A. Appendix

This is an Online Appendix to the paper "Technology enabled agent-choice and uptake of social assistance programs: Evidence from India's food security program". It contains nine sections providing supplementary analysis as referenced in the paper.

A.1. Summary statistics and potential outliers

In this section, we report more elaborate summary statistics in Table A1 and re-estimate (1) to test if our results are solely driven by the presence of outliers.

Table A1 Summary Statistics on key operational variables of interest

	AP (Control)				TS (Treatment)			
	Mean	Min	Median	Max	Mean	Min	Median	Max
Uptake (kg)	289,207	160	241,712	2,735,150	296,168	41,624	226,430	4,156,065
Quantity allocated (kg)	295,492	2,939	247,702	2,752,024	330,470	73,665	253,342	4,062,120
Number of choice users	-	-	-	-	2,103	0	687	88,852

Uptake of 160 kg in a sub-district month in AP is likely to be an outlier as the corresponding 1^{st} and the 5^{th} percentiles are 13,040.5 kg and 69,533.9 kg, respectively. To test if our results are solely driven by this value, we take two approaches to re-estimate our main model: (1) exclude observations falling below the 1^{st} percentile and above the 99^{th} percentile and, (2) winsorise the observations at 1 % levels on both tails of the data (Cox 2006). Table A2 and A3 show the results for the dependent variables as uptake and uptake as a ratio of quantity allocated, respectively. Our results continue to be positive and significant, but slightly lower in magnitude, 17,812.63 kg and 18,286.55 kg from exclusion and winsorization, respectively vis-ã-vis 20,273 kg. Corresponding values for ratio of allocated grains collected are 0.042 and 0.047, respectively vis-ã-vis 0.053. The lower magnitudes are likely because sub-district-months greater than 99^{th} percentile in uptake correspond to large urban centers with high population density where the impact of agent-choice is almost four times higher, and our outlier treatments either omit or subdue these observations. Moreover, except for the value 160 kg, we find no contextual reason to believe that the other values of uptake correspond to outliers.

Table A2 Outlier analysis with Uptake as DV

	(1) Base model	(2) Excluding outliers	(3) Winsorizing outliers
Impact of agent-choice (δ)	20,273.82*** (1,975.32)	17,812.63*** (1,355.99)	18,286.55*** (1,408.06)
Impact of quantity allocated (θ_1)	0.94*** (0.01)	0.938*** (0.01)	0.930*** (0.008)
Impact of strike (θ_2)	-11,362.66*** (1349.59)	-10,452*** (1,073.45)	-11,118.58*** (1,143.68)
Adjusted R^2	0.898	0.911	0.926
Observations	28,232	27,541	28,232

Notes: **1)** Column (1) shows the base model results from estimating equation (1). **2)** Column (2) shows the results from estimating equation (1) for a sub-sample by excluding the observations falling below the 1st percentile and above the 99th percentile of the data. **3)** Column (3) shows the results from estimating equation (1) where the observations at 1 % levels on both tails of the data are winsorized. **4)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3 Outlier analysis with $\left(\frac{UPTK_{it}}{QA_{it}}\right)$ as DV

	(1) Base model	(2) Excluding outliers	(3) Winsorizing outliers
Impact of agent-choice (δ)	0.053*** (0.003)	0.042*** (0.003)	0.047*** (0.003)
Impact of strike (θ_2)	-0.027*** (0.003)	-0.024*** (0.003)	-0.025*** (0.003)
Adjusted R^2	0.19	0.28	0.27
Observations	28,232	27,541	28,232

Notes: **1)** Column (1) shows the base model estimating equation (1) with the dependent variable $\left(\frac{UPTK_{it}}{QA_{it}}\right)$. **2)** Column (2) shows the results from estimating equation (1) for a sub-sample by excluding the observations falling below the 1st percentile and above the 99th percentile of the data. **3)** Column (3) shows the results from estimating equation (1) where the observations at 1 % levels on both tails of the data are winsorized. **4)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A.2. Discussion on stockouts

In a national survey, 10% of agents reported at least one instance of stockout in the past three months due to increased demand variability induced by agent-choice (Dalberg 2022). Consequently, it is possible that agent-choice has a negative impact on uptake mediated through increase in the number of stockouts at agents. In this section, we explore this possibility using data on closing balance of grains at each agent by the 15^{th} of each month from our control state for a period of six months. We consider agents with closing balance less than 35 kgs as stocked out since the agent does not have adequate inventory to serve one additional beneficiary.

Table A4 Analysis of stockouts

Month	(1) Number of stocked out agents (% of all agents)	(2) Number of stocked out agents with a stocked agent < 1 km distance (% of agents in Column (1))	(3) Average distance (in km) of the nearest stocked agent < 1 km distance
1	3,258 (13%)	3,024 (93%)	0.309
2	2,589 (10%)	2,434 (94%)	0.301
3	3,123 (12%)	2,928 (91%)	0.299
4	3,659 (15%)	3,321 (91%)	0.316
5	3,605 (14%)	3,296 (91%)	0.319
6	4,540 (18%)	3,959 (87%)	0.334

Notes: **1)** We consider agents with closing balance < 35 kgs as stocked out agents.**3)** Closing balances of 25,114 agents in the state are considered in six months. **3)** Haversine formula is used compute the distance.

In Columns (1) and (2) of Table A4, we show the number of agents that experienced a stockout in each month and the number of agents that have a nearby agent with stock within one kilometer distance from the focal agent, respectively. Evidently, over 90% of the agents that witnessed a stockout have another agent with stock where the beneficiaries could have collected their entitlements. Moreover, in Column (3), we show that the average distance of the nearest agent with stock from the focal agent is only about 300 meters. Since the nearest agent with stock is in such close proximity, we believe that beneficiaries witnessing a stockout at an agent may have collected their grains from a nearby agent. In fact, we learned from our interactions with the beneficiaries that they do make an attempt to visit other agents in the vicinity when their pre-assigned agent runs out of stock. Consequently, we believe that instances of stockouts at an agent level are unlikely to change total uptake at the sub-district level.

However, there is a possibility of beneficiaries not being able to collect entitlements due to instances of stockouts for reasons such as lack of information or collusion among agents. Therefore, our results represent the overall impact of agent-choice despite the negative effect of stockouts.

A.3. Beneficiary level transaction dataset

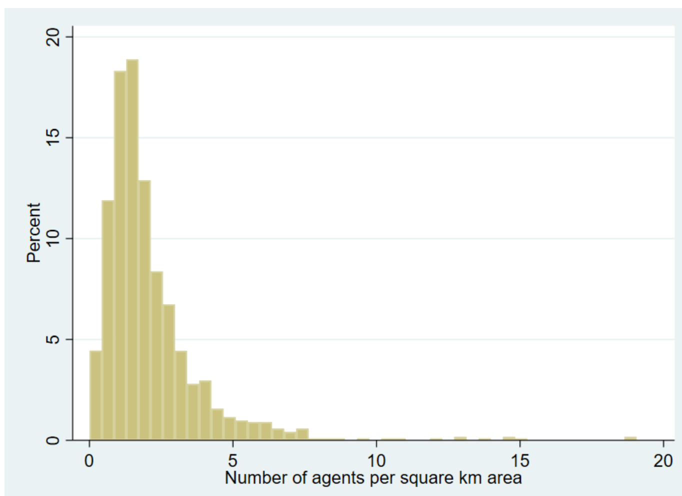
The data used by Ganesh et al. (2022) contains $500,000$ randomly chosen transactions made by beneficiaries in the state of Andhra Pradesh for a period of six months (March 2018 to August 2018). The data contains transaction ID which is unique for every transaction, beneficiary ID which is unique for every beneficiary, shop number where the transaction occurred, date of the transaction, quantity of food grains collected and the quantity of food grains entitled. Upon comparing the quantity of grains collected and the quantity allocated, we find that 99.98% of the transactions recorded beneficiaries collecting their full entitlement.

A.4. Role of agent density

In this section, we generate results in Table 3 with two alternate definitions of agent density: agent density defined as a continuous variable and agent density defined as urban regions.

Figure A1 shows the distribution of agent density.

Figure A1 Distribution of number of agents per square km area



Notes: Distribution of number of agents per square km area.

Table A5 shows results of estimating (2) using density as a continuous variables. All our results shown in Table 3 hold.

Next, estimate (2) using an alternate definition of DEN_i as sub-districts identified as urban in the District Census dataset published in 2011. We define URB_i as a binary variable which takes the value 1 if a sub-district is identified as "urban" in Census data and 0 otherwise. We estimate the following triple difference model using these measures to test our hypothesis:

$$UPTK_{it} = \alpha_i + \beta_t + \delta \;\; CHC_{it} + \phi_u \;\; CHC_{it} * URB_i + \theta_1 QA_{it} + \theta_2 STRK_{it} + \gamma D_{st} + \epsilon_{it} \quad (1)$$

where, δ is the impact of providing choice in non-urban sub-districts and $\delta + \phi_u$ is impact in urban sub-districts. We expect $\phi > 0$ and significant.

Column (3) in Table A6 shows the result of this estimation. We find that the impact of choice in non-urban regions is $16,767$ kgs (δ) and in urban regions is $32,569$ Kgs ($\delta + \phi_u$).

Table A5 Heterogeneity: Differential impact of Agent-Choice by agent density (continuous model)

	(1) Base model	(2) Agent density
Overall impact of agent-choice (δ)	20,273*** (1,975)	15,950*** (2,469)
Interaction of choice and density (ϕ)		2,069*** (630)
Impact of quantity allocated (θ_1)	0.94*** (0.01)	0.93*** (0.01)
Impact of strike (θ_2)	-11362.66*** (1349.59)	-11,390*** (1,350)
Adjusted R^2	0.898	0.903
Observations	28,232	28,180

Notes: **1)** Results shown are for Sub-district-month panel regressions with Sub-district and month fixed effects, Sub-district time dummies and standard errors clustered at the sub-district level. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** Column (1) shows the overall impact (results from estimating (1) with the dependent variable Uptake). **4)** Column (2) shows the results from estimating (2) with the dependent variable Uptake and agent density as a continuous variable. For every additional shop in a square km area, the impact on increase in uptake is 2,069 kgs per sub-district per month.

Table A6 Heterogeneity: Differential impact of Agent-Choice (Urban areas)

	(1) Base model	(2) Agent density	(3) Location
Overall impact of agent-choice (δ)	20,273*** (1,975)	16,295*** (2,178)	16,767*** (2,248)
Incremental impact in high-density areas (ϕ)		43,734*** (12,399)	
Incremental impact in urban areas (ϕ_u)			15,802*** (5,334)
Impact of quantity allocated (θ_1)	0.94*** (0.01)	0.94*** (0.01)	0.94*** (0.01)
Impact of strike (θ_2)	-11362.66*** (1349.59)	-11,392*** (1,350)	-11,364*** (1,350)
Adjusted R^2	0.898	0.901	0.899
Observations	28,232	28,180	28,232

Notes: **1)** We identify sub-districts with large concentration of urban centers from Census of India. We define URB_i as a binary variable which takes the value 1 if a sub-district is identified as "urban" in Census data and 0 otherwise. **2)** Results shown are for Sub-district-month panel regressions with Sub-district and month fixed effects, Sub-district time dummies and standard errors clustered at the sub-district level. **3)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **4)** Column (1) shows the overall impact (results from estimating (1) with the dependent variable Uptake). **5)** Column (2) shows the incremental impact in high density areas (results from estimating (1) with the dependent variable Uptake). **6)** Column (3) shows the incremental impact in urban areas (results from estimating (1) with the dependent variable Uptake).

A.5. Ordinal Specifications

In this section, we replicate our analyses in Section 5.4 by replacing the average number of days agents within a distance range keep their shops open with the number of days the n^{th} nearest agent keeps the shop open.

From the shop open days dataset, we identify the nine nearest neighbors of each agent within a haversine distance of seven kilometers. We restrict our consideration to seven kilometers as 80% of choice users choose an agent within this distance from their pre-assigned agents (Ganesh et al. 2022). Consequently, agents beyond this distance threshold are less likely to influence the number of days the focal agent keeps the shop open. In our control state, $\sim 95\%$ of the agents have at least nine shops within seven kilometers haversine radius. We conduct our analysis only on these agents.[1] Our analysis data contains $25,114$ agents and $149,157$ agent-month observations.

Let s and t denote subscripts for shop and month, respectively. Let $OPENDAYS_{st}$ denote the number of days shop s is kept open in month t, $OPENDAYS_{s_nt}$ denote the number of days the n^{th} nearest neighbor of shop s is kept open in month t, n takes the values one to five. Let $UPTK_{st}$ denote uptake of grains at shop s in month t. Table A7 shows summary statistics of key variables in the data. During our study period, the average monthly uptake at an agent is $7,013$ kg and the average number of shop open days each month is 12.6 days.

Table A7 Summary statistics

	(1) Mean	(2) Min	(3) P25	(4) P50	(5) P75	(6) Max
Uptake (kg), $UPTK_{st}$	7,013	0	4,534	6,545	8,845	42,940
Number of days open (days), $OPENDAYS_{st}$	12.6	1	11	14	15	19
Distance of the neighbors (in km)						
1^{st} nearest neighbor	0.683	0	0.003	0.237	1.157	5.584
2^{nd} nearest neighbor	1.214	0	0.278	1.073	1.950	5.739
3^{rd} nearest neighbor	1.648	0	0.514	1.601	2.436	6.463
4^{th} nearest neighbor	1.981	0	0.811	1.971	2.856	6.658
5^{th} nearest neighbor	2.266	0	1.153	2.281	3.184	6.664
6^{th} nearest neighbor	2.521	0	1.471	2.554	4.397	6.843
7^{th} nearest neighbor	2.754	0	1.725	2.791	3.781	6.871
8^{th} nearest neighbor	2.969	0	1.934	3.001	4.038	6.980
9^{th} nearest neighbor	3.169	0	2.121	3.205	4.303	6.998

Notes: Summary statistics of agent-month level data obtained from our control state (AP)

A.5.1. Analysis 1: We estimate the association between uptake at an agent with the number of days the n^{th} nearest agent keeps the shop open using the following equation:

$$UPTK_{st} = \alpha_s + \beta_t + \gamma_0 \; OPENDAYS_{st} + \sum_{n=1}^{9} (\gamma_{s_n} \; OPENDAYS_{s_n,t}) + \epsilon_{st} \tag{2}$$

where, α_s denotes agent fixed effects, β_t denotes month fixed effects. We include month and agent level fixed effects to capture aspects such as seasonal patterns, shop location and shop size that are likely to impact uptake at the shop. γ_0 captures the marginal increase in uptake at a shop

[1] The choice of nine nearest neighbors is made to avoid dropping additional agents from our analysis. A choice of ten nearest agents would have resulted in inclusion of only $\sim 92\%$ of the agents as the rest do not have at least ten shops within seven kilometers of haversine radius.

for each extra day the shop is open, γ_{s_1} to γ_{s_9} are the associations of interest. ϵ_{st} is clustered at an agent level.

Results of this estimation are shown in Table A8. We make three observations. First, a one day increase in the number of shop open days of the focal agent is associated with an increase in uptake of 66.417 kg. Second, uptake at the focal agent is negatively associated with number of shop open days of neighboring agents. A one day decrease in the number of shop open days of the nearest neighbor is associated with a 9.357 kg increase in uptake at the focal agent. Third, the strength of association is highest for the first nearest neighbor and decreases in ordinal rank.

Table A8 Association between uptake at an agent and the number of days n^{th} nearest shop is kept open

	(1) $UPTAKE_{st}$
$n=0$, the focal shop (γ_0)	66.417 *** (0.996)
$n=1$, (γ_{s_1})	-9.357 *** (1.000)
$n=2$, (γ_{s_2})	-7.543 *** (0.951)
$n=3$, (γ_{s_3})	-6.286 *** (0.964)
$n=4$, (γ_{s_4})	-3.987 *** (0.957)
$n=5$, (γ_{s_5})	-5.539 *** (0.957)
$n=6$, (γ_{s_6})	-2.493 *** (0.962)
$n=7$, (γ_{s_7})	-5.144 *** (0.955)
$n=8$, (γ_{s_8})	-1.812 ** (0.953)
$n=9$, (γ_{s_9})	-4.099 *** (0.957)
Agent and month fixed effects	Yes
Number of agents	25,114
Number of observations	149,157

Notes: **1)** *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

A.5.2. Analysis 2: We use a poisson regression specification on the shop open days dataset using the following log-linear function to estimate the association between the number of shop open days of an agent and the number of days the n^{th} nearest agent keeps the shop open:

$$\log(\lambda_{st}) = \alpha_s + \beta_t + \gamma_0' \; OPENDAYS_{s,t-1} + \sum_{n=1}^{9}\left(\gamma_{s_n}' \; OPENDAYS_{s_n,t-1}\right) + \epsilon_{st} \tag{3}$$

where, α_s denotes agent fixed effects, β_t denotes month fixed effects. We include month and agent level fixed effects to capture aspects such as seasonal patterns, shop location and shop size that are likely to impact the number of days the shop is kept open. We include the number of days the shop is kept open in the previous month ($OPENDAYS_{s,t-1}$) to account for potential time-varying omitted variables that may systematically influence the number of days an agent keeps the shop open in this month. γ_{s_1} to γ_{s_9} are the associations of interest. ϵ_{st} is clustered at an agent level.

Results of this estimation are shown in Column (1) of Table A9. We make two observations. First, the number of shop open days of the focal agent is positively associated with lagged number of shop open days of the neighboring agents. Specifically, a one day increase in the number of open days of the nearest neighbor results in a 0.25% $((e^{\gamma_{s_1}} - 1) * 100)$ increase in the expected number of shop open days of the focal agent. Considering the average number of shop open days as 12.6 days from Table A7, this translates to an increase of 0.032 days. Second, the strength of association significantly drops beyond the first two nearest neighbors. The decrease suggests that shops that are farther away have lesser influence on the number of days the focal agent keeps the shop open.

Table A9 Association between the number of days an agent keeps the shop open and the number of days the n^{th} nearest shop is kept open in the previous month

	(1) $OPENDAYS_{st}$
$n = 0$, the focal shop (γ'_0)	0.063 *** (0.0004)
$n = 1$, (γ'_{s_1})	0.0025 *** (0.0003)
$n = 2$, (γ'_{s_2})	0.0026 *** (0.0002)
$n = 3$, (γ'_{s_3})	0.0011 *** (0.0002)
$n = 4$, (γ'_{s_4})	0.0016 *** (0.0002)
$n = 5$, (γ'_{s_5})	0.0016 *** (0.0002)
$n = 6$, (γ'_{s_6})	0.0014 *** (0.0002)
$n = 7$, (γ'_{s_7})	0.0009 *** (0.0002)
$n = 8$, (γ'_{s_8})	0.0011 *** (0.0002)
$n = 9$, (γ'_{s_9})	0.0012 *** (0.0002)
Agent and month fixed effects	Yes
Number of agents	25,114
Number of observations	149,157

Notes: **1)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A.6. Identification strategy and parallel trends

In this section, we provide clarification around the parallel trends assumption in our context and more generally, for studies using the reverse DID approach.

A widely used formal test for parallel trends in conventional DID settings involves estimating event-study coefficients by interacting the treatment indicator with the time dummies for pre-treatment periods (Pischke 2005, Jacobson et al. 1993, He and Wang 2017, Jacobson et al. 1993). Each coefficient is interpreted as the difference between the treatment and control in that time period relative to a chosen base time period. Consequently, if the magnitude of the interaction coefficients are ≈ 0, or statistically insignificant, we conclude that the identification condition of parallel trends assumption is satisfied, conditional on all other covariates in the model (Roth et al. 2023, Callaway and Sant'Anna 2021, Braghieri et al. 2022, Dee et al. 2021).

We adapt this approach for our reverse DID setting, where the treatment and control are comparable *after* the intervention and therefore, the test for parallel trends must be conducted after the intervention. Accordingly, in our study setting, we test whether treatment and control move in parallel after the 15th month, when agent-choice was introduced in our treatment state, TS (and was always present in the control state AP). We check for statistical significance of coefficients obtained from interacting the treatment indicator with time dummies for *post-treatment periods*. Each post treatment interaction coefficient is interpreted as the difference between the treatment and control in that time period relative to the first period. Hence, the stabilization of post-treatment coefficients around a constant value, after few initial post-treatment periods would imply that the assumption of parallel trends holds in a reverse DID setting in *post-treatment periods*.

Accordingly, we set up our event-study specification as follows -

$$UPTK_{i,t} = \alpha_i + \beta_t + \sum_{j=1}^{8} \eta_j \, MBAC_{i,j,t} + \sum_{j=0}^{12} \delta_j \, MPAC_{i,j,t} + \theta_1 \, Q^a_{it} + \theta_2 \, STRK_{it} + \gamma \, D_{st} + \varepsilon_{i,t}, \quad (4)$$

where the treatment indicator is interacted with time dummies for post-treatment periods and pre-treatment periods separately. Specifically, Months Before Agent choice, $MBAC_{i,j,t}$ takes the value 1 if agent-choice is introduced j periods after t at sub-district i, and 0 otherwise. Months Post Agent Choice, $MPAC_{i,j,t}$ takes the value 1 if t represents the j^{th} month after introduction of agent-choice at sub-district i, and 0 otherwise. Other variables are as defined in Equation (1). The expectation is that, for the first few periods following the intervention, the effect size may increase (based on the context) and the post-treatment coefficients will stabilize around a constant value thereafter. We consider this stabilization of post-treatment coefficients around a constant value as a test for parallel trends in reverse DID settings.

The results of the above estimation are shown in Table A10. The findings indicate that there is a noticeable rise in uptake by 15,091 and 35,166 kg in the month when agent-choice is introduced and the subsequent month, respectively. Between the second and sixth month post treatment, there is a modest decrease in interest subsequent to the initial enthusiasm (2 MPAC to 6 MPAC). After seven months post treatment, the impact of this change remains stable (7 MPAC to 12 MPAC). Four out of the six coefficients exhibit statistical significance and demonstrate coefficient values ranging between 11,648 and 15,366. We note that 4 MPAC cannot be estimated as it is the month in which agents in our treatment state went on a strike, i.e., $STRK_{it} = 1$. As for 9 MPAC, we learned that AP, our control state, distributed commodities in addition to rice in that month and consequently, may not be appropriate for comparison.

Table A10 Impact of agent-choice over time

Variable	Impact of agent-choice
During the month of agent-choice introduction	15,091*** (4830)
1 MPAC	35,166 *** (4724)
2 MPAC	3670*** (883)
3 MPAC	5492*** (886)
4 MPAC	- (-)
5 MPAC	3015 (4171)
6 MPAC	7186* (4240)
7 MPAC	11468*** (4205)
8 MPAC	16813*** (4108)
9 MPAC	-7229 (4680)
10 MPAC	15366 *** (4196)
11 MPAC	4284 (4201)
12 MPAC	14843*** (4384)
Adjusted R^2	0.898
Observations	28,232

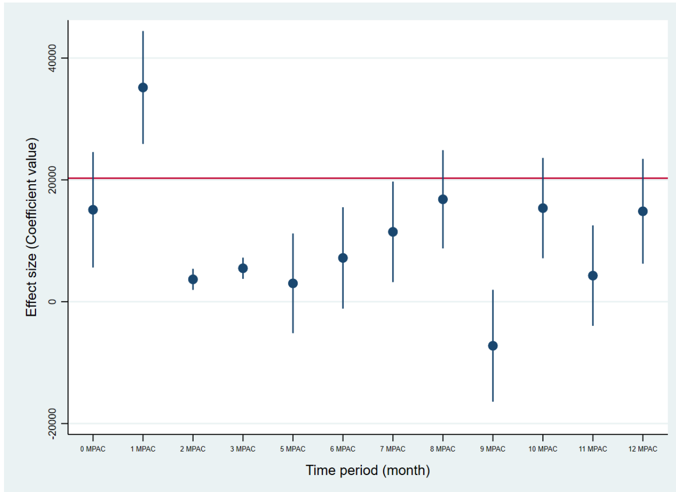
Notes: **1)** Results shown are for model estimated using (4). **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** MPAC - "Month(s) Post Agent Choice". **4)** We omit coefficients of the months before agent-choice (MBAC) for brevity.

We plot the coefficients shown in the Table A10 along with their 90 percent confidence intervals in Figure A2. The figure shows that the treatment effect gradually increases in the first 8 months after intervention and stabilizes during the final four months.

Further, one may argue that the effect we observe in the first few months is the consequence of a growth or learning phase. In such a scenario, we would be interested in the stabilized effect size which excludes these effects. Therefore, we perform an additional verification check by estimating our main model in Equation (1) by excluding the first six months following agent-choice (to exclude potential growth or learning phase), in order to determine whether we can recover the average treatment impact. We find that the impact of agent-choice on this sub-sample is 20,328 kg and is comparable to the effect size we obtained from our main model, 20,273 kg estimated from Equation (1). The stabilization of post-treatment coefficients in a narrow band after the initial few periods post treatment suggests that the parallel trends assumption is reasonable in our context.

Sensitivity analysis. We also perform sensitivity analysis of our results to potential violations in parallel trends by adapting Rambachan and Roth (2023) to the context of reverse DID. We consider the largest violation of parallel trends between two consecutive periods after the introduction of

Figure A2 Graph of impact of agent-choice over time in the post intervention period



Notes: **1)** The X and Y axes indicate MPAC and impact of agent-choice (with 90% confidence intervals) respectively. **2)** The red line indicates the average treatment effect value of 20,273.

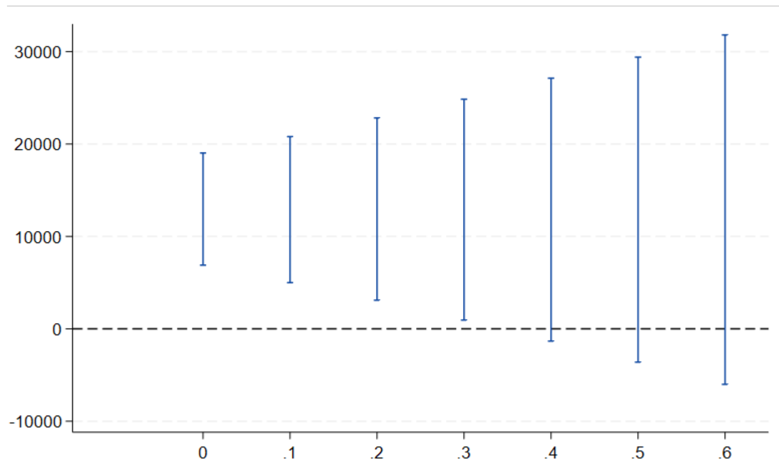
agent-choice from the event-study coefficients, and consider the uptake in AP and counterfactual TS with agent-choice present before April 2018 to diverge by a constant multiple of this violation in successive time periods. We denote the constant multiple as $\bar{M}$ and re-estimate our effect size for different choices of $\bar{M}$.

For each choice of $\bar{M}$, we generate robust confidence sets of our re-estimated effect size by adapting *honestdid* package in Stata by Bravo (2022) to the context of reverse DID. In Figure A3, we show the range of the robust confidence set on the vertical axis at each choice of $\bar{M}$ on the horizontal axis. The figure shows that if we choose $\bar{M} = 0.4$, the robust confidence set of the impact of agent-choice includes a zero, i.e., $\bar{M} = 0.4$ is the "breakdown value". Since the largest violation in parallel trends after the introduction of choice is about $16,813$ kg between 8MPAC and 9MPAC, the value of 0.4 means that if we consider uptake in AP and the counterfactual TS with agent-choice before April 2018 to diverge by about $0.4 \times 16,813$ kg in successive time periods, agent-choice may not result in a significant impact on uptake.

The obtained breakdown value is notably higher than that of other studies which have tested for similar sensitivity and report robustness for $\bar{M}$ ranging between 0.001 and 0.15 (Ang 2021, Miller et al. 2021, Alpert et al. 2022, Bleemer 2022). In conclusion, our results are positive and significant in a broad range of possible deviations in parallel trends.

Falsification checks. Finally, we test for the possibility of the relationship between agent-choice and uptake being a result of pure chance. We randomly assign sub-districts to treatment group and randomly choose a month as a period in which agent-choice was introduced in the treatment group. We repeat this procedure 1,000 times by drawing uniform random variables and assigning agents to control and treatment groups based on the realization. We conduct 1,000 such runs and estimate our main model in each run. If the impact of agent-choice on uptake is pure chance, one would expect statistically significant estimates of comparable magnitude (20,273 per month in a sub-district) in most of the simulations. We find that the effect size averaged over all 1,000 simulations is - 4.9 kgs in a sub-district per month, as shown in Table A11. A t-test for the equality of mean of these 1,000 estimates with 20,273, our estimate from the main model, is rejected with a p-value of 0.0000. Further, agent-choice has a significant effect on uptake, i.e., 0 is not included in 90 % confidence interval of the estimate, in only 11% of the simulations, thereby suggesting that our estimate may not be a manifestation of pure chance.

Figure A3 **Range of robust confidence sets for each choice of $\bar{M}$**



Notes: **1)** The figure shows the sensitivity analysis of estimated effects to potential violations of the parallel trends assumptions per Rambachan and Roth (2023). **2)** The obtained "breakdown value" for a null effect is at $\bar{M} = 0.4$.

Table A11 **Random assignment of treatment and control**

Mean of the pseudo - port coefficient (δ_{pseudo})	-4.9
Mean of p value associated with δ_{pseudo}	0.48
Percentage instances of δ_{pseudo} significant at p<0.1	10.7
Percentage instances of δ_{pseudo} significant at p<0.05	5.7

Notes: The table represents the results over 1,000 simulation runs where sub-districts were randomly assigned to treatment and control groups. The estimated value of the original coefficient of CHC_{it} is 20,273***

Thus, the estimates from event study specification and the sensitivity analysis, taken together with the falsification tests demonstrate that our effect size is reasonably robust.

A.7. Matching

A.7.1. Description of Nightlights dataset Night-lights data, provides a numerical measure of brightness of the earth during the night. It is viewed as an indication of human activities and as a consequence a proxy for economic development (Donaldson and Storeygard 2016). Henderson et al. (2012) use night lights as a tool to augment the GDP series of 188 countries. Luminosity data is available at the '1 deg latitude x 1 deg longitude grid-cell level'

We collect this data from sub-districts in AP and TS for the month April 2019. We then calculate the mean value of all the grid cells over the sub-district to arrive at our measure, which we use as a covariate in the logistic regression for development-based matching.

A.7.2. Matching We use logistic regression to compute the probability of each sub-district obtaining agent-choice. Results in Table A12 indicate that the variable mean radiance significantly predicts the chances of getting selected into treatment.

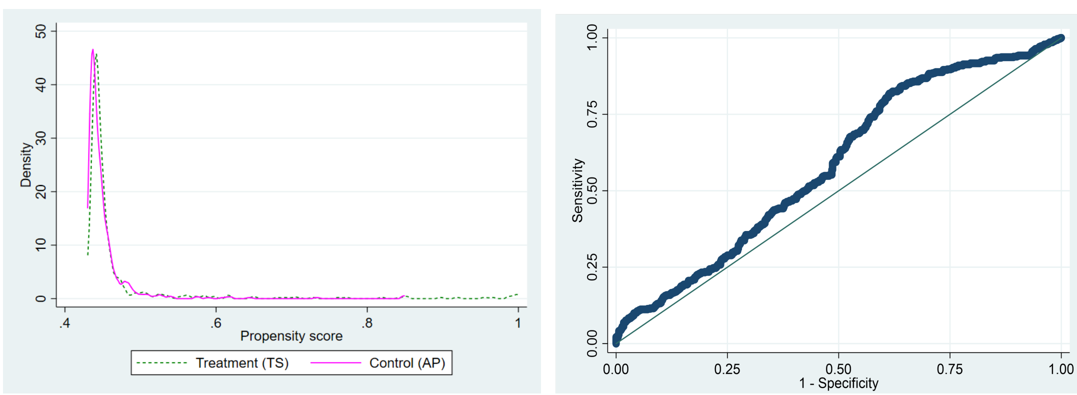
Figure A4 (left) shows the distribution of estimated probabilities for the treatment (TS) and control (AP) groups. We find that the common support, range of estimated probabilities that are common to both control and treatment groups, is between 0.425 and 0.85, and comprises approximately 97% of our observations. Restricting attention to only sub-districts with propensity scores in the common support region results in 26,644 observations for our empirical analysis.

Table A12 Logistic regression using night lights data

Variable	Coefficient
Mean radiance (lumens)	1.12 *** (0.038)
Constant	0.753*** (0.049)
L R Chi-square	23.83
Prob > Chi-square	0.00
Pseudo R^2	0.014
Observations	1206

Notes: **1)** The results shown are from estimating the logistic regression model with robust standard errors.. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure A4 Distribution of propensity scores over treatment and control groups (left) Receiver Operating Curve (ROC) of the logistic regression model (right)



Notes: (**1**) Left: Common support is between propensity score of 0.425 and 0.85, comprises ∼97% of observations. (**2**) Right: Area under the Receiver Operating Curve shown is 0.585.

Figure A4 (left) (right) shows the Receiver Operator Characteristic (ROC) curve for the logistic regression to assess the goodness of fit. The area under the ROC curve is 0.587, which indicates that sub-districts in both states are *more-or-less* similar in terms of night-lights and economic development as a corollary. This strengthens our choice of AP and TS for our evaluation.

We use the probabilities to match sub-districts based on three approaches. First, the *nearest neighbor matching* which selects a sub-district in AP that is nearest in terms of propensity score for each sub-district in TS. Second, *caliper matching* where all sub-districts from AP within a predefined caliper distance (0.1 units), in terms of propensity score for each sub-district in TS. Lastly, *inverse propensity weighting* where sub-districts are made comparable across the two states by assigning low weight to sub-districts in TS with high propensity score and those in AP with low propensity score (Stuart 2010). Results of estimating (1) with the matching techniques are shown in Column (3), (4) and (5) of Table 7 and in Table A13 for uptake as a proportion of quantity allocated as the dependent variable. We find that the effect size is comparable across three methods and to the estimate obtained from our main model, thereby demonstrating the robustness of our result. These results also strengthen our choice of using two neighboring states for evaluation.

Table A13 Exogeneity of Intervention - with $\frac{UPTK}{Q_a}$ as the alternate DV

		Proximity		Development	
	(1) Base model	(2) Proximity	(3) NNM	(4) CM	(5) IPW
Impact of agent-choice (δ)	0.053*** (0.003)	0.126*** (0.021)	0.054*** (0.003)	0.053*** (0.003)	0.051*** (0.003)
Impact of strike (θ_2)	-0.027*** (0.003)	-0.085*** (0.021)	-0.033*** (0.003)	-0.033*** (0.003)	-0.029*** (0.003)
Adjusted R^2	0.19	0.407	0.19	0.19	0.19
Observations	28,232	1,187	24,900	23,907	27,644

Notes: **1)** Results shown are for Sub-district-month panel regressions with Sub-district and month fixed effects, Sub-district time dummies and standard errors clustered at the sub-district level. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** Column (1) shows the base model (results from estimating Equation 1 in the manuscript with the alternate dependent variable $\frac{UPTK}{Q_a}$). **4)** Column (2) shows the results from estimating Equation 1 with the alternate dependent variable for a sub-sample including only bordering districts. **5)** Column (3) shows the results from estimating Equation 1 with the alternate dependent variable by using nearest neighbour matching (NNM) using night lights data. **6)** Column (4) shows the results from estimating Equation 1 with the alternate dependent variable by using Caliper matching (CM) using night lights data. **7)** Column (5) shows the results from estimating Equation 1 with the alternate dependent variable by using inverse probability weights (IPW) using night lights data.

Finally, in Table A14, we present standardized differences among both the groups before and after matching. Results show that our matching procedures improve the comparability of the treatment and the control groups.

Table A14 Summary statistics and standardised differences of radiance before and after matching techniques

	Control Mean (Standard Deviation)	Treatment Mean (Standard Deviation)	Standardised difference
Before matching	0.568 (1.350)	1.824 (8.336)	0.210
Proximity	0.344 (0.314)	0.421 (0.469)	0.193
NNM	0.652 (1.448)	1.824 (8.336)	0.194
CM	0.637 (1.469)	0.719 (1.563)	0.054
IPW	0.810 (2.246)	1.172 (5.764)	0.082

Notes: **1)**The first two columns record the mean and standard deviation of the control and treatment groups. The last column records the standardised difference. The first row pertains to the summary statistics values before matching and the next four rows pertain to the summary statistics values after matching for various matching techniques. **2)** Following Austin (2009), standardized difference (also referred to as Cohen's effect size) is computed 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 the standard deviations of treatment and control group observations, respectively.

A.8. Field visit - Interview guide

We conducted 4 field visits comprising 80 hours spent on the field interviewing beneficiary households, FPS dealers and officials of food and civil supplies ministry, to understand both implementation and validate the evidence gathered from our analysis. We include the relevant part of the interview guide with different stakeholders below:

Households: General questions

- What is the household size, composition and occupation of the earning members?
- What type of ration card do you have?
- For how long have you been in possession of this ration card?
- How many FPSs are there in your village?
- How far are the FPS from the place of your residence?
- Do the FPSs provide services other than grains? If yes, what?

Households: Questions related to PDS transactions

- How much grain did you buy from the PDS last month?
- Which family member collected grain last month? Does the same person collect every month?
- Last month, which shop did you use to purchase your grains? Is that your registered shop? If yes, did you know you could have used another shop? If yes, why did you still choose your registered shop? If no, which shop did you choose and why?
- Have you ever been denied grain? Do you know the reason? How did you respond?
- Do you/your family migrate for work? What is the duration of migration?
- From where did the migrant purchase grains?

Agents

- How long have you been running this shop?
- What do you do other than running the FPS?
- What are your main month-on-month activities to run an FPS?
- Do you provide grains to beneficiaries not allocated to you?
- Typically, how many foreign beneficiaries purchase grains at your shop?
- Typically, how many of your beneficiaries do not purchase ration at your shop?
- What do you do differently because of portability that you did not do before?
- Do you“feel” portability has helped you? If yes, how?, If no, why?
- What do you do when a customer’s transaction fails?
- Does your position in the society or the FPS owner association change if you provide ration to a beneficiary allocated to your neighbouring shop?
- Do you offer any services other than disbursing grains to the visiting customers? If yes, what?
- Do you engage in any activities to attract beneficiaries?

Ministry of civil supplies

- What are your main month-on-month activities to run the network of shops under you?
- How do you plan and monitor the stock levels across the shops?
- How are the FPS activities monitored?
- Do beneficiaries get monthly SMSs regarding stock movements?
- Was there any resistance from the FPS owners when portability was introduced? If yes, why, what was the response from Civil Supplies?
- What is the commission provided to the FPS owner? Has it changed with after portability?
- How has the administrative pressure to run the PDS operations changed due to portability?

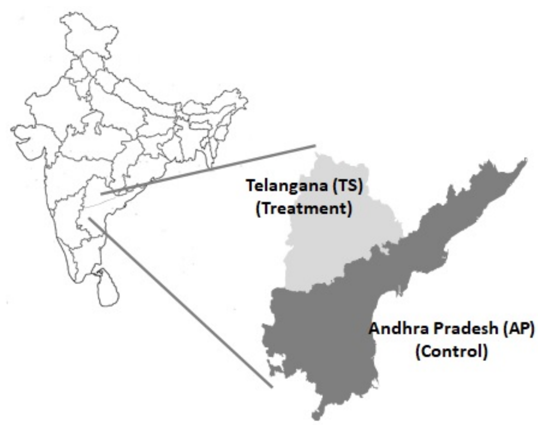
A.9. Other tables and figures

Table A15 Provision of biometric devices - with matching

	(1) Based on BA devices	(2) NNM	(3) CM
Impact of agent-choice (δ)	17,968*** (2,147)	18,508*** (1,935)	18,024*** (1,752)
Impact of quantity allocated (θ_1)	0.94*** (0.01)	0.94*** (0.01)	0.94*** (0.01)
Impact of strike (θ_2)	-9,737*** (1,439)	-11,638*** (14999.59)	-11,869*** (1,455)
Adjusted R^2	0.90	0.91	0.94
Observations	25,264	21,932	20,972

Notes: **1)** Estimation includes sub-district and month fixed effects, sub-district time dummies and standard errors clustered at the sub-district level. **2)** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. **3)** Column (1) shows the results from estimating (1) for a sub-sample in which all agents were provided biometric devices at least 4 months before treatment. **4)** Columns (2), (3) show the results from estimating (1) the sub sample using nearest neighbour (NNM) and caliper matching (CM) respectively.

Figure A5 The neighboring states of AP and TS in India



Notes: **1)** Control and treatment are divided into 659 and 572 sub-districts, respectively.

Figure A6 Scale-up of digitisation of FPSs over time

Notes: **1)** Change in cumulative percentage of agents equipped with biometric devices over time in the treatment.

References

Alpert A, Evans WN, Lieber EM, Powell D (2022) Origins of the opioid crisis and its enduring impacts. *The Quarterly Journal of Economics* 137(2):1139–1179.

Ang D (2021) The effects of police violence on inner-city students. *The Quarterly Journal of Economics* 136(1):115–168.

Austin PC (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.

Bleemer Z (2022) Affirmative action, mismatch, and economic mobility after california's proposition 209. *The Quarterly Journal of Economics* 137(1):115–160.

Braghieri L, Levy R, Makarin A (2022) Social media and mental health. *American Economic Review* 112(11):3660–3693.

Bravo MC (2022) Honestdid: Stata module implementing the honestdid r package .

Callaway B, Sant'Anna PH (2021) Difference-in-differences with multiple time periods. *Journal of Econometrics* 225(2):200–230.

Cox N (2006) Winsor: Stata module to winsorize a variable .

Dalberg (2022) Fulfilling the promise of one nation one ration card. URL `https://policyinsights.in/wp-content/uploads/2022/04/ONORC-report_April-2022-2.pdf`, [Online; accessed 3-May-2022].

Dee TS, Huffaker E, Phillips C, Sagara E (2021) The revealed preferences for school reopening: Evidence from public-school disenrollment. *American Educational Research Journal* 00028312221140029.

Donaldson D, Storeygard A (2016) The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives* 30(4):171–98.

Ganesh M, Allu R, Deo S, Devalkar S (2022) Portability in the public distribution system. *Economic and Political Weekly* Vol. 57(Issue No. 2).

He G, Wang S (2017) Do college graduates serving as village officials help rural china? *American Economic Journal: Applied Economics* 9(4):186–215.

Henderson JV, Storeygard A, Weil DN (2012) Measuring economic growth from outer space. *American Economic Review* 102(2):994–1028.

Jacobson LS, LaLonde RJ, Sullivan DG (1993) Earnings losses of displaced workers. *The American Economic Review* 685–709.

Miller S, Johnson N, Wherry LR (2021) Medicaid and mortality: new evidence from linked survey and administrative data. *The Quarterly Journal of Economics* 136(3):1783–1829.

Pischke (2005) Empirical methods in applied economics lecture notes. URL `https://econ.lse.ac.uk/staff/spischke/ec524/evaluation3.pdf`, [Online; accessed 13-March-2023].

Prakash A, Shukla AK, Bhowmick C, Beyer R (2019) Night-time luminosity: Does it brighten understanding of economic activity in india. *Reserve Bank of India Occasional Papers* 40(1):1–24.

Rambachan A, Roth J (2023) A more credible approach to parallel trends. *Review of Economic Studies* rdad018.

Roth J, Sant'Anna PH, Bilinski A, Poe J (2023) What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics* 235.

Stuart EA (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.