of states and ultimately advancing to nation-wide portability, where beneficiaries will not be bound by state boundaries and can buy grains from any FPS across the country (Mishra, 2018; Today, 2019). This will only increase the demand variability even further, thereby intensifying the challenges associated with the supply of grains. Moreover, a substantial gain in utility depends on being able to avoid stock outs under the existing replenishment policy with portability. Therefore, change in replenishment policy design to mitigate stock outs is required for the realisation of welfare benefits in it's entirety, from portability. Consequently, replenishment and inventory planning become increasingly essential to ensure that appropriate amount of inventory is available at the FPSs at the right time and that the ultimate goal of beneficiary welfare is achieved. Future research should explore and design easily implementable alternate replenishment policies that can better match supply and demand.

Tables and Figures

Table 4.1: Summary statistics

	Mean	**Standard deviation**	**Median**
Number of alternatives (0.5 km)	3.19	3.46	2
Number of alternatives (1 km)	5.89	8.2	2
Number of alternatives (2 km)	12.21	18.87	5
Distance travelled from home FPS (km)	0.65	0.43	0.65
FPS open days	12.23	3.98	14

Table 4.2: Effects of drivers of choice

Variable	**Coefficient estimate**
Distance	-2.019***
	(0.021)
FPS open days	0.011***
	(0.001)
Average stock	0.000***
	(0.000)
Opening balance	0.000***
	(0.000)
Home FPS	2.426***
	(0.014)
Open market	-6.689***
	(0.319)
Pseudo R^2	0.570
Wald test (p-value)	0.000
Observations	1,502,464

Notes: 1) Results shown for the model described in equation (4.2) with robust standard errors. 2) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4.3: Heterogeneous preferences of households

	Differential preference on drivers of choice by		
	AAY households	**Female dominated households**	**Elderly households**
Distance	-2.059***	-2.003***	-2.037***
	(0.022)	(0.026)	(0.022)
FPS open days	0.011***	0.012***	0.012***
	(0.001)	(0.002)	(0.001)
Average stock	0.000***	0.000***	0.000***
	($1.5e^{-06}$)	($1.5e^{-06}$)	($1.5e^{-06}$)
Opening balance	0.000***	0.000***	0.000***
	($7.9e^{-07}$)	($8.8e^{-07}$)	($7.9e^{-07}$)
Home FPS	2.391***	2.436***	2.429***
	(0.014)	(0.017)	(0.015)
Open market	-6.581***	-6.673***	-6.956***
	(0.319)	(0.382)	(0.381)
Incremental differential preference on distance	1.021***	-0.104**	0.231*
	(0.112)	(0.049)	(0.138)
Incremental differential preference on FPS open days	-0.001	-0.001	-0.011
	(0.008)	(0.003)	(0.008)
Incremental differential preference on home FPS	0.951***	-0.021	0.201**
	(0.086)	(0.032)	(0.091)
Incremental differential preference on open market	-11.299***	-14.406***	-11.592***
	(0.365)	(0.387)	(0.433)
Pseudo R^2	0.570	0.571	0.574
Wald test (p-value)	0.000	0.000	0.000
Observations	1,501,193	1,439,672	1,478,482

Notes: 1) Results shown for the models described in equation (4.3) with robust standard errors. 2) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 3) The marginal effect of the differential preference is calculated as the linear combination of the marginal effects of the main effect and the interaction term. For instance, the marginal effect on distance for female dominated households is calculated as the sum of -18.15 and -1.03 which results in -19.1 indicating that a 0.1 km increase in distance to an FPS leads to a 19.1% decrease in likelihood of usage for female dominated households.

Table 4.4: Welfare and variability metrics - Pre and post portability

Measure	**Pre-portability**	**Post-portability**	**Percentage change**
Percentage PDS purchase	93	98	5.37
Average household utility	2.55	2.86	12.04
Percentage of instances of stock out at the most preferred FPS	0	5.96	-
Coefficient of demand variation	0.037	0.06	62.16

Notes: The table shows the values of the four metrics for both pre and post portability scenarios.

Table 4.5: Analysis of transactions in the presence of stock outs

		Transacted alternative				
		Home	**Most preferred non-home**	**Second/third preferred non-home**	**Open market**	**Total**
Preferred alternative	**Home**	99,128	3,513	531	0	103,172
	Non-home	3,606	29,838	699	0	34,143
	Open market	0	0	0	2,685	2,685
	Total	102,734	33,351	1,230	2,685	140,000

Notes: The table shows a break-up of the transactions with respect to their preferred alternative and transacted alternative.

Table 4.6: Analysis of utility loss in the presence of stock outs

Preferred FPS	**Transacted FPS**	**Utility loss in percentage**	**Percentage of transactions**
Home	Most preferred non-home	5.51	2.51
Home	Second/third preferred non-home	0.89	0.38
Non-home	Home	5.83	2.57
Most preferred non-home	Second/third preferred non-home	1.14	0.49
Total loss		**13.36**	**5.96**

Notes: The table shows the break-up of the utility loss from not being able to transact at the preferred FPS due to stock outs.

4.8 Appendix

4.8.1 Steps in construction of pre and post portability scenarios for simulation analysis

1. For generating the pre-portability scenario, retain only two choices - home FPS and the open market. Populate the values of covariates from empirical data for all months for all variables except availability related covariates (Average stock and Opening balance). Populate 'Day of purchase' from original data, for all months. For month, $t = 2$, populate the historical average stock from month $t = 1$.

2. Calculate monthly replenishment quantity for month $t = 2$ which arrives at the beginning of the month based on the defined current replenishment policy. The policy is similar to a base stock policy, where the order-up to level is equal to the gross requirement (GR) of the FPS. Gross requirement is calculated as the total entitlement quantity of all the households affiliated to the FPS, as defined before portability was introduced. This will represent the opening balance of first day of purchase at that FPS for month $t = 2$.

3. For month $t = 2$, for every FPS, arrange the choice instances of households in the ascending order of day of purchase. For the smallest value of day of purchase, use the populated covariate values, coefficient estimates from the multinomial logit model and the random disturbance (random numbers generated using Type 1 extreme value distribution with location parameter $\eta = 0$ and scale parameter $\mu = 1$) to calculate the utility for each household for all alternatives in the consideration set, using the indirect utility function. For instance, let the covariates $Distance_{h,f}$, $FPS\ open\ days_{f,t-1}$, $Average\ stock_{f,t-1}$ and $Opening\ balance_{f,t}$ take the values 1.22 kms, 12 days, 11,618 kgs and 6,385

kgs respectively. Let the alternative in question be a home FPS and $\epsilon_{h,f,t}$ take the value 1.086. In this case, based on the coefficient values ($\beta_1 = -2.019$, $\beta_2 = 0.011$, $\beta_3 = 0.000$, $\beta_4 = 0.000$, $\beta_5 = 2.426$ and $\beta_6 = -6.689$), the utility derived by the household from this alternative is 1.18 units.

4. Based on the maximum utility that the household derives from amongst all alternatives in the consideration set, indicate the most preferred alternative and predict the store choice for every household. If there is a stock out at the most preferred FPS, indicate an instance of stock out and make the household choose from the second preferred alternative. If the second preferred alternative is also stocked out, move to the next preferred alternative. Continue this, until the household makes a purchase from one of the FPSs in the PDS or finally the open market.

5. Aggregate the sales from all households for the smallest day of purchase for each FPS. Use this sales to calculate the FPS wise closing balance for that day (opening balance for next day).

6. Repeat steps 3, 4 and 5 for all the days of purchase in month $t = 2$. Finally calculate the average daily inventory for all days in month $t = 2$ for each FPS.

7. Repeat steps 2, 3, 4, 5 and 6 for all months. Calculate the average utility derived by all households - based on the utility derived from the choice made. Calculate the demand variance and coefficient of variation of demand for each FPS. Finally, calculate the percentage instances transactions from the PDS and percentage instances of stock outs at the most preferred alternative.

8. For generating the post-portability scenario, retain all alternatives in the consideration set and repeat the above procedure. Finally, compute and compare the values of average utility, coefficient of variation of demand,

instances of stock outs and percentage instances of PDS transactions between the pre and post portability scenarios.

4.8.2 Model validation

We carry out two types of model validation - (i) In-sample validation (using the sample used to estimate coefficients) to test how well the model fits the data that it has been trained on and (ii) Out-of-sample validation (cross validation using a different test sample) - using 'new' data which is not found in the dataset used to build the model. Table 4.7 shows the results of model validation. Columns (2) and (3) show the results for in-sample validation and compare the metrics from the choice model and simulation with the empirical data. Similarly, columns (4) and (5) show the results for out-of-sample validation and compare the metrics from the choice model and simulation with the empirical data. We find that the prediction accuracy is satisfactory and other metrics are comparable between the model and empirical data.

Table 4.7: Model validation

	In-sample validation		**Out-of-sample validation**	
Metric	**Model**	**Empirical data**	**Model**	**Empirical data**
Prediction accuracy (percentage correct predictions)	72.4	-	62.5	-
Percentage portability transactions	25	28	24.11	22.3
Percentage of instances of stock out in the most preferred FPS	5.96	-	5.2	-
Percentage of instances of stock out (FPS-month level)	69.2	74	56	55.2

References

Ackerman, D. S. and Gross, B. L. (2006). How many choices are good? measurement of the effects of course choice on perceptions of a marketing option. *Journal of Marketing Education*, 28(1):69–80.

Ahmed, A. U., Quisumbing, A. R., Nasreen, M., Hoddinott, J., and Bryan, E. (2009). Comparing food and cash transfers to the ultra poor in bangladesh. Technical report.

Ai, C. and Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics letters*, 80(1):123–129.

Ali, J. and Kumar, S. (2011). Information and communication technologies (icts) and farmers' decision-making across the agricultural supply chain. *International Journal of Information Management*, 31(2):149–159.

Ali, R. (2018). Nationwide ration portability soon. *Times of India*. [Online; accessed 13-March-2021].

Allu, R., Deo, S., and Devalkar, S. (2019). Alternatives to aadhaar-based biometrics in the public distribution system. *Economic & Political Weekly*, 54(12):31.

Ananda, Jonathan (2017). Food for nought: when your hand fails to feed you. [Online; accessed 05-June-2018].

Angrist, J. D. and Pischke, J.-S. (2008). Mostly harmless econometrics: An empiricist's companion. pages 221–246. Princeton university press.

Appleby, J. and Dixon, J. (2004). Patient choice in the nhs: Having choice may not improve health outcomes. *BMJ: British Medical Journal*, 329(7457):61.

Atalı, A., Lee, H., and Ozer, O. (2005). If the inventory manager knew: Value of rfid under imperfect inventory information. Technical report, Working paper, Dept. of Management Science and Engineering, Stanford ....

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, 28(25):3083–3107.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3):399–424.

Austin, P. C. (2013). The performance of different propensity score methods for estimating marginal hazard ratios. *Statistics in medicine*, 32(16):2837–2849.

Austin, P. C. and Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine*, 34(28):3661–3679.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics*, 21(1):1–42.

Azen, R. and Traxel, N. (2009). Using dominance analysis to determine predictor importance in logistic regression. *Journal of Educational and Behavioral Statistics*, 34(3):319–347.

Babich, V. and Tang, C. S. (2012). Managing opportunistic supplier product adulteration: Deferred payments, inspection, and combined mechanisms. *Manufacturing & Service Operations Management*, 14(2):301–314.

Bailey, S. and Hedlund, K. (2012). The impact of cash transfers on nutrition in emergency and transitional contexts: A review of evidence. *HPG Synthesis Paper. London: Overseas Development Institute.*

Bajaj, V. (2012). As Grain Piles Up, India's Poor Still Go Hungry. *The New York times.* [Online; accessed 10-November-2017].

Banerjee, A., Duflo, E., Imbert, C., Mathew, S., and Pande, R. (2014). Can e-governance reduce capture of public programs? experimental evidence from a financial reform of india's employment guarantee. *economics. mit. edu/files/10557 (accessed July 1, 2015).*

Banerjee, A., Hanna, R., Kyle, J., Olken, B. A., and Sumarto, S. (2018). Tangible information and citizen empowerment: Identification cards and food subsidy programs in indonesia. *Journal of Political Economy*, 126(2):451–491.

Banerjee, A., Hanna, R., Kyle, J., Olken, B. A., and Sumarto, S. (2019). Private outsourcing and competition: Subsidized food distribution in indonesia. *Journal of Political Economy*, 127(1):101–137.

Bank, W. (2003). *World development report 2004: making services work for poor people.* The World Bank.

Barnwal, P. (2016). Curbing Leakage in Public Programs with Direct Benefit Transfers. *The World Bank, Working paper.* [Online; accessed 10-November-2017].

Bayus, B. L. and Putsis Jr, W. P. (1999). Product proliferation: An empirical analysis of product line determinants and market outcomes. *Marketing Science*, 18(2):137–153.

Bell, D. R. and Lattin, J. M. (1998). Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer edlp. *Marketing Science*, 17(1):66–88.

Ben-Akiva, M. E., Lerman, S. R., and Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand*, volume 9. MIT press.

Bergolo, M. and Galván, E. (2018). Intra-household behavioral responses to cash transfer programs. evidence from a regression discontinuity design. *World Development*, 103:100–118.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Briesch, R. A., Chintagunta, P. K., and Fox, E. J. (2009). How does assortment affect grocery store choice? *Journal of Marketing research*, 46(2):176–189.

Briesch, R. A., Dillon, W. R., and Fox, E. J. (2013). Category positioning and store choice: The role of destination categories. *Marketing Science*, 32(3):488–509.

Brynjolfsson, E., Hu, Y., and Simester, D. (2011). Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science*, 57(8):1373–1386.

Business-Standard (2018). Im-pds approved for implementation during 2018-19 and 2019-20. [Online; accessed 12-September-2018].

Business Today (2016). Govt to computerise all ration shops by March 2017: Ram Vilas Paswan. [Online; accessed 10-November-2017].

Cachon, G. P. and Fisher, M. (2000). Supply chain inventory management and the value of shared information. *Management science*, 46(8):1032–1048.

Cachon, G. P., Terwiesch, C., and Xu, Y. (2008). On the effects of consumer search and firm entry in a multiproduct competitive market. *Marketing Science*, 27(3):461–473.

Cameron, A. C. and Trivedi, P. K. (2013). *Regression analysis of count data*, volume 53. Cambridge university press.

CCS (2018). Education for all: School choice campaign.

Chandra, N. (2014). India's PDS leaks leave poor hungry. *India today.* [Online; accessed 10-November-2017].

Chandukala, S. R., Kim, J., Otter, T., and Allenby, G. M. (2008). *Choice models in marketing: Economic assumptions, challenges and trends.* Now Publishers Inc.

Chatterjee, C. B. and Sheoran, G. (2007). *Vulnerable groups in India.* Centre for Enquiry into Health and Allied Themes Mumbai, India.

Chernev, A. and Hamilton, R. (2009). Assortment size and option attractiveness in consumer choice among retailers. *Journal of Marketing Research*, 46(3):410–420.

Cho, S.-H., Fang, X., Tayur, S., and Xu, Y. (2019). Combating child labor: Incentives and information disclosure in global supply chains. *Manufacturing & Service Operations Management.*

Chowdhury, G. and Koya, K. (2017). Information practices for sustainability: Role of ischools in achieving the un sustainable development goals (sdgs). *Journal of the Association for Information Science and Technology*, 68(9):2128–2138.

Clarke, J., Newman, J., and Westmarland, L. (2008). The antagonisms of choice: New labour and the reform of public services. *Social Policy and Society*, 7(2):245–253.

CMS (2005). India corruption study. *Transparency International India.*

Commission, P. et al. (2015). Performance evaluation of targeted public distribution system (tpds). Technical report.

Correspondent (2016). Pds portability lets card holders avail rations from any outlet. *Deccan Herald*.

Correspondent (2018). Ration dealers will be paid increased commission: Eatala.

Cui, R., Allon, G., Bassamboo, A., and Van Mieghem, J. A. (2015). Information sharing in supply chains: An empirical and theoretical valuation. *Management Science*, 61(11):2803–2824.

Currie, J. and Gahvari, F. (2008). Transfers in cash and in-kind: Theory meets the data. *Journal of Economic Literature*, 46(2):333–83.

Damera, V. K. (2017). Choice for the poor or poor choice? experimental evidence from implementation of india's school choice policy. *BSG Working Paper Series*.

Dash, Dipak, K. (2015). Govt to check urea diversion to industry. *The Times of India*. [Online; accessed 10-November-2017].

Davis, J. (2004). Corruption in public service delivery: experience from south asia's water and sanitation sector. *World development*, 32(1):53–71.

Davis, P. (2006). Spatial competition in retail markets: movie theaters. *The RAND Journal of Economics*, 37(4):964–982.

Del Ninno, C., Dorosh, P. A., and Subbarao, K. (2007). Food aid, domestic policy and food security: Contrasting experiences from south asia and sub-saharan africa. *Food Policy*, 32(4):413–435.

Demirguc-Kunt, A. and Klapper, L. (2012). *Measuring financial inclusion: The global findex database*. The World Bank.

Dennis, C., Marsland, D., and Cockett, W. (1999). Objects of desire: attraction and distance in shopping centre choice. In *Proceeding of the 10th International*

Conference on Research of Distributive Trades, Institute for Retail Studies, Sterling. USA.

Department of Food and Public Distribution, India (2013). Coverage and entitlement under NFSA. [Online; accessed 10-November-2017].

Desai, R. M. and Olofsgård, A. (2019). Can the poor organize? public goods and self-help groups in rural india. *World Development*, 121:33–52.

Dhanaraj, S. and Gade, S. (2012). Universal pds: efficiency and equity dimensions. *Available at SSRN 2020130*.

Dreze, J. and Khera, R. (2015). Understanding leakages in the public distribution system. *Economic & Political Weekly*, 50(7):39–42.

Ekirapa-Kiracho, E., Waiswa, P., Rahman, M. H., Makumbi, F., Kiwanuka, N., Okui, O., Rutebemberwa, E., Bua, J., Mutebi, A., Nalwadda, G., et al. (2011). Increasing access to institutional deliveries using demand and supply side incentives: early results from a quasi-experimental study. *BMC international health and human rights*, 11(1):1–14.

El Zarwi, F., Vij, A., and Walker, J. L. (2017). A discrete choice framework for modeling and forecasting the adoption and diffusion of new transportation services. *Transportation Research Part C: Emerging Technologies*, 79:207–223.

Elbahnasawy, N. G. (2014). E-government, internet adoption, and corruption: an empirical investigation. *World Development*, 57:114–126.

Express, I. (2018). Krishna anganwadi centres in sorry state of affairs, reveals vigilance raids.

Fan, T., Tao, F., Deng, S., and Li, S. (2015). Impact of rfid technology on supply chain decisions with inventory inaccuracies. *International Journal of Production Economics*, 159:117–125.

FAPDA (2019). Food and agriculture policy decision analysis tool . *FAO*. [Online; accessed 13-March-2021].

Ferlie, E., Freeman, G., McDonnell, J., Petsoulas, C., and Rundle-Smith, S. (2006). Introducing choice in the public services: some supply-side issues. *Public Money and Management*, 26(1):63–72.

Fernández, M. J. G. (2015). *Improving Food Access for Poor Households in Indonesia: Cash Transfers and the Raskin Program Reform*. PhD thesis, Harvard University.

Fischer, R., Gonzalez, P., and Serra, P. (2006). Does competition in privatized social services work? the chilean experience. *World Development*, 34(4):647–664.

Fisher, M. L. and Ittner, C. D. (1999). The impact of product variety on automobile assembly operations: Empirical evidence and simulation analysis. *Management science*, 45(6):771–786.

Fotaki, M. and Boyd, A. (2005). From plan to market: a comparison of health and old age care policies in the uk and sweden. *Public Money & Management*, 25(4):237–243.

Fotaki, M., Boyd, A., Smith, L., McDonald, R., Roland, M., Sheaff, R., Edwards, A., and Elwyn, G. (2005). Patient choice and the organisation and delivery of health services: scoping review. In *Workshop 1 15th March*, volume 8.

Funk, M. J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M. A., and Davidian, M. (2011). Doubly robust estimation of causal effects. *American Journal of Epidemiology*, 173(7):761–767.

Ganesh, M., Deo, S., and Devalkar, S. K. (2019). Leveraging digital technology to improve monitoring and planning in public sector supply chains: Evidence from india's food security program. *Indian School of Business*.

Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S., and Aldridge, M. D. (2014). Methods for constructing and assessing propensity scores. *Health Services Research*, 49(5):1701–1720.

Gaur, V., Giloni, A., and Seshadri, S. (2005). Information sharing in a supply chain under arma demand. *Management science*, 51(6):961–969.

Gentilini, U. (2007). *Cash and food transfers: A primer*. World Food Programme Rome.

Gentilini, U., Honorati, M., and Yemtsov, R. (2014). The state of social safety nets 2014. *The World Bank*. [Online; accessed 13-March-2021].

Gentilini, U. and Omamo, S. W. (2011). Social protection 2.0: Exploring issues, evidence and debates in a globalizing world. *Food policy*, 36(3):329–340.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., and Vermeersch, C. M. (2016). *Impact evaluation in practice*. The World Bank.

Goldfarb, A. and Tucker, C. E. (2014). Conducting research with quasi-experiments: A guide for marketers. *Rotman School of Management Working Paper*, (2420920).

Gössling, S. and Michael Hall, C. (2019). Sharing versus collaborative economy: how to align ict developments and the sdgs in tourism? *Journal of Sustainable Tourism*, 27(1):74–96.

Government of India (2015). Recommendations of High Level Committee on restructuring of FCI. *Ministry of Consumer Affairs, Food & Public Distribution*. [Online; accessed 10-November-2017].

Government of Karnataka (2015). FPS/WSD Statistics . *Food Civil Supplies and Consumer Affairs Department, Government of Karnataka*. [Online; accessed 10-November-2017].

Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician*, 61(2):139–147.

Guadagni, P. M. and Little, J. D. (1983). A logit model of brand choice calibrated on scanner data. *Marketing science*, 2(3):203–238.

Gulati, A. and Saini, S. (2015). Leakages from public distribution system (PDS) and the way forward. *ICRIER, Working paper*. [Online; accessed 10-November-2017].

Hamaguchi, Y., Kawagoe, T., and Shibata, A. (2009). Group size effects on cartel formation and the enforcement power of leniency programs. *International Journal of Industrial Organization*, 27(2):145–165.

Harvey, P. and Savage, K. (2006). No small change. oxfam gb malawi and zambia emergency cash transfer projects: a synthesis of key learning. *ODI. London*.

Heeks, R. (2001). Understanding e-governance for development.

Heeks, R. (2002). Failure, success and improvisation of information systems projects in developing countries. *Development Informatics Working Paper*, (11).

Hidrobo, M., Hoddinott, J., Peterman, A., Margolies, A., and Moreira, V. (2014). Cash, food, or vouchers? evidence from a randomized experiment in northern ecuador. *Journal of development Economics*, 107:144–156.

Hindu, T. (2018). Ration portability in Telangana from today. *The Hindu*. [Online; accessed 13-March-2021].

Hindu, T. (2019). State portability under PDS to be strengthened'. *The Hindu*. [Online; accessed 13-March-2021].

Hosmer Jr, D. W., Lemeshow, S., and Sturdivant, R. X. (2013). *Applied logistic regression*, volume 398. John Wiley & Sons.

Huang, L., Song, J.-S. J., and Swinney, R. (2017). Managing social responsibility in multitier supply chains. *SSRN, Working paper.* [Online; accessed 10-November-2017].

Imbens, G. and Wooldridge, J. (2007). Difference-in-differences estimation. *National Bureau of Economics Research.* [Online; accessed 10-November-2017].

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

Jain, N. and Tan, T. (2021). M-commerce, sales concentration, and inventory management. *SMU Cox School of Business Research Paper Forthcoming.*

Jayasankara Prasad, C. and Ramachandra Aryasri, A. (2011). Effect of shopper attributes on retail format choice behaviour for food and grocery retailing in india. *International Journal of Retail & Distribution Management*, 39(1):68–86.

Jebaraj, P. (2019). Graft shadow over welfare scheme. *The Hindu.*

Joshi, A., Sinha, D., and Patnaik, B. (2016). Lessons from core pds reforms in chhattisigarh: Credibility and portability? *Economic and Political Weekly*, 51:51–59.

Kabra, A., Belavina, E., and Girotra, K. (2020). Bike-share systems: Accessibility and availability. *Management Science*, 66(9):3803–3824.

Kanda, M. (2018). Health, education need govt patronage.

Kaushik, P. and Singh, N. (2004). Information technology and broad-based development: preliminary lessons from north india. *World Development*, 32(4):591–607.

Keane, M. P. and Wolpin, K. I. (2007). Exploring the usefulness of a nonrandom holdout sample for model validation: Welfare effects on female behavior. *International Economic Review*, 48(4):1351–1378.

Kennedy, P. (2003). *A guide to econometrics*. MIT press.

Khera, R. (2011). Trends in diversion of grain from the public distribution system. *Economic and Political Weekly*, pages 106–114.

Khera, R. (2014). Cash vs. in-kind transfers: Indian data meets theory. *Food Policy*, 46:116–128.

Koul, S. and Mishra, H. G. (2013). Customer perceptions for store attributes: A study of traditional retail stores in india. *Journal of Business & Economics*, 5(1):79.

Krishna, A. and Schober, G. (2014). The gradient of governance: distance and disengagement in indian villages. *Journal of Development Studies*, 50(6):820–838.

Kulkarni, D. (2014). Food grains, kerosene not reaching beneficiaries as they should. *Daily News Service*. [Online; accessed 10-November-2017].

Kumar, N., Raghunathan, K., Arrieta, A., Jilani, A., Chakrabarti, S., Menon, P., and Quisumbing, A. R. (2019). Social networks, mobility, and political participation: The potential for women's self-help groups to improve access and use of public entitlement schemes in india. *World Development*, 114:28–41.

Kumar, R. and Best, M. L. (2006). Impact and sustainability of e-government services in developing countries: Lessons learned from tamil nadu, india. *The Information Society*, 22(1):1–12.

Lakhani, Somya (2018). Devices at ration shops in Delhi fail to recognise fingerprints, govt says no one will be turned away. [Online; accessed 05-June-2018].

Lal, N. (2015). Poor Bear the Brunt of Corruption in India's Food Distribution System. *Inter Press Service*. [Online; accessed 10-November-2017].

Le Grand, J. (2009). *The other invisible hand: Delivering public services through choice and competition*. Princeton University Press.

Lechner, M. et al. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, 4(3):165–224.

Lee, H. L., Padmanabhan, V., and Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management science*, 43(4):546–558.

Lee, H. L., So, K. C., and Tang, C. S. (2000). The value of information sharing in a two-level supply chain. *Management science*, 46(5):626–643.

Lee, Y. M., Cheng, F., and Leung, Y. T. (2004). Exploring the impact of rfid on supply chain dynamics. In *Proceedings of the 2004 Winter Simulation Conference, 2004.*, volume 2, pages 1145–1152. IEEE.

Lentz, E., Ouma, R., and Mude, A. (2016). Does peer monitoring influence choices between cash and food? findings from a field experiment. *Food Policy*, 65:21–31.

Levi, R., Singhvi, S., and Zheng, Y. (2018). Economically motivated adulteration in farming supply chains. *Forthcoming in Management Science*. [Online; accessed 10-November-2017].

Library Congress (2016). Constitutional right to an education: India.

Lusk, J. L. and Weaver, A. (2017). An experiment on cash and in-kind transfers with application to food assistance programs. *Food Policy*, 68:186–192.

Mahajan, S. and Van Ryzin, G. (2001). Inventory competition under dynamic consumer choice. *Operations Research*, 49(5):646–657.

Masiero, S. (2015). Redesigning the indian food security system through e-governance: The case of kerala. *World Development*, 67:126–137.

Masiero, S. (2016). The origins of failure: seeking the causes of design–reality gaps. *Information Technology for Development*, 22(3):487–502.

Masiero, S. (2017). Digital governance and the reconstruction of the indian anti-poverty system. *Oxford Development Studies*, 45(4):393–408.

Mayer, K. J., Nickerson, J. A., and Owan, H. (2004). Are supply and plant inspections complements or substitutes? a strategic and operational assessment of inspection practices in biotechnology. *Management Science*, 50(8):1064–1081.

Michelson, H., Lentz, E. C., Mulwa, R., Morey, M., Cramer, L., McGlinchy, M., and Barrett, C. B. (2012). Cash, food, or vouchers? an application of the market information and food insecurity response analysis framework in urban and rural kenya. *Food Security*, 4(3):455–469.

Mishra, P. (2018). Ration cards: Government may soon launch inter-state mobility. *Financial express*.

Morton, G. D. (2019). The power of lump sums: Using maternity payment schedules to reduce the gender asset gap in households reached by brazil's bolsa família conditional cash transfer. *World Development*, 113:352–367.

Mu, L., Dawande, M., Geng, X., and Mookerjee, V. (2016). Milking the quality test: Improving the milk supply chain under competing collection intermediaries. *Management Science*, 62(5):1259–1277.

Muralidharan, K., Niehaus, P., and Sukhtankar, S. (2016). Building state capacity: Evidence from biometric smartcards in india. *The American Economic Review*, 106(10):2895–2929.

Nagavarapu, S. and Sekhri, S. (2016). Informal monitoring and enforcement mechanisms in public service delivery: Evidence from the public distribution system in india. *Journal of Development Economics*, 121:63–78.

Naik, G., Chitre, C., Bhalla, M., and Rajan, J. (2020). Impact of use of technology on student learning outcomes: Evidence from a large-scale experiment in india. *World Development*, 127:104736.

NCAER (2015). Evaluation study of targeted public distribution system in selected states. [Online; accessed 10-November-2017].

Newman, K. and Thorat, S. K. (2010). *Blocked by caste: Economic discrimination in modern India*. Oxford University Press.

NIC (2018). [Online; accessed 12-September-2018].

OFT (2010). Choice and Competition in Public Services . *The Office of Fair Trading and Frontier Economics*. [Online; accessed 13-March-2021].

Olken, B. A. (2006). Corruption and the costs of redistribution: Micro evidence from indonesia. *Journal of public economics*, 90(4-5):853–870.

O'Neill, S., Kreif, N., Grieve, R., Sutton, M., and Sekhon, J. S. (2016). Estimating causal effects: considering three alternatives to difference-in-differences estimation. *Health Services and Outcomes Research Methodology*, 16(1-2):1–21.

Pingali, P., Aiyar, A., Abraham, M., and Rahman, A. (2019). *Transforming food systems for a rising India*. Springer Nature.

Pingali, P., Mittra, B., and Rahman, A. (2017). The bumpy road from food to nutrition security–slow evolution of india's food policy. *Global food security*, 15:77–84.

Pischke, J. (2005). Empirical methods in applied economics: Lecture notes. *LSE*. [Online; accessed 10-November-2017].

Plambeck, E. L. and Taylor, T. A. (2015). Supplier evasion of a buyer's audit: Implications for motivating supplier social and environmental responsibility. *Manufacturing & Service Operations Management*, 18(2):184–197.

Pradhan, M. and Rao, N. (2018). Gender justice and food security: The case of public distribution system in india. *Progress in Development Studies*, 18(4):252–266.

Pradhan, M., Roy, D., and Sonkar, V. (2015). What is the effective delivery mechanism of food support in india. *Economic and Political Weekly*, 54(42):7–8.

Prasad, C. J. (2010). Effect of consumer demographic attributes on store choice behaviour in food and grocery retailing-an empirical analysis. *Management and Labour Studies*, 35(1):35–58.

Raghunathan, S. (2001). Information sharing in a supply chain: A note on its value when demand is nonstationary. *Management science*, 47(4):605–610.

Rajan, P., Chopra, S., Somasekhar, A., and Laux, C. (2016). Designing for food security: Portability and the expansion of user freedoms through the corepds in chhattisgarh, india. *Information Technologies & International Development*, 12(3):pp–1.

Raju, S., Rampal, P., Bhavani, R., Rajshekar, S., et al. (2018). Introduction of millets into the public distribution system: Lessons from karnataka. *Journal*, 8(2):120–136.

Ramakrishnan (2017). You may soon be able to buy from the fair price shop of your choice. *The Hindu*.

Ray, S. and Mukherjee, A. (2007). Development of a framework towards successful implementation of e-governance initiatives in health sector in india. *International journal of health care quality assurance*.

Rekik, Y. (2010). Inventory inaccuracies in supply chains: How can rfid improve the performance? *Wiley Encyclopedia of Operations Research and Management Science*.

Rosamma, Thomas (2016). Machine fails to read fingerprints, 1.4cr Rajasthanis go without ration. [Online; accessed 05-June-2018].

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Rozindar, F. (2016). BPL shopping bag has rice, wheat...but why detergent? [Online; accessed 05-June-2018].

Rui, H. and Lai, G. (2015). Sourcing with deferred payment and inspection under supplier product adulteration risk. *Production and Operations Management*, 24(6):934–946.

Sabates-Wheeler, R. and Devereux, S. (2010). Cash transfers and high food prices: Explaining outcomes on ethiopia's productive safety net programme. *Food Policy*, 35(4):274–285.

Sabharwal, N. S. (2011). Caste, religion and malnutrition linkages. *Economic and Political Weekly*, pages 16–18.

Sachdev, V. (2018). Aadhaar authentication for Govt services fails 12 *The Quint*. [Online; accessed 17-March-2019].

Saha, D. (2015). What 13 million fake cooking-gas consumers reveal. *Business Standard*. [Online; accessed 10-November-2017].

Sahas, J. (2009). Exclusion and Inclusion of Dalit Community in Education and Health - A Study. *Jan Sahas Social Development Society*. [Online; accessed 13-March-2021].

Salunke, P. (2015). Maharashtra: Welfare schemes victims of graft. *Hindustan Times*.

Sargar, R., Kumar, A., Nakade, V., and Borkar, N. (2014). Public distribution system in solapur district of maharashtra: A case study. *International Journal of Research in Engineering & Advanced Technology*, 2(3):8.

Sati, V. P. (2015). Issues and options of food security and poverty: an empirical study of mizoram, the eastern extension of the himalaya. *Journal of Food Security*, 3(4):107–114.

Saxena, K. B. C. (2005). Towards excellence in e-governance. *International Journal of Public Sector Management*.

Sen, A. (2012). Chhatisgarh goes for smart card solution to plug pds leaks. *The Economic Times*. [Online; accessed 10-November-2017].

Sharma, N. and Gupta, S. (2019). An investigation of it-intervention adoption in public distribution system: A stakeholder and agency theory perspective. *Information Development*, 35(2):203–219.

Shoag, D. et al. (2010). The impact of government spending shocks: Evidence on the multiplier from state pension plan returns. *unpublished paper, Harvard University*.

Sinha, P. K., Banerjee, A., and Uniyal, D. P. (2002). Deciding where to buy: Store choice behaviour of indian shoppers. *Vikalpa*, 27(2):13–28.

Słoczyński, T. and Wooldridge, J. M. (2018). A general double robustness result for estimating average treatment effects. *Econometric Theory*, 34(1):112–133.

Smith, A. D. (2014). Case studies of rfid practices for competitive inventory management systems. In *Management Science, Logistics, and Operations Research*, pages 1–25. IGI Global.

Special correspondent (2014). A.p. moots ration card portability on lines of chhattisgarh. *The Hindu*.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 25(1):1.

Swait, J. and Erdem, T. (2002). The effects of temporal consistency of sales promotions and availability on consumer choice behavior. *Journal of Marketing Research*, 39(3):304–320.

Tan, T. F., Netessine, S., and Hitt, L. (2017). Is tom cruise threatened? an empirical study of the impact of product variety on demand concentration. *Information Systems Research*, 28(3):643–660.

The Hindu (2015). Funds for malnourished anganwadi children diverted. [Online; accessed 10-November-2017].

Thompson, D. V., Hamilton, R. W., and Rust, R. T. (2005). Feature fatigue: When product capabilities become too much of a good thing. *Journal of marketing research*, 42(4):431–442.

Tilakaratna, G. and Sooriyamudali, C. (2017). Food-based social assistance programs in sri lanka: Evolution and transition to cash transfers.

Tjoa, A. M. and Tjoa, S. (2016). The role of ict to achieve the un sustainable development goals (sdg). In *IFIP World Information Technology Forum*, pages 3–13. Springer.

Today, I. (2019). Implement one nation, one ration card system in 1 year without fail: Centre to states, UTs. *India Today*. [Online; accessed 13-March-2021].

Ton, Z. and Raman, A. (2010). The effect of product variety and inventory levels on retail store sales: A longitudinal study. *Production and Operations Management*, 19(5):546–560.

Tonidandel, S. and LeBreton, J. M. (2010). Determining the relative importance of predictors in logistic regression: An extension of relative weight analysis. *Organizational Research Methods*, 13(4):767–781.

Torkelson, E. (2020). Collateral damages: Cash transfer and debt transfer in south africa. *World Development*, 126:104711.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Vaidya, A., Somasekhar, A., and NIC, C. (2014). Core pds: Empowering with portability. *National Institute for Smart Government. Retreived from: http://nisg.org/files/documents/UP1418303314. pdf*.

Vian, T., Savedoff, W., and Mathisen, H. (2009). Preventing drug diversion through supply chain management. Kumarian press.

Vij, A. and Krueger, R. (2017). Random taste heterogeneity in discrete choice models: Flexible nonparametric finite mixture distributions. *Transportation Research Part B: Methodological*, 106:76–101.

Wan, X., Evers, P. T., and Dresner, M. E. (2012). Too much of a good thing: The impact of product variety on operations and sales performance. *Journal of Operations Management*, 30(4):316–324.

Wang, Y. and Bell, D. R. (2015). Consumer store choice in asian markets. *Marketing Letters*, 26(3):293–308.

Xia, N. and Rajagopalan, S. (2009). Standard vs. custom products: Variety, lead time, and price competition. *Marketing science*, 28(5):887–900.

Yadav, Anumeha (2016). Rajasthan presses on with Aadhaar after fingerprint readers fail: We'll buy iris scanners. [Online; accessed 05-June-2018].

Zentner, A., Smith, M., and Kaya, C. (2013). How video rental patterns change as consumers move online. *Management Science*, 59(11):2622–2634.