American Economic Review 2023, 113(2): 514–547
https://doi.org/10.1257/aer.20210461

Electronic Food Vouchers: Evidence from an At-Scale Experiment in Indonesia†

By Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Elan Satriawan, and Sudarno Sumarto*

We compare how in-kind food assistance and an electronic voucher-based program affect the delivery of aid in practice. The Government of Indonesia randomized across 105 districts the transition from in-kind rice to approximately equivalent electronic vouchers redeemable for rice and eggs at a network of private agents. Targeted households received 46 percent more assistance in voucher areas. For the bottom 15 percent of households at baseline, poverty fell 20 percent. Voucher recipients received higher-quality rice, and increased consumption of eggs. The results suggest moving from a manual in-kind to electronic voucher-based program reduced poverty through increased adherence to program design. (*JEL* H53, I18, I32, I38, O12)

Targeted food programs aiming to provide nutritional assistance to the poor are one of the most common forms of social welfare programs in the world (World Bank 2018). There are two broad approaches to these programs. The first approach is an in-kind program, which distributes a set amount of free or subsided foods to poor households, such as India's public distribution system and Egypt's baladi bread program (Alderman, Gentilini, and Yemtsov 2018). An alternative approach is a voucher program, which provides poor families with a voucher, or increasingly, a prefilled electronic debit card, which can be used to purchase particular food items at participating shops, such as the United States' "food stamps" program, now known as the Supplemental Nutrition Assistance Program (SNAP).

* Banerjee: MIT (email: banerjee@mit.edu); Hanna: Harvard Kennedy School (email: rema_hanna@hks.harvard.edu); Olken: MIT (email: bolken@mit.edu); Satriawan: Gadjah Mada University and TNP2K (email: elan.satriawan@tnp2k.go.id); Sumarto: TNP2K and SMERU (email: ssumarto@smeru.or.id). Henrik Kleven was the coeditor for this article. This project was a collaboration involving many people. We thank Chaerudin Kodir, Ivan Mahardika, Lina Marliani, Alexa Weiss, and Poppy Widyasari for their outstanding work implementing the project and Aaron Berman, Robbie Dulin, and Michelle Han for their excellent research assistance. We thank our many Indonesian government colleagues, particularly Bambang Widianto, Vivi Yulaswati, Andi Z.A. Dulung, T.B Achmad Choesni, Maliki, M. Oni Royani, Nurul Farijati, Herbin, Gantjang Amanullah, Priadi Asmanto, Ardi Adji, Sri Kusumastuti Rahayu, Jurist Tan, as well as many other colleagues from Bappenas, BPS, the Ministry of Social Affairs, the Coordinating Ministry for Human Development and Cultural Affairs and the Indonesian National Team for the Acceleration of Poverty Reduction for their cooperation implementing the project and data collection. We thank Amy Finkelstein, Jesse Shapiro, and three anonymous referees for helpful comments. This project was financially supported by the Australian Government, Development Innovation Ventures at USAID, and the JPAL Governance Initiative. This RCT was registered in the American Economic Association Registry for randomized control trials under trial AEARCTR-0004675. The IRB of record was MIT COUHES, under protocol 1701819226. All views expressed in the paper are those of the authors, and do not necessarily reflect the views of the many institutions or individuals acknowledged here.

† Go to https://doi.org/10.1257/aer.20210461 to visit the article page for additional materials and author disclosure statements.

Economists often focus on the price-theoretic reasons for why these two types of programs may differ—for example, if the in-kind program provides more of a particular type of good than people would consume otherwise, it may mechanically lead to an increase in consumption of that good. In-kind programs also change supply, whereas vouchers only directly affect demand (Cunha, De Giorgi, and Jayachandran 2019), so price effects could differ.

But, the mechanics of how these programs are *administered* is also fundamentally different—government warehouses' constantly moving around millions of tons of food for direct distribution each month is more cumbersome than electronically refilling electronic vouchers each month. Governments using electronic vouchers may have more control over whether targeted beneficiaries actually get their full benefits rather than having some share diverted to ineligible households, since a debit card is hard to subdivide, whereas in-kind food stocks can be divided by local government officials who control the final delivery of the food to the beneficiary. On the other hand, electronic voucher cards raise their own set of logistical challenges: availability of cell-phone signals for debit card machines, challenges with PINs or other authentication mechanisms, or even challenges in printing and ensuring that the correct debit cards reach the correct households (Banerjee et al. 2018; Muralidharan, Niehaus, and Sukhtankar 2016, 2022). In states with limited administrative capacity, these administrative differences may be first order both in terms of how programs are implemented in practice and their ultimate impacts on the poor.

To study the differences between these two types of programs in a limited state capacity setting, we conducted a unique policy experiment in cooperation with the Indonesian government. Starting in 2017, Indonesia began a national reform to replace its largest antipoverty program—"*Rastra*" (an abbreviation of *beRAS sejahTeRA*, or Rice for Welfare), an in-kind food program that delivered 10 kilograms (kg) of free rice per month to 15 million targeted households nationwide—with an electronic voucher-based program, named "BPNT" (*Bantuan Pangan Non-Tunai*, or Non-Cash Food Assistance), that aimed to provide the same targeted households with a debit card that allowed them to purchase a similar value of rice and eggs from any eligible private provider. The conversions were rolled out district by district, with the districts staggered into waves, such that the government could have the funds and logistical ability to switch over each set of districts. For the 2018 round of district conversions, 105 districts were deemed ready to convert, but funds were only budgeted to cover a fraction of these districts. This allowed for random assignment: 42 of these districts were randomly assigned to receive the program over three waves in 2018, while the remaining 63 districts were randomly assigned to receive the program in 2019.

The scale of this experiment—the 105 districts in the experiment have a combined population of 53 million, or about one-fifth of Indonesia's population, with over 3.4 million beneficiary households—allows us to study how in-kind transfers compare to electronic voucher programs in a real world setting at a large enough scale to incorporate general equilibrium effects (Muralidharan and Niehaus 2017; Egger et al. 2019). To study these issues, we used two primary datasets. First, we wrote a special module that was included in three waves of the national social-economic survey (SUSENAS). Second, we merged the SUSENAS survey with the administration targeting database at the household level to both identify which households were to be targeted by the programs, as well as to provide us with baseline characteristics of these households.

We find that switching from in-kind transfers to electronic vouchers led to a substantial change in the allocation of aid. Whereas in-kind food aid was often subdivided and distributed to many more than the targeted number of households in villages, electronic vouchers were not. A similar amount of total aid was distributed in both voucher and in-kind districts, but conditional on receiving benefits, a household received a subsidy worth 85 percent more in voucher areas compared to in-kind areas. Indeed, in electronic voucher districts, almost all of the households who received assistance in a given month received the full amount they were entitled to under the program's design, which was by no means true in in-kind areas. The share of households receiving a transfer fell in voucher areas: poorer households (at baseline) were 16 percent less likely to receive benefits in voucher areas, indicating an increase in exclusion error. However, the share was disproportionately larger among better-off households (i.e., large reductions in inclusion error), with wealthier households (at baseline) being 49 percent less likely to receive benefits in voucher areas. We also show that the in-kind program led to substantially more aid received by those who are actually poor, rather than being an artifact of who is on the baseline eligibility list. On net, while the number of households (including some poor households) receiving any assistance declined, the vouchers led to increased concentration of benefits among the poor: poorer households (at baseline) as a group received 46 percent more assistance in voucher areas than in in-kind districts.

The increased concentration of benefits among the poor in voucher areas led to a large reduction in poverty, despite the fact that some poor households did lose benefits. For households in the bottom 15 percent at baseline, the share of households below the poverty line fell by 20 percent (4.3 percentage points). Thus, on net, by better concentrating benefits to targeted households, the electronic voucher program reduced poverty.

Why did the voucher program lead to larger poverty reductions? While both subsidy programs aimed to provide food to the poor, we argue that the voucher program made it easier for the national government to implement the program as designed and make sure aid was delivered to the intended beneficiaries, rather than redistributed to other members of local villages. This is because the voucher program's debit cards, redeemable through a network of private agents, were provided directly to the targeted households, essentially cutting out local government officials who maintained control over the distribution within the in-kind program. We then pose a simple model that illustrates how the features of this program could lead to the much higher fealty to the program design—in particular, a stark increase in the share of households receiving the full amount they are entitled to—that we observed in the voucher program.

Of course, while the administrative aspects of the programs differed, there could also be substantial welfare differences between the two programs due to more classical price-theoretic mechanisms associated with the switch from in-kind to vouchers. We explore three such mechanisms. First, voucher programs typically offer more flexibility for beneficiaries than a physically delivered in-kind food basket. If a household would normally consume less of a particular type of food that is included in the in-kind bundle absent the transfer, in-kind transfers could constrain their ultimate consumption decisions. Vouchers, by contrast, could allow households to adjust the mix of items consumed to best suit their needs (see for example, Leroy et al. 2013; Hidrobo et al. 2014; Cunha 2014; Aker 2017; Gentilini 2016).

We do observe some changes in the types of food consumed with the switch to the electronic vouchers. Importantly, more than 97 percent of households consume more rice than what was provided under the in-kind transfers. Given this, one would not expect a switch in the consumption bundle from a transition from the in-kind program to the voucher program, which also allowed households to purchase eggs, since households should have been unconstrained in their rice consumption and able to adjust on other margins. Yet, the switch to vouchers led eligible households to increase their *total* consumption of egg-based proteins, by about 4.3 percent, suggesting some type of stickiness from the types of food included in the voucher programs. This suggest that governments seeking to improve nutrition can potentially do so through adjusting the set of foods they chose to include or not include in vouchers, even if, in fact, basic price theory says that unconstrained households could undo these constraints through their other purchase decisions.[1] That said, these differences in nutrition due to the flexibility from the vouchers—a 4.3 percent increase in egg-based proteins—are small relative to the overall gains eligible households receive from simply receiving a higher amount of subsidy under the voucher program.

A second price-theoretic consideration is that voucher-based programs could lead to increases in prices relative to in-kind programs, particularly in remote areas (Cunha, De Giorgi, and Jayachandran 2019). This is because with in-kind food programs, the government intervenes in the market by supplying the requisite quantity of the goods in question, whereas with a voucher-based program, supply is left to the private market. If supply is elastic, this should not matter, but with inelastic supply (such as, perhaps, in rural or remote areas), this could make a difference (Coate, Johnson, and Zeckhauser 1994; Cunha, De Giorgi, and Jayachandran 2019; Jiménez-Hernández and Seira 2021). On average, however, we do not find any change in rice prices associated with the transition. We observe modest price increases in remote areas—for example, in villages in the seventy-fifth percentile or more in terms of distance we observe a 3.5 percent increase in rice prices—but even taking the price increase into account, poor households are still substantially better off under vouchers because they receive so much more assistance.

A third price-theoretic consideration is that in-kind programs could have important targeting properties through self-selection into the programs (Nichols and Zeckhauser 1982; Currie and Gahvari 2008). Often the food provided in these programs is a type of food that richer households may not particularly need a lot of, or of much lower quality, so much so that richer households may choose not to avail themselves of these programs. But in our case, the higher concentration of benefits to the poor occurred *despite* the fact that the food in the voucher program was rated as higher quality, and aid was more fungible, both reasons that vouchers could have been more appealing to richer households.

[1] This finding is reminiscent of the findings from Hastings and Shapiro (2018) that show that households have a higher marginal propensity to consume food out of SNAP benefits than out of cash. Our findings suggest that even the *types* of foods consumed can be influenced by what is included in vouchers, even when this is not binding.

Taken together, it is likely that the stark reductions in poverty that we observe come from the administrative gains of the voucher program. On top of this, the electronic voucher program is also cheaper to administer. In general, the administrative costs are low even for the in-kind program—4 percent of the benefits disbursed—but, the vouchers are even lower—between 0.75 to 2 percent of benefits disbursed, depending on what portion of the financial transaction infrastructure one apportions to the program relative to other transactions.

In short, switching from in-kind food delivery to vouchers made a real difference in the lives of the poor as it reduced poverty among the poorest households by about 20 percent. This occurred because the vouchers allowed the government to enforce fealty to program design. In the previous, in-kind program, communities had spread the in-kind assistance around to many households in the village; this led to low exclusion error (most people got something), but the poor received comparatively little. By contrast, once vouchers and electronic debit cards were introduced, those on the official eligibility list received the full amount they were entitled to, but those who did not make it onto the list, for whatever reason, received nothing (see also Muralidharan et al. 2022 for related results in India). In this context, this led to, on net, a substantial increase in the amount of aid delivered to those who need it the most, but how much this would be in other contexts depends on the degree to which government's underlying targeting data are accurate and the degree to which communities, when given discretion, use it only to plug exclusion error gaps among the very poor or instead use it to distribute aid to the comparatively well-off.

The rest of the paper proceeds as follows. Section I describes the setting, research design, and empirical methods. Section II presents our findings on the effect of voucher versus in-kind subsidies on the delivery of assistance, targeting and poverty outcomes, and also includes a discussion on administrative capacity mechanisms. Section III examines alternative mechanisms with regards to consumption choices and prices. Section IV explores overall leakages and the relative costs of the programs, while Section V concludes.

I. Setting, Experimental Design, Data, and Estimation

A. *Setting*

In 1998, during the Asian Financial Crisis, the Indonesian government launched a new program to provide a safety net to poor and near-poor households in the form of heavily subsidized rice, delivered directly to beneficiary households through village and local governments.[2] With an annual budget of US$1.5 billion, the subsidized rice program, called Rastra, aims to provide subsidized food assistance to 15 million beneficiary households per month.[3] While the program has changed slightly over

[2] Rastra is implemented by the lowest government administrative unit, known as a *desa* (village) in rural areas or *kelurahan* in urban areas. Both *desa* and *kelurahan* refer to the fourth tier of local government in rural and urban areas, respectively, below provinces, districts/cities, and subdistricts. We henceforth refer to these government units as "villages" throughout the paper, but it is important to note that the program operates identically in both rural and urban areas nationwide.

[3] The program was launched in 1998 under the name *Operasi Pasar Khusus* (OPK, or Special Market Operations). It was renamed *Raskin* (an abbreviation of *beRAS misKIN*, or Rice for the Poor) in 2002, and renamed *Rastra* in 2015.

time, at the time of the study, households were entitled to receive 10 kg of free rice per month, with a value of about Rp 100,000 (US$7) per month, depending on the market price of rice (approximately Rp 9,700 in our period). This is a substantial subsidy: it is about 6.5 percent of the poverty line for a family of 4.

In theory, to be eligible for Rastra, targeting is conducted through proxy-means testing carried out by the central government, with the recipient list provided by the Ministry of Social Affairs. The proxy-means test (PMT) relevant to the study period was computed by conducting a census of more than 28 million households in 2015, which collected a variety of indicators that are predictive of a household's consumption level (assets, education, etc.; see Alatas et al. 2012 and Banerjee et al. 2020 for more details).[4] The National Team for the Acceleration of Poverty Reduction combines these indicators into a PMT score, and then computes percentiles of this score within the entire population; we use this percentile-normed score throughout the paper. In general, households with lower PMT scores are eligible for the program, with the rules varying by district and urban/rural status. Furthermore, villages are allowed some minimal changes to the recipient list through village meetings to replace households who have moved, died, or been double counted. Excluded households can also apply through the local social works agency if they believe they have been excluded from the proxy-means testing in error (Ministry of Social Affairs 2018). According to the 2018 official rules, villages are not allowed to subdivide the rice—benefit recipients must receive the full 10 kg.

Rice is procured by the government through the central logistics agency (BULOG) in rice producing areas, transported (if necessary) to areas that have more beneficiary households than available government rice supply, stored at district-level warehouses, and then delivered to villages (often in 50 kg sacks). BULOG seeks to procure rice at approximately the market price; in our period, BULOG procures rice at Rp 8,600 per kilogram, which corresponds to the average sale price at rice mills for nonquality rated rice in 2017. While the program is as large as an antipoverty program, it represents only about 6 percent of the total rice market.[5] Local level officials are tasked with dividing the rice and delivering it to beneficiaries on the list, either at the village office, at neighborhood distribution points, or door-to-door.

In 2017, the government began their largest social assistance reform in nearly 20 years as they removed Rastra and replaced it with a targeted voucher program. The government did this for a variety of reasons. First, while there is a robust targeting procedure, it was not often followed in practice, with many nonpoor households receiving rice and many beneficiaries only receiving a fraction of their entitlement (Banerjee et al. 2018). The redistribution to nonpoor households in the in-kind program had been a perennial challenge that the government was unable to fully resolve despite repeated attempts to do so. Moreover, it was believed that the voucher program could reduce the improper inclusion of nonbeneficiary households. Second, and related to the first

[4] To comprise the sample, the government first lists everyone on the previous targeting list and asks local leaders to help update the sample list with poor and near-poor households (i.e., the goal is to exclude richer households from the list). In 2015, this resulted in 28 million households being surveyed on household composition and assets variables.

[5] To calculate this, we divide the total amount of Rastra rice allocated to our study districts by an estimate of total rice consumed in those districts using baseline data from the SUSENAS. We estimate that Rastra allocations represent 6.2 percent of the total rice consumed in in-kind districts in 2018.

point, Rastra was known for having high levels of overall leakages—a fair share of the rice went missing (Olken 2006), and there was a belief that the switch to vouchers could reduce this leakage. Third, the quality of Rastra rice was very low—dusty, off-color, full of rocks (Banerjee et al. 2019)—and while this could have self-targeting properties, there was a belief that the move to vouchers could improve program satisfaction by increasing rice quality. Finally, Indonesia has high levels of stunting, even conditional on income levels (Cahyadi et al. 2020), and there was a belief that the vouchers could increase food diversity by allowing households to also purchase eggs.[6]

The new voucher program was called BPNT, or "Non-Cash Food Aid." In principle, the program eligibility for BPNT was the same as Rastra. Instead of receiving rice, households received a monthly voucher of Rp 110,000 on a debit card that was issued to the female adult in the household. The amount was chosen to be approximately equal to the value of the subsidized Rastra rice.[7]

Households could use the voucher for purchases of rice or eggs at a network of eligible small shops,[8] i.e., private shops that were registered as remote agents for the state-owned bank chosen to implement the program in the district.[9] The bank provided each agent with a debit-card reader machine that was connected to the bank's network over a cellular connection and which could directly debit the amount redeemed from the household's voucher account. The voucher could not be cashed out; it could be saved for future months, but that was not always encouraged and rare in practice. The government required banks to increase the number of agents, with a goal of 1 agent per 250 beneficiaries in every village and a minimum of 2 shops in every village. While the agent network did increase, and almost all villages had at least one shop, not all villages met the government-mandated standards (Banerjee et al. 2021). Agents were allowed to source rice and eggs from the private market as they saw fit. Though in practice there was pressure in some areas to source rice from

[6] In fact, in 2020 (subsequent to the period we study here), the BPNT voucher program later morphed into a program called Program Sembako, which allowed households to buy a much wider variety of foods, including carbohydrates (rice, but also locally preferred carbohydrates such as corn or cassava), animal proteins (eggs, and also beef, chicken, and fish), vegetable proteins (nuts, tofu, and tempeh), as well as fruits and vegetables.

[7] The Rastra rice was valued at 10 kg $\times$ the market price of rice, which fluctuates around Rp 10,000 per kg. Since Raskin rice is often delivered in neighborhoods or door-to-door, whereas BPNT can be redeemed at one or two shops in the village, the BPNT amount was slightly higher to compensate for slightly higher transportation costs borne by households. Given the fluctuating price of rice, it is possible that food transfers provide some insurance against a price increase (see, for example, Gadenne et al. 2021), but in this setting, this effect is small relative to the distributional changes we find here. To see this, note that if we examine rice price data from the last ten years, the price of rice never rose more than 6.4 percent higher and never fell more than 15.4 percent lower than the price we observe at the time of this study. Cross-sectional variation is also low: the tenth percentile of rice prices is 16 percent below the mean, and the ninetieth percentile of rice prices is 14 percent above it. Thus, as we will show below, these differences in price are small relative to the substantial increase in quantity that households receive with the vouchers. Moreover, the Indonesian government can adjust the BPNT value to compensate for such macroeconomic shocks and, in fact, did exactly this in the recent COVID-19 pandemic in which the benefit of the BPNT was increased by 25 percent (note that this is outside of the time frame that we study), similar to other countries' expansions of social assistance during COVID-19 (e.g., Londoño-Vélez and Querubin 2022).

[8] While the central government has no way of electronically monitoring that the agents in fact only redeem vouchers for rice and eggs, they could in principle audit this. In practice, households in the SUSENAS report that more than 99 percent of the vouchers were spent on rice and eggs, and less than 1 percent on other items. The set of allowable items was expanded substantially to include other food essentials in 2020, subsequent to the period we study here.

[9] The debit cards were issued and administered by one of the four state-owned banks, Bank Mandiri, BRI, BNI, and BTN. The government assigned one bank to administer the program in each district. As such, they could only be redeemed at agents of the designated bank.

the government logistics agency, this is secondary, and the large-scale movement around the country of government-procured rice by the government logistics agency under the Rastra program was dramatically scaled back under the BPNT program.

B. *Sample*

The conversion from Rastra to BPNT was rolled out over approximately four years, as the government had budget targets in place for the yearly conversion process. The voucher program was first rolled out to 44 cities as a pilot in 2017. For the second phase, in 2018, the government first chose a number of districts (primarily in East Java) to definitely be converted in 2018. The government then ranked all of the remaining districts by a readiness indicator; based on this indicator, 105 districts were deemed potentially ready to be converted next. However, given the government's budget—they planned to convert the program for about 8.3 million beneficiary households total in 2018, including those in districts that were definitely chosen—the budget was insufficient to convert all of the 105 potentially ready districts to BPNT in the same year.

Therefore, the government randomized which of the 105 districts would be treated in 2018, and which ones would subsequently be treated in 2019 (see online Appendix Figure 1). These 105 districts—spread across Indonesia's diverse landscape—comprise our sample (see online Appendix Figure 2), encompassing about a fifth of Indonesia's population, with a total population size of 53 million individuals and about 3.4 million targeted beneficiary households. Given the scope, as well as the fact that these were real programs conducted by the government, this vast sample gives us a unique opportunity to study both the general equilibrium effects and administration outcomes of moving from an in-kind to a voucher program.

C. *Randomization Process and Compliance*

We randomized at the district level, as this is the level of administration for both the in-kind and voucher programs. Forty-two districts were randomly chosen to be converted in 2018, while the remaining 63 were scheduled to be converted in 2019 (see online Appendix Figure 1). The treatment districts were additionally randomized to be converted in three waves: 10 districts in May 2018, 18 in October 2018, and 14 in November 2018 (see online Appendix Figure 3).

Note two aspects of the randomization. First, we grouped districts by geography, and we then stratified our randomization based on this geographic grouping (a map showing the randomization results is in online Appendix Figure 2). Second, the government added a constraint that it wanted the randomization results to come as close as possible to exhausting the budget of 8.3 million households (including the districts that were chosen for sure, as well as randomly chosen districts). To accomplish this, in setting up the geographic strata, we reserved the 20 districts with the smallest number of beneficiaries and put them in a special 'holdout' stratum, which we randomize last. For this last 'holdout' stratum, we randomized order of districts, and assigned districts to treatment until assigning one more district (in randomized order) would have exceeded the 8.3 million total budget. Therefore, while whether a given district is treated is always random, the treatment probability in this last holdout, stratum is different than in other strata; to account for this, we include strata fixed effects in all regressions. We also show below robustness results dropping this holdout stratum entirely; results are very similar.

In practice, the randomization was followed almost exactly. All of the districts randomly chosen to receive vouchers were converted to vouchers in 2018. Only 3 of the 63 districts randomized to be converted in 2019 were in fact converted in 2018; given this small number, we study the reduced form effect of being assigned to the vouchers (i.e., the intent-to-treat effect).

D. *Data*

In order to evaluate the experimental switch from in-kind transfers to the voucher program, we worked with the Government of Indonesia to design a special social protection module to be added to the SUSENAS, the Indonesian national household survey. The SUSENAS is a nationally representative household survey, completed twice annually—a large wave in March covering about 300,000 households and a smaller wave in September covering about 75,000 households—by the Indonesian Census Bureau (BPS). It is a repeated cross section of households.

We use two waves of the SUSENAS for our main analysis: March 2018 and March 2019. We focus on the March 2019 SUSENAS as our endline, as all of the districts randomly assigned to vouchers had been converted before March 2019, whereas the districts randomly assigned to later conversion had not (see timeline in online Appendix Figure 3). The special social protection module collected questions on whether households received each program (Rastra and/or BPNT), how much they received from each program, prices and quantities of subsidized rice and eggs, the quality of rice, and so forth. In addition to questions from this social protection module, we also analyzed the separate, detailed consumption module that is normally administered in the survey, which not only allowed us to look at total household consumption of different commodities (e.g., rice, eggs, cigarettes, etc.), but also the household poverty measures. Note that these two modules—the social protection module, which covers social assistance, and the consumption module, which covers consumption from all sources—are distinct sections of the survey. For all variables, we construct district $\times$ rural/urban averages of the same variables in the March 2018 SUSENAS, which we merge in at that level as potential baseline control variables.

The second dataset that we use is administrative data from the Unified Targeting Data Base (UDB). As described above, this database was constructed for targeting purposes in 2015 and forms the basis of eligibility for both the in-kind and voucher programs. The government merged UDB variables into the SUSENAS data using each individual's national identification number, which is collected in both datasets; we then analyzed the deidentified version of this merged data. The merge is important for two reasons. First, it allows us to independently identify who in the SUSENAS is likely to be eligible for the program, since targeting was done by the Ministry of Social Affairs using UDB data. Second, as the SUSENAS is a repeated cross section, we do not have household-level baseline variables, which are useful to examine heterogeneous treatment effects and to absorb residual variation. The UDB variables, since they were collected in 2015, give us baseline measures of household well-being (PMT percentile scores) for heterogeneity analysis, as well as baseline family composition and assets that can be used as potential control variables. In particular, the PMT percentile score is the predicted percentile of the consumption distribution the household would be in, based on the household's assets, composition, etc. Finally, we merged the 2018 village census (PODES) to provide village-level baseline control variables, as well as to be able to explore heterogeneity by baseline village characteristics.[10]

E. *Experimental Validity*

We examine potential differences between the in-kind and voucher districts using baseline data from the 2018 SUSENAS. We chose the variables for the randomization check in January 2019, covering per capita consumption, caloric consumption and basic assets. In general, looking at the differences in online Appendix Table 1 across the variables considered, we find that treatment and control groups appear balanced at baseline. Only 1 of the 11 variables is statistically significant, which is what one would expect by chance, and the randomized inference p-value of a joint F-test is 0.384.

F. *Estimation*

We estimate the impact of the experimental switch from in-kind assistance to vouchers using equation (1):

$$y_{hvds} = \beta_0 + \beta_1 Voucher_{ds} + \mathbf{X}'_{hvds}\gamma + \alpha_s + \epsilon_{hvds}, \tag{1}$$

where y_{hvds} is the relevant outcome variable; $Voucher_{ds}$ is the randomization into the voucher conversion in 2018; $\mathbf{X}'_{hvds}$ is control variables selected using a double LASSO (least absolute shrinkage and selection operator) (Belloni, Chernozhukov, and Hansen 2014); and α_s is strata (s) fixed effects. In the LASSO, we include baseline control variables, such as household (h) covariates from the UDB, village-level covariates from the village (v) census and district $\times$ urban/rural averages of variables from the March 2018 SUSENAS. A full list of the control variables used as input into the LASSO is shown in online Appendix B. We estimate the intent to treat using the original randomization results (only three of our control districts were treated during the study period, so treatment-on-treated models would be very similar). Standard errors are clustered by district (d), as that is the level of randomization; we report randomization inference p-values in all tables (Young 2019).

Following our preanalysis plan, we differentiate results based on households that are likely to be eligible for subsidized assistance and households that are

[10]The 2018 PODES was enumerated from May 2 to May 31, 2018. Since BPNT distribution in the first set of units did not occur until the end of the month (starting May 25, 2018), and since PODES captures long-moving variables such as population, infrastructure, and so on, we consider the 2018 PODES as baseline variables. (See https://www.tribunnews.com/bisnis/2018/04/22/bantuan-pangan-non-tunai-tahap-2-cair-bulan-mei-2018-penerimanya-nambah-2-juta.) We present robustness results below that drop as well as other control variables; results are qualitatively similar.

unlikely to be eligible. To determine this, we use the baseline PMT score each household received in 2015 in the UDB (recall, as described above, that the PMT score is the household's predicted percentile in the national per capita consumption distribution based on household assets and household composition). The Ministry of Social Affairs drew up the final list of eligible beneficiaries in each district using this dataset as a basis. Since the program was aimed at households in approximately the bottom 30 percent of the population, following our preanalysis plan we divide households into two groups: those with a PMT score of 30 or below, and those who either had a PMT score above 30 or were not included in the UDB.[11] We also examine several additional cuts of the PMT score distribution, and in particular those with a PMT score of 15 or below, i.e., below the approximate poverty line.

While imperfect, the 2015 PMT score cutoff of 30 has strong predictive power both for consumption/poverty levels and for program eligibility. Online Appendix Table 2 provides summary statistics from the March 2018 SUSENAS data (i.e., the baseline household survey), separately for the two main groups we analyze those with PMT scores $\leq$ 30 (i.e., the approximate target group), and those with PMT $>$ 30 or no PMT score. The data reveal that households with PMT scores $\leq$ 30 are substantially poorer on virtually all dimensions. In particular, for households with PMT scores $\leq$ 30, average per capita consumption is Rp 690,000 (US$49) per month; for those with PMT score $\geq$ 30, average per capita consumption is Rp 1,116,000 (US$80) per month—i.e., 62 percent higher. Online Appendix Figure 4 plots this relationship as a bin-scatter, showing average baseline per capita consumption levels for households by PMT scores (in bins of five). There is a very strong, nearly linear relationship between PMT scores and mean baseline consumption levels. This is true within the set of households who have PMT scores—households with a PMT score from 0 to 5 have mean consumption of about Rp 590,000 per month, compared to Rp 940,000 or more for those with PMT scores 40 and above. Households who were not included in the PMT process (shown as PMT $=$ 100 on the graph) have even higher average consumption, averaging almost Rp 1,200,000 per capita.

Online Appendix Table 2 also analyzes the relationship between PMT scores and a number of other variables, including some not included in the PMT formula (i.e., poverty rate, ownership of a flat screen television) and variables included in the PMT formula (e.g., low-quality wall, floor, or roof material, ownership of refrigerators). Households with low PMT scores appear poorer on both sets of variables.

The PMT cutoff that we use is also strongly associated with program eligibility: the households in the PMT $\leq$ 30 are substantially more likely to have been officially eligible for the Rastra program in 2017 (75 percent eligible for those with PMT $\leq$ 30, compared to only 5.7 percent for those with PMT $>$ 30), and to have received the Rastra program in 2018 (71 percent for those with PMT $\leq$ 30, compared to only

[11] A full preanalysis plan can be found at https://www.socialscienceregistry.org/trials/4675. We prespecified the regression in (1), as well as prespecified the separation by baseline PMT score $\leq$ 30 or not. The primary outcome variables specified in the preanalysis plan are total subsidy received, self-reported quality of rice, and food insecurity indicators. The "populated analysis plan" (Banerjee et al. 2020) for the primary outcome variables can be found in online Appendix C.

33 percent for those with PMT $>$ 30). Given these results, we interpret the split based on household preperiod PMT score as indicative of both official program eligibility and overall poverty levels.

II. Delivery of Assistance, Targeting, and Poverty Impacts

A. *Fidelity to Program Design*

We begin by plotting the distribution of the total amount of subsidized assistance received per month for those households that receive anything at all. Specifically, we calculate the monthly value of assistance that households receive from either the voucher or in-kind program, in rupiah.[12] For the in-kind program, we value rice received in-kind at the market rate in each rural or urban area of the given island group, measured from the SUSENAS consumption module.[13] For the voucher program, we take the average reported receipt in rupiah; if the monetary value is missing in the survey but the household reports the amount of subsidy spent on rice, eggs, or other goods, we sum these values.[14]

Figure 1 plots histograms of the amount of subsidy received (for those who receive assistance) in both in-kind and voucher districts. Panel A in Figure 1 shows all respondents who receive assistance; panel B restricts the analysis just to those with PMT score $\leq$ 30.

Figure 1 shows a remarkable shift in the distribution of aid received, showing dramatically increased fidelity to program design in terms of the amount received for those who receive assistance. Specifically, in voucher areas, almost everyone who receives assistance receives the full amount. In fully 81 percent of months in which households in voucher areas report receiving any subsidy, the amount reported is exactly Rp 110,000. The truth is probably even higher: if we include those who report amounts of BPNT spent on rice, eggs, and other goods on the survey rather than rupiah amounts of BPNT received (and for which we need to impute a rupiah amount) but which are within 10 percent of Rp 110,000 (i.e., accounting for reasonable local variation in prices), then fully 92 percent of those who receive assistance receive the full amount.[15]

[12] Note that US$1 averaged Rp 14,000 during the period of this study.

[13] By island groups, we refer to the main regions of Indonesia. These units are even larger than provinces and much larger than the districts which are the unit of analysis here. The island groups we use in our data are Sumatra, Java, Kalimantan, Sulawesi, Maluku, Papua, and Bali/Southeastern Indonesia.

[14] In principle, using the monetary values is preferred since it captures the market value of what recipients receive and what they choose to buy and where they choose to buy it. One might be concerned, however, that this approach could overstate the value received if the voucher-program led to higher prices. As discussed in Section IID below, we do not, however, find prices for rice change on average. Nevertheless, as a robustness check, we consider an alternative definition of the subsidy variable where we use the quantities of rice and eggs received, which we then value at the same fixed set of prices for both programs. The results, presented in online Appendix Table 3, panel A, are broadly similar to the main results in Table 1.

[15] Figure 1 is conditional on receiving assistance. To verify that this is not driven by differential selection, online Appendix Figure 5 reproduces Figure 1 unconditionally, i.e., including all households, including those that receive zero subsidy. One can still see dramatic changes in the amount received, and in particular, the dramatic increase in the share of households receiving the full amount, in this unconditional figure. For example, focusing on panel B (those with PMT $<$ 30) in the in-kind program, less than 14 percent of all households with PMT $\leq$ 30 receive the full amount in a given month (i.e., Rp 95,000 per month or greater). By contrast, in the voucher program, this rises to 35 percent.

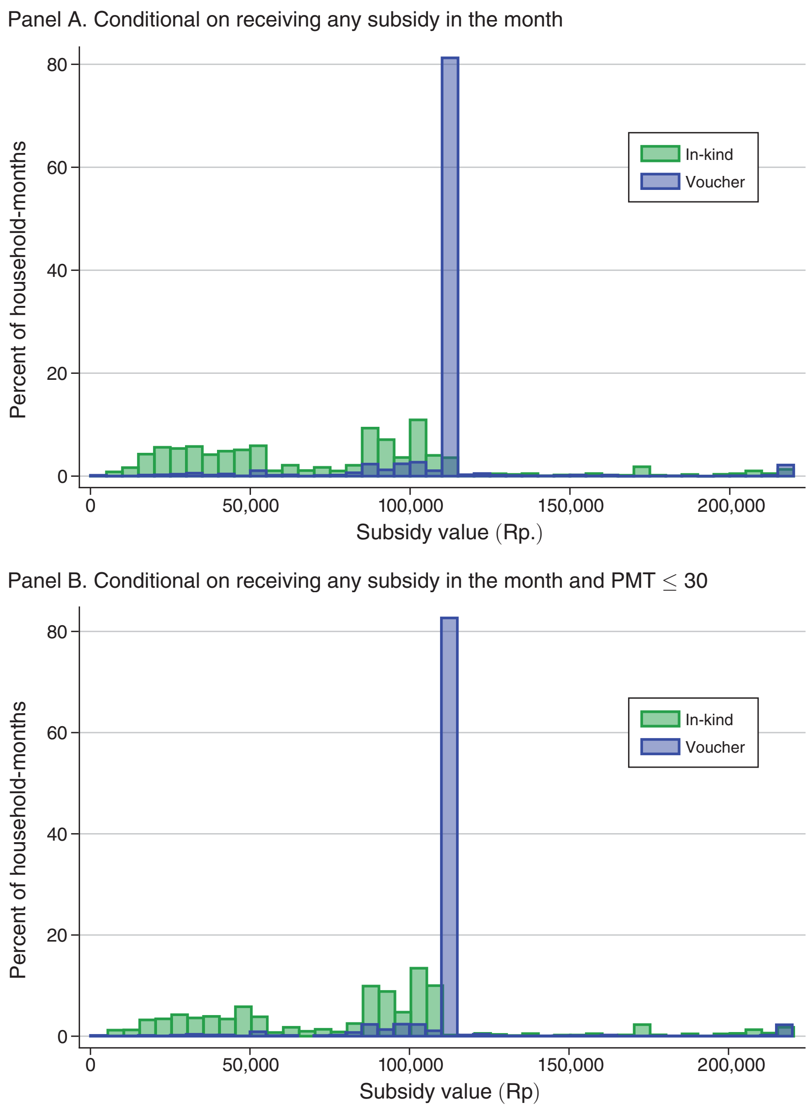

Figure 1. Distribution of Subsidy Amounts Received in Month

Notes: Observations are at the household-month level. For the purpose of illustration, subsidy values above 220,000 have been top coded. Panel A: $N = 44{,}567$. Panel B: $N = 24{,}344$.

By contrast, the Rastra program looks very different. As described above, the full value of the monthly Rastra distribution was approximately Rp 97,000 at then-prevailing market prices. Figure 1 shows that in only 24 percent of months do Rastra-recipient households report an amount within 10 percent of this. Instead, large numbers of households report receiving only 20–50 percent of this amount. Indeed, the mean amount received (conditional on receiving something) is Rp 112,000 in BPNT districts versus only Rp 67,000 in Rastra districts. Thus, the BPNT areas showed a dramatic difference in fidelity to program design, at least for those who receive assistance.

B. *Who Gets Assistance, and How Much?*

Of course, the previous figure is only indicative since it is conditional on a household receiving assistance. To evaluate the impacts of the switch from the Rastra to BPNT programs more systematically, we examine the impacts on the receipt of subsidized assistance for all households (i.e., not conditional on program receipt). First, we examine the average monthly value of assistance that households receive from either the voucher or in-kind program, in rupiah, averaged over the four months prior to the survey. Note that households who received nothing from either program are included with a value of zero.

Figure 2, panel A begins by showing the treatment effect on value of the total subsidy received nonparametrically. We break the households into bins based on household preperiod PMT scores (i.e., PMT scores 0–5, 6–10, and so on, with those who do not have PMT scores—typically the wealthy—assigned the maximum score of 100).

Switching from an in-kind program to vouchers greatly increased the ability of the government to target aid to eligible households. Poorer households, as measured by having lower baseline PMT scores, experience substantial gains in the total amount of subsidy received. The gains are most pronounced for those with PMT scores of 20 or below, which roughly corresponds to the bottom 18 percent of the population. Households in the middle of the distribution received about the same amount of subsidy on average. By contrast, households at the top—i.e., those without PMT scores (shown as PMT $=$ 100, which are largely households that local communities saw as too wealthy to be measured as possibly poor)—saw the amount of subsidy they receive fall considerably.

Table 1 presents the regression results from estimating equation (1) for total subsidy, shown for all households (column 1), households with PMT scores $\leq$ 30 (i.e., targeted beneficiaries), and PMT score $>$ 30 (i.e., those not targeted).[16] The key result is that the voucher program reallocated aid toward those with low PMT scores, i.e., the poorer, targeted group.[17] Specifically, as shown in column 2, targeted households received about 46 percent more in subsidy in voucher districts (Rp 13,496 additional per month on average, compared to Rp 29,219 per month in in-kind areas; $p < 0.001$). Conversely, the vouchers led to a reduction in subsidy received for those above 30. Those with PMT scores $>$ 30 already received substantially less in in-kind areas—only Rp 9,162 on average—and this fell by 28 percent in voucher areas ($p = 0.002$).

[16] Table 1 examines outcomes in the March 2019 SUSENAS. The September 2018 SUSENAS also included part of the new social protection module. It was a smaller wave in terms of both sample size and included questions (e.g., the quality question was not included). Perhaps more importantly, only 10 of our 42 treatment districts were randomized to be treated by the time of the survey, and so we would expect to have less statistical power. Nonetheless, if we run our analysis only in September (defining as treated the 10 districts randomly selected to be treated prior to September), or in the pooled September and March data, we find very similar results. See online Appendix Tables 4 and 5.

[17] Our results are remarkably robust to specification choices, such as dropping the controls selected via LASSO (online Appendix Table 6), dropping the holdout sample (online Appendix Table 7), or winsorizing the subsidy outcome (online Appendix Table 8). We also decompose the results separately based on whether the district was randomly chosen to be converted to in-kind in May, October, or November 2018. The results in online Appendix Table 9 show similar results in all three waves, with if anything qualitatively larger effects for those districts that had the voucher program the longest (i.e., the May 2018 group).

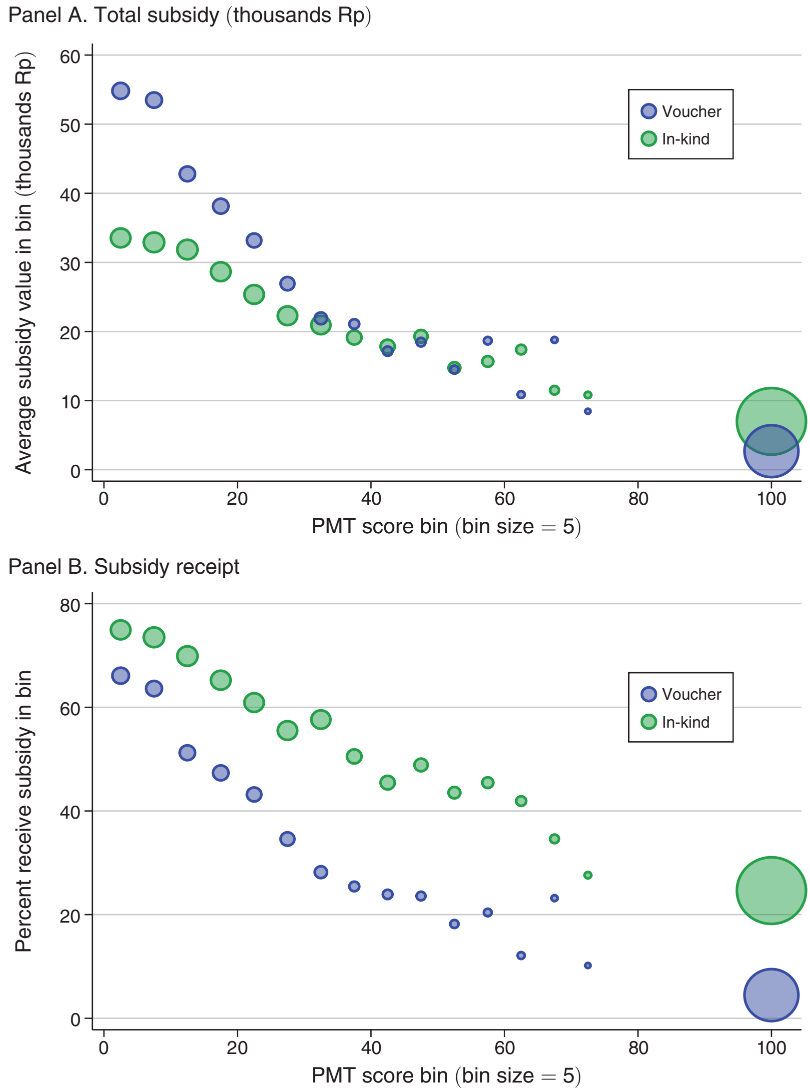

Figure 2. Relationship between PMT Score and Subsidy Received

Notes: These figures graph the relationship between a household's PMT score and, respectively, their total subsidy received (panel A) and an indicator variable for received subsidy (panel B), by treatment. PMT scores are binned in groups of five, with those who have no PMT score grouped with those with a score of 100. Markers are scaled by the number of households in each bin. Data on outcomes are from the March 2019 SUSENAS, while PMT data are from the UDB.

Overall, the mean amount received across all households was slightly higher in voucher areas, by about 10 percent (column 1, $p = 0.063$). This likely reflects the fact that, due to fluctuations in the market price of rice, at the time of our survey the market price of rice was only Rp 9,700 per kilogram nationwide, so 10 kg of in-kind rice would be valued at Rp 97,000, compared to Rp 110,000 for the voucher. In online Appendix Table 10, we redo Table 1, but scaling the amount received in the in-kind program by 110/97 so that the amounts between the two programs are

Table 1—Experimental Difference between Voucher and In-Kind Districts on Subsidy Outcomes

	Total subsidy (Rp)			Receive any subsidy			Recipients only: Total subsidy (Rp)	Recipients only: Rice quality
	All	PMT $\leq$ 30	PMT $>$ 30	All	PMT $\leq$ 30	PMT $>$ 30	All	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Voucher	1,404.537	13,495.899	−2,531.862	−0.134	−0.105	−0.145	31,332.589	0.203
	(617.436)	(1,908.590)	(564.413)	(0.019)	(0.021)	(0.020)	(3,190.397)	(0.020)
	[0.063]	[0.000]	[0.002]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	66,494	16,327	49,566	66,496	16,329	49,566	19,355	19,260
Stratum fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Double lasso	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DV mean (control)	14,461.335	29,218.903	9,162.138	0.393	0.669	0.293	36,930.909	0.630

Notes: This table experimentally estimates the difference in subsidy outcomes for voucher versus the in-kind districts. Total subsidy (Rp) is the sum of Rastra and BPNT subsidy values, while received subsidy is an indicator variable for receiving any amount of subsidy. In columns 2 and 3 we present the results on total subsidy disaggregated by whether the household is targeted (a PMT score below or equal to 30) or not targeted (score above 30); we do the same in columns 3 and 4 for received subsidy. The quality of rice measure is standardized between 0 and 1, where 1 is the highest quality. For continuous outcome variables, we drop any value greater than 12 standard deviations from the mean. The outcome data come from the March 2019 SUSENAS; the PMT data come from the UDB. We used a double LASSO to choose the control variables (all potential variables used as inputs for the LASSO are listed in online Appendix B). Standard errors are clustered at the district (*kabupaten*) level and displayed in parentheses. Randomization inference p-values are from 1,000 permutations of the treatment assignments and are displayed in brackets.

comparable. When we make this change, column 1 shows that the overall amount received is the same, but we still see the same dramatic increases in subsidy received by eligible households and declines among ineligible households.

As described above, change came about because households who received the transfer received substantially more in voucher areas; i.e., the aid was substantially more concentrated to the poor in voucher areas. On the one hand, Figure 2, panel B shows that the probability of receiving any aid is lower for *all* groups in voucher areas compared to in-kind areas, regardless of their PMT score. This difference in the probability of receiving any aid is smallest for those with the lowest PMT scores, but more substantial for those with PMT scores above 30. Table 1, columns 4–6, show the results in a regression format. The probability of receiving any assistance falls 16 percent (10.5 percentage points; $p < 0.001$) in voucher areas for those with PMT scores $\leq$ 30; for those with PMT scores $>$ 30, the probability of receiving assistance falls by 49 percent (14.5 percentage points; $p < 0.001$).[18] Thus, while everyone across the income distribution has a lower probability of receiving the voucher program than the in-kind program, the reductions were much more pronounced for those who were not targeted.

On the other hand, those who do receive assistance receive substantially more in assistance in voucher areas. Column 7 of Table 1 shows that, conditional on receiving assistance of some form in any of the last four months, recipients in voucher

[18] The fact that households with low PMT scores were slightly less likely to receive assistance under the voucher program may be in part a consequence of the fact that the PMT lists we use here for analysis are not the exact final lists used for program distribution. As described above, the Ministry of Social Affairs made some revisions to the initial targeting lists we use here before sharing them with local governments, but we were unable to obtain these data. It is possible then that some of the households that we observe not receiving the program were in fact ineligible households according to the final Ministry of Social Affairs list.

areas received 85 percent more than in in-kind areas on average over the previous four months (Rp 31,333 more; $p < 0.001$). In particular, as shown in Figure 1, conditional on receiving anything, recipients in voucher areas are substantially more likely to receive the full amount they are entitled to.

Combined, these results show a clear pattern. In in-kind districts, assistance is subdivided into much smaller amounts, so that the typical beneficiary receives only about a third of the intended transfer size. The additional funds are used to give many more households assistance, despite the fact that this is explicitly against the stated rules. In voucher areas, this practice is substantially less prevalent. Somewhat fewer households receive assistance, but those that receive assistance receive substantially more, usually the full amount they were entitled to.[19] In short, the voucher program exhibited dramatically higher fealty to program design, and as a result, transferred substantially more to those in the lowest percentiles of PMT scores.[20]

We next show that this represents substantially more aid received by those who are actually poor and is not just an artifact of who is on the baseline eligibility list. Specifically, one potential concern could be that the baseline PMT scores do not fully capture true consumption, and that the targeting adjustments that communities were doing in the in-kind program were toward poor households excluded by the PMT scores, an idea suggested by Alatas et al. (2012). To examine this issue, Figure 3 graphs the subsidy received by per capita consumption percentile bins, rather than PMT bins. Note that since this measure is contemporaneous consumption (unlike our PMT data, which is measured several years prior), we subtract out the value of subsidy received from either the in-kind or voucher programs prior to calculating this graph in order to capture pre-subsidy consumption levels. Figure 3 shows the same pattern as our main results—the voucher program increases the amount of subsidy received by the poorest households—particularly so for those in the bottom decile, but also for those in the twentieth and thirtieth percentiles, and it does so by reducing the amount of subsidy received by wealthier households, particularly those in the sixty-fifth to ninetieth percentiles of the per capita consumption distribution. This analysis shows that the voucher program really did deliver more aid to the poor.

To further probe this issue and understand the targeting of the vouchers, we look within the voucher areas, and compare consumption levels for households who did and did not receive the BPNT program. These results are shown in online Appendix Table 13. Columns 1–3 present results with no controls, with district fixed effects, and controlling for dummies for each PMT score level (i.e., a dummy for each integer of PMT scores, i.e., 1, 2, 3, etc.). The outcome variable is per capita consumption less the value of the subsidy received (i.e., a reasonable measure of what consumption would have been in the absence of either program). Columns 4–6 present the same results, but with raw per capita consumption as measured at endline (i.e., not making

[19] Related to this, we next check whether the reform appeared to have larger effects in areas where subsidy receipt was low at baseline. We examine this in online Appendix Table 11, exploring the effect of the vouchers in areas where targeted households got the least amount of subsidy at baseline. We find no difference, but this may be because the level of subsidy received by households everywhere was pretty low: at baseline, less than half of the total subsidy was provided to targeted households in 85 percent of district-urban/rural units (and less than 60 percent of the subsidy for 95 percent of units).

[20] One may also be concerned that by concentrating benefits to poorer households, there may be dissatisfaction within the village that is expressed through protests or voting in new local leaders. Online Appendix Table 12 shows no observable differences in either of these outcomes.

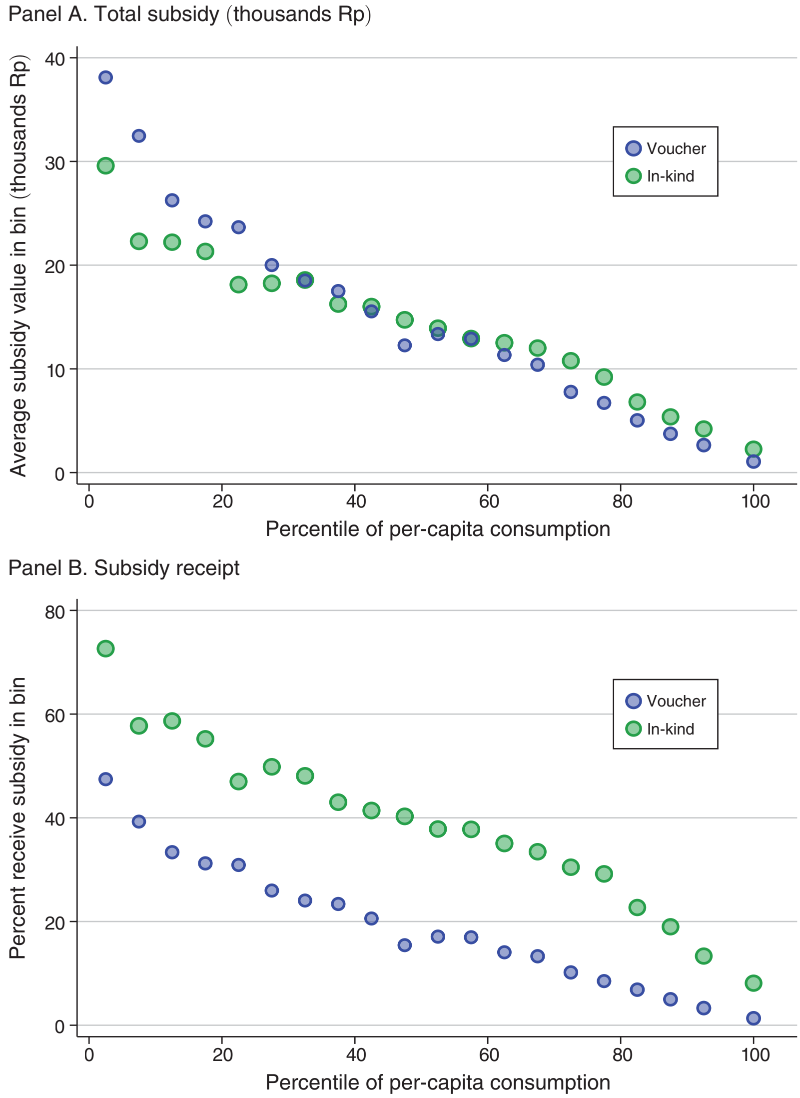

Figure 3. Relationship between Percentile of Consumption and Subsidy Received

Notes: This figure replicates Figure 2, plotting bins of per capita consumption on the *x*-axis instead of bins of PMT score. Data are from the March 2019 SUSENAS. See Figure 2 for details.

any adjustment for program receipt). Panel A considers the entire sample and panel B restricts to households with PMT scores < 30.

We find that, within voucher districts, households who receive the BPNT voucher are substantially poorer than those who do not. Remarkably, this is true even controlling flexibly for the PMT score—households who receive BPNT are about 18 percent poorer than households who do not, even holding PMT scores constant (column 3). This suggests that, to the extent there are local deviations from the raw PMT scores in allocating vouchers, they go in the direction of including poorer households.

It is also important to note that while there is exclusion error in the voucher program, most of those excluded are not among the very poorest. For example, in voucher districts, only 12.8 percent of targeted (PMT $\leq$ 30) households that receive no subsidy are poor (i.e., below the national poverty line) and 15.4 percent of targeted (PMT $\leq$ 30) households are near poor (i.e., between 100 and 125 percent of the poverty line).

To investigate whether local officials target on other dimensions (as in Galasso and Ravallion 2005; Alatas et al. 2012; and Basurto, Dupas, and Robinson 2020), we further explore differential targeting of the two programs on other household characteristics in online Appendix Table 14. Focusing on the combined regressions in column 8, we find that the voucher program distributes more benefits than the in-kind program to poorer households, those with more kids, and those with lower levels of education, but that widows are less likely to receive the program under the voucher program compared with the in-kind program.

C. *Administrative Mechanisms*

The previous section showed a striking finding: in the original Rastra system, there was a wide distribution of amount received by beneficiaries. In the voucher-based BPNT program, by contrast, virtually all beneficiaries who receive a transfer receive the full amount. What may explain this?

One thing to note is that information about who is entitled to benefits does not necessarily explain this phenomenon. In prior work (Banerjee et al. 2018), we investigated, in a similar program, the implications of the government sending identification cards to Rastra (then called Raskin) beneficiaries showing them what their rights were. That intervention—information alone—reduced leakage from the program and shifted the distribution of amount received to the right, but it did not result in a large point mass at the official eligibility amount. In fact, that intervention—providing official eligibility cards to the eligible in the context of the in-kind Rastra program—had, in fact, already been scaled up prior to the period we studied here, as the government mailed cards to all Rastra beneficiaries (over 15 million) starting in 2014.

There are, however, several important features that may matter. First, there is an important difference in the divisibility of what was given out to beneficiaries. In the in-kind program areas, sacks of rice were brought to local government warehouses, where they were then packaged into individual sized portions and delivered to beneficiaries. This manual process affords little ability for central government oversight. By contrast, the voucher program used individualized debit cards, with the intended recipient's name preprinted on the card (see online Appendix Figure 6; the name and debit card number are blurred on the photo to protect confidentiality). A debit card is not easily subdivided—the number of cards was fixed, and each card received the full transfer each month electronically, so local officials could not easily reallocate a part of their value to other beneficiaries.

Second, the administration of the program was switched from government officials to banks and private shops who signed up as remote bank agents. We collected data on the types of actors that were involved in the distribution of both programs through a phone survey to district governments. The results, in online Appendix

Table 15, show that the distribution points for the in-kind program in 88 percent of districts are government run, while in 99 percent of districts, the voucher program distribution sites are private bank agents.[21] Private agents may be less sensitive to political pressures than government officials to distribute aid to those other than beneficiaries. Moreover, the designation of private vendors as agents was run as part of a bank's remote bank agent program. Those programs are governed by banking regulations that also regulate other debit cards, which require banks to verify that debit cards are issued to the person who is listed on the account (i.e., know-your-customer rules), and require bank agents to verify that this person is indeed the person redeeming the card. The move from a program administered directly by local governments, to a program effectively administered by banks, may be a second reason for increased fidelity to distribution rules. It is worth noting that the involvement of banks in issuing and managing debit cards for government transfers is not unique to Indonesia; JP Morgan Chase, for example, issues and manages SNAP electronic benefit debit cards for US food assistance in 24 states, while other private debit card issuers manage them in all other states.[22]

Together, these changes may have changed the underlying bargaining process by changing the defaults. That is, by default in the new system the local government officials are not involved in the distribution unless they take an action to do so (since distribution occurs through a network of agents), and the default is that each individual receives an indivisible card. In Rastra, by contrast, distribution occurs through the village government, so local government officials are involved regardless, and the official needs to decide how much to allocate each person.

To illustrate how powerful this change can be, consider a simple Nash bargaining setup. The beneficiary is entitled to b from the program according to the eligibility rules. The village head, however, has a threat that he can impose a penalty X_i on beneficiary i (i.e., excluding her from village activities, and so on). This cost is socially wasteful. If the village head has bargaining weight α, then the beneficiary and village head share the surplus from agreement, i.e. the penalty X_i, with fixed bargaining weights. The beneficiary therefore receives a net transfer $b - (1 - \alpha)X_i$ and the village head keeps αX_i. Note that if the amount of penalty the village head can impose on various beneficiaries differs across beneficiaries i, the amount of net benefits received will also vary. There will also be, in general, no beneficiaries who end up with exactly b unless $X_i = 0$.

To capture the difference between the two programs, we assume that there is a fixed cost F that the village head needs to pay for each interaction with a villager. In the existing Rastra program, since the village head is forced to run the program anyway and hand out the rice, F is sunk, so this does not affect the allocation. Increasing information about the program, as in Banerjee et al. 2018 (which presumably reduces X_i; see Banerjee et al. 2018 for a model along these lines) shifts the distribution of

[21] Unfortunately, the SUSENAS did not ask recipients of the in-kind food subsidy program where they received their rice (due to the fact that it was all being distributed through the government). However, the results for location of where voucher recipients redeemed the voucher for food is confirmed in the SUSENAS data, where only 11 percent of BPNT recipients stated that they received it at the government offices.

[22] Data as of 2012; see http://www.g-a-i.org/wp-content/uploads/2012/10/GAI-Report-ProfitsfromPoverty-FINAL.pdf.

final outcomes, but does not affect the extensive margin of whether to negotiate at all, since the village head is still obligated to run the program, and F remains fixed.

In the switch to the BPNT program, however, the cost is no longer sunk, because by default the village head is not part of the distribution process. In particular, suppose that the village head first has to pay the cost F to begin discussions with the villager. At that point (i.e., once the village head starts discussions with a villager), F has been sunk, so the bargaining outcome is unaffected, so conditional on beginning negotiations, the village head would receive αX_i and the beneficiary would keep the rest, just as before. However, in this case, if $F > aX_i$, the village head will choose not to bother beginning negotiations, and individual i will keep the entire benefit b. Thus, with the new system, the distribution will feature a large point mass of beneficiaries receiving the full amount b, a gap just to the left of b, and then a distribution of lower amounts far to the left.

D. *Quality and Targeting*

In addition to how *much* assistance people received, there is also a question of the *quality* of the assistance that people received. The poor quality of rice under the in-kind program was a frequent complaint about the Raskin program, as well as other government-delivered programs more broadly (Jacoby 1997). Even though most rice consumed in developed economies like the United States is of high quality, in Indonesia, rice has many quality grades, and lower-quality rice often contains small stones, is off-color, has broken kernels, or smells bad, all of which are complaints that have surfaced about the Raskin program (Banerjee et al. 2019).

The SUSENAS household survey module asked respondents about the quality of that rice they received, on a three-point Likert scale, which we transformed to percentiles between 0 and 1. The question was asked identically for both the in-kind and voucher program recipients. On this scale, column 8 in Table 1 shows that households rated the voucher rice substantially higher quality, by about 32 percent.[23] This suggests that the quality-adjusted amount of subsidy received by targeted beneficiaries in voucher areas is even higher than the 46 percent increase discussed above. Putting a monetary value on this quality improvement is challenging, but a back-of-the-envelope calculation suggests that quality adjusting the values of subsidy received by targeted households in voucher areas would further increase the amount of the transfer received by about 20 percent.[24]

[23] While the quality of the rice the beneficiaries receive had increased, we do not expect this to lead to differential changes in the market premium for higher quality rice received by farmers. The reason is both because the subsidized rice in these programs is only 6 percent of the total rice market, and because the low-quality rice in the in-kind program may at least in part be the result of quality degradation in government warehouses. While we did not observe differential rice prices by quality level for each district, nationally, we see no difference in the price premium for high versus low quality rice over the period we studied here (e.g., it averaged Rp 646 per kilogram [7.4 percent] in 2016, before the transition began, and Rp 518 [5.6 percent] in 2019).

[24] To compute the quality adjusted price, we take the point estimate from the subsidized rice quality results and transform this into a z-score. That is, the quality improvements in Table 1 translate to an increase of 0.71 standard deviations on the quality measure. We then calculate the rice price distribution paid by residents within each district-urban area, which may capture heterogeneity in rice quality. Finally, we assign a monetary value of the quality improvements equal to that number of standard deviations of the rice price (i.e., if the standard deviation of rice price is Rp 1,000 per kilogram, and the quality measure in Table 1 represents 0.25 standard deviations, we would assign a price equivalent of Rp 250 per kilogram). On average, this calculation suggests that the quality improvements in Table 1 would be worth an additional Rp 1,060 per kilogram, which is anecdotally consistent

The quality improvements in the voucher program are likely related to the fact that the sourcing of the rice was more flexible. In the in-kind program, all rice was sourced and distributed by BULOG, the government logistics agency. By contrast, in the voucher program, agents were largely able to source rice where they see fit—which resulted in a very different sourcing pattern. Specifically, using data from a survey of agents conducted by the World Bank, we find that in the voucher program, 7 percent of agents reported sourcing their rice only from the government (BULOG), 27 percent report a mix of BULOG and the private market, and 65 percent reporting sourcing it entirely from the private market.

The improvement in the quality of the assistance makes the improved targeting even more remarkable. A common argument in favor of providing lower-quality food assistance is that the low quality serves as a screening device, so that it can be better targeted to the poor (Nichols and Zeckhauser 1982; Besley and Kanbur 1988). What is remarkable here, however, is that targeting improved *despite* the food being substantially higher quality, which would give the rich more reason to try to be included in the program. In this case, the improved administrative process of the voucher program swamped whatever targeting benefits may have been accruing from lower-quality in-kind rice.

E. *Impacts on Poverty*

The previous analysis showed that the vouchers channeled additional resources to the poorest households. We next ask whether this effect was large enough to affect household poverty rates.

To examine this, for each household, we compute whether they are below the official Government of Indonesia per capita consumption poverty line, which is set separately for rural and urban households in each province. Since richer households are largely above the poverty line to begin with, one might expect stronger results for those households who have lower PMT scores.

We therefore begin by presenting the results nonparametrically in Figure 4, using the same bin-scatters as above. As discussed before, baseline PMT score is strongly predictive of poverty status, in both voucher and in-kind areas. About 27 percent of households with PMT scores 0–5 are below the poverty line in in-kind areas. This falls to 21 percent for households with PMT scores 5–10, about 16 percent for households in the next few PMT bins, and further at higher levels.

Figure 4 shows substantial reductions in poverty associated with the move from the in-kind program to the voucher program. This is particularly apparent for the poorest households—poverty rates fall by about 6.5 percentage points for households in the poorest group (baseline PMT score 0–5) and about 4 percentage points for households in the next-poorest group (baseline PMT score 5–10). This occurs both because more of these households were close to the poverty line to

with typical differences in market prices between low and average quality rice. Online Appendix Table 3, panel B shows that, when we used a quality-adjusted price compared to a fixed price, the benefits of the voucher program for households with PMT $\leq$ 30 are about 39 percent higher.

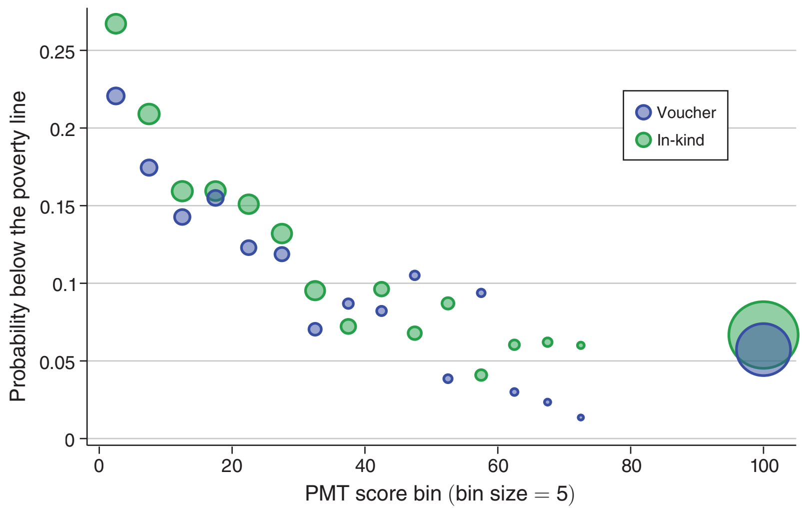

Figure 4. Relationship between PMT Score and Poverty Status

Notes: This graph provides the relationship between whether the household is below the poverty line and PMT score, by treatment status. PMT scores are binned in groups of five, with those who have no PMT score grouped with those with a score of 100. Markers are scaled by the number of households in each bin. Data on outcomes are from the March 2019 SUSENAS, while PMT data are from the UDB.

begin with, and also because—as shown above—the impact of the transfer was the largest for these groups.[25]

Table 2 shows the results for poverty in regression form. We show results for the full sample (column 1), for households with PMT scores ≤ 30 at baseline (column 2), and for households with PMT scores > 30 at baseline (column 3). In the remaining columns, we zoom in further on the poor, restricting to households with PMT ≤ 25, ≤ 20, and so on.

The results confirm substantial reductions in poverty for those households with low PMT scores. For the full sample of all households with PMT scores ≤ 30 at baseline, the share in poverty in voucher areas fell by 12.7 percent (2.3 percentage point reduction from 18 percent in in-kind areas, $p = 0.134$; column 2). As we focus our attention on households with lower PMT scores, we see (as expected) higher poverty rates in the in-kind areas, and substantially larger reductions in poverty in the voucher areas. For example, for households with PMT scores ≤ 15 at baseline—which approximately corresponds to the bottom 15 percent of the population (slightly higher than the overall poverty rate in the entire sample), we find a reduction in poverty of 20 percent (4.3 percentage point reduction from a mean of 21.0 percent in voucher areas, $p = 0.028$). For the very poorest—those with PMT scores 0–5 at baseline—we see a 24 percent reduction in the share of households in

[25] Online Appendix Tables 16 and 17 further break down the impact on total subsidy received and rice quality by subgroup as we zoom in on poorer households, as defined by baseline PMT status. As shown in online Appendix Table 16, households with PMT scores 0–5 saw their mean subsidy increase by Rp 19,648 in voucher areas compared to control ($p < 0.001$), compared to Rp 13,496 for the entire group with PMT ≤ 30.

TABLE 2—EXPERIMENTAL DIFFERENCE BETWEEN VOUCHER AND IN-KIND DISTRICTS BEING BELOW THE POVERTY LINE

	All (1)	PMT $\leq$ 30 (2)	PM $\leq$ 25 (3)	PMT $\leq$ 20 (4)	PMT $\leq$ 15 (5)	PMT $\leq$ 10 (6)	PMT $\leq$ 5 (7)
Voucher	−0.009	−0.023	−0.025	−0.034	−0.043	−0.052	−0.065
	(0.008)	(0.015)	(0.016)	(0.017)	(0.018)	(0.020)	(0.024)
	[0.205]	[0.134]	[0.166]	[0.078]	[0.028]	[0.020]	[0.012]
Observations	66,496	16,329	13,707	11,072	8,307	5,529	2,788
Stratum fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Double lasso	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DV mean (control)	0.098	0.180	0.189	0.198	0.210	0.237	0.267

Notes: This table explores the impact of being below the consumption poverty line, by PMT groupings. The outcome variable, below poverty line, is an indicator for whether a household is below the poverty line in its province by urban/rural area, as measured by per capita consumption. The outcome data come from the March 2019 SUSENAS; the PMT data come from the UDB. We used a double LASSO to choose the control variables (all potential variables used as inputs for the LASSO are listed in online Appendix B). Standard errors are clustered at the district (*kabupaten*) level and displayed in parentheses. Randomization inference *p*-values are from 1,000 permutations of the treatment assignments and are displayed in brackets.

poverty (6.5 percentage points, $p = 0.012$).[26] We find broadly similar results when we examine other utility metrics such as log consumption, or constant relative risk aversion utility of consumption with $\rho = 2$ or $\rho = 3$ (online Appendix Table 20).

The poverty impacts are not driven by price changes associated with the switch from in-kind to vouchers. To investigate this, we recompute household consumption—instead of using the household's reported value of rice and eggs consumed, we use the household's reported *quantities* of rice and eggs consumed, which we then value at a common, fixed set of prices for each island $\times$ rural/urban area, that is, using the same prices for these goods for both in-kind and voucher districts. We then calculate poverty rates based on this new consumption measure that holds rice and egg prices fixed. The results, shown in online Appendix Table 21, are virtually identical to the results shown in Table 2.

In combination, the results here suggest that the improved targeting of assistance in the voucher program had a substantial impact: substantially more assistance to targeted households, particularly for the very poor, which in turn led to a substantial reduction in the share of these households below the poverty line.

III. Alternative Mechanisms

While the administrative features of the programs differed, there could nonetheless be substantial welfare effects due to more classic price-theoretic mechanisms associated with the switch from in-kind to vouchers. Thus, we explore two additional channels below—consumption choices and price effects.

[26] Online Appendix Table 18 also shows similar patterns to the poverty results when we examine the poverty gap or poverty gap squared. We find no observable impact on the food insecurity index (online Appendix Table 19).

A. *Consumption Choices*

We now turn to understanding how the move from in-kind subsidies to vouchers affected the patterns of consumption choices. Recall that the in-kind subsidy program only provided rice, where the voucher program allowed for more flexibility by allowing households to also purchase eggs if they so choose. The government made this change as it wanted to encourage protein consumption in addition to starches. Does this increased flexibility translate into changes in the patterns of consumption? And if so, how large are these relative to the overall increase in subsidy households receive due to better fidelity to program design in the voucher program?

Before we turn to the experimental results, we first examine consumption patterns in the in-kind areas. Theory would predict that if households are already consuming more rice than they would receive in the transfer, this would imply they were unconstrained by the fact that the in-kind transfer was limited to rice. As such, one would expect that the switch from the in-kind transfer to the more flexible voucher (where they can also choose to purchase eggs) would not affect consumption patterns as long as the amount of the transfer is the same. That is, under a standard consumption model, if a household was consuming more than 10 kg of rice per month, it was doing so not because of the in-kind transfer of 10 kg rice per month per se, but rather because its optimal consumption of rice was above 10 kg. As such, one would expect households to consume the same bundle regardless of whether they received the transfers of rice in-kind or the more flexible vouchers. The amount of rice consumed is shown in online Appendix Figure 7. Panel A plots out rice consumption in the in-kind areas, and it shows that almost everyone is consuming more than 10 kg of rice. Examining households with PMT scores at or below 30 (panel B), we find similar patterns; in fact, only 3.3 percent of these households consume less than 10 kg of rice.[27] Thus, we would expect that the switch from in-kind to more flexible vouchers would not mechanically affect consumption patterns, as virtually all households are unconstrained even under the in-kind program.

Turning to our experimental results, we examine the effect of the conversion to the voucher program on consumption of both subsidized rice and eggs (Table 3, panel A), and *total* monthly consumption of rice and eggs (Table 3, panel B). These are measured separately in the survey. Information on subsidized consumption of rice and eggs come from questions about the Rastra and BPNT programs specifically. Information about total consumption comes from a completely separate module in the SUSENAS which collected information on the consumption of the food products in question (as well as over 200 other items of consumption) over the previous week from all sources. In both cases, for eggs, we report the total egg protein consumption in grams, which is the Indonesian government's aggregated estimate totaling all different types of eggs consumed (e.g., farmed chicken eggs, free-range chicken eggs, duck eggs, quail eggs).

[27] Moreover, in practice the in-kind program did not distribute 10 kg to beneficiaries—as discussed above, the typical beneficiary received only 3.9kg of rice—and over 99 percent of households with PMT scores ≤ 30 consumed more than this amount of rice. These patterns are similar among the poorest of the poor: for example, among households with PMT score of less than 5, we find that households consume about 36 kg of rice on average (online Appendix Table 22)

Table 3—Experimental Difference between Voucher and In-Kind Districts on Food Consumption

Panel A. Subsidized food consumption

	Subsidized rice (kilograms)			Subsidized egg protein (grams)		
	All	PMT $\leq$ 30	PMT $>$ 30	All	PMT $\leq$ 30	PMT $>$ 30
	(1)	(2)	(3)	(4)	(5)	(6)
Voucher	−0.300	0.062	−0.424	10.932	32.719	3.362
	(0.066)	(0.205)	(0.058)	(1.534)	(4.648)	(0.463)
	[0.002]	[0.773]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	66,495	16,328	49,566	66,423	16,270	49552
Stratum fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Double lasso	Yes	Yes	Yes	Yes	Yes	Yes
DV mean (control)	1.494	2.987	0.957	0.140	0.484	0.015

Panel B. Total food consumption

	Total rice (kilograms)			Total egg protein (grams)		
	All	PMT $\leq$ 30	PMT $>$ 30	All	PMT $\leq$ 30	PMT $>$ 30
	(1)	(2)	(3)	(4)	(5)	(6)
Voucher	−0.012	−0.411	0.143	3.454	9.279	0.566
	(0.314)	(0.478)	(0.304)	(3.110)	(4.750)	(3.781)
	[0.971]	[0.492]	[0.704]	[0.354]	[0.100]	[0.891]
Observations	66,496	16,329	49,566	66,483	16,327	49,555
Stratum fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Double Lasso	Yes	Yes	Yes	Yes	Yes	Yes
DV Mean (Control)	27.627	31.586	26.170	226.384	213.652	230.738

Notes: This table examines consumption of rice (columns 1–3) and egg protein from all types of eggs (columns 4–6). Panel A is any consumption from Rastra or BPNT, while panel B is total consumption. For continuous outcome variables, we drop any value greater than 12 standard deviations from the mean. The outcome data come from the March 2019 SUSENAS; the PMT data come from the UDB. We used a double LASSO to choose the control variables (all potential variables used as inputs for the LASSO are listed in online Appendix B). Standard errors are clustered at the district (*kabupaten*) level and displayed in parentheses. Randomization inference *p*-values are from 1,000 permutations of the treatment assignments and are displayed in brackets.

Looking first at the bundle of subsidized food households receive, we see no change in consumption of subsidized rice for households with PMT at or below 30 (panel A, column 2). We do, however, observe that targeted voucher households consume more of their subsidy in the form of eggs. This is, of course, virtually zero in in-kind areas, but increases to about 33 grams per month of egg protein for households with PMT scores $\leq$ 30 ($p < 0.001$), and by a much smaller amount (but still positive) for households with PMT scores $>$ 30.

The key question is what happens to the patterns of *total* consumption of specific types of food, which we explore in panel B. We find no impact on total rice consumption for either of these samples.[28] However, we do observe that eligible voucher households consume more total egg protein. Specifically, for households with PMT scores $\leq$ 30, we observe an increase in total egg protein of about 4.3 percent (9.3 grams, *p*-value 0.10). Combined with the estimates in panel A, this implies that about 28 percent of the increase in subsidized egg protein consumed represents a net increase in consumption.

[28] Online Appendix Table 22 replicates the rice consumption results by different cuts of PMT score, i.e., different poverty scores. We find no observable difference in rice consumption even in the poorest groups (i.e., PMT score less than 10 or less than 5).

Online Appendix Table 23 replicates the egg consumption results, by different cuts of the PMT score, i.e., different poverty levels. Total egg protein consumption increases the most for the poor; for example, those with PMT score below 10 experience an 8.4 percent increase ($p = 0.053$) while those with PMT score below 5 experience an 11.7 percent increase ($p = 0.033$). The "stickiness" of the transfer increases as we look at poorer households—for households with PMT $\leq$ 10, the change in total egg consumption is about 45 percent of the change in subsidized egg protein. For the very poorest households—those with PMT $<$ 5—fully 61 percent of the increase in subsidized egg protein represents a change in net consumption. These results imply that flexibility with the vouchers—combined with the labeling for eggs—may substantially affect real consumption decisions.

An important question is whether the change in egg consumption represents an income effect for the set of households who received the larger subsidy, as opposed to some specific "stickiness" coming from the set of goods included in the voucher. To examine this, we explore consumption of a variety of other food items, as well as "temptation goods" such as cigarettes and alcohol (online Appendix Table 24). We do not observe systematic increases in other food consumption (p-value of the joint test $=$ 0.223); if anything, there is a small reduction in salt consumption. We find no observable change in cigarette consumption nor alcohol consumption. This suggests the results we find on egg consumption really are coming from the fact that eggs, and only eggs, were added to the set of goods that the transfer could be used on, rather than a more general income effect.

We also examine another potential feature of vouchers—they let households change their consumption in response to local prices. Specifically, we compute, for those households who have PMT $>$ 30 (i.e., those who are unlikely to be eligible for vouchers) the average log unit price paid for eggs and for rice in each district $\times$ urban/rural cell. We then take the households with PMT $\leq$ 30 (i.e., a different sample of households) and look at how the share of voucher spent on rice depends on the rice price, with different levels of geographic fixed effects to control for differential demand. The one important caveat here is that we only have cross-sectional variation in prices, so this could also be picking up local demand differences. Nevertheless, the results are suggestive of some substitution. Online Appendix Table 25 suggests that households use a greater share of their voucher for rice in districts where egg prices are high (see columns 2–5, which include different types of geographic fixed effects). This suggests that households are responsive to local prices in their relative consumption decisions with the vouchers. But overall, it is also important to note that the changes we observe in this section are relatively small in magnitude compared to the 46 percent increase in *total* subsidy eligible households receive due to the increased fidelity to program design.

B. *Prices*

Changes in prices that could occur as a result of the switch could also affect the observed poverty levels. While we might expect that, if anything, the differences in prices would go in the direction of higher levels of poverty under the voucher, we explore the impact of the switch on prices below. One benefit of this unique policy experiment is that given that the unit of randomization was at a high geographic

Table 4—Experimental Difference between Voucher and In-Kind Districts on Price

				Measures of isolation			
	Main effect only (1)	Above median supply shock (2)	Above 75th pct. supply shock (3)	Nonasphalt road (4)	Road not always passable (5)	Above median time to district capital (6)	Above 75th pct. time to district capital (7)
Voucher	129.282	65.998	50.561	118.131	124.157	51.898	50.316
	(130.238)	(172.217)	(133.334)	(140.103)	(131.202)	(145.295)	(137.099)
	[0.309]	[0.708]	[0.702]	[0.372]	[0.322]	[0.711]	[0.696]
Voucher × [*variable*]		172.624	539.234	52.554	195.171	151.902	333.929
		(257.483)	(475.145)	(125.464)	(179.316)	(116.163)	(136.253)
		[0.530]	[0.138]	[0.677]	[0.317]	[0.233]	[0.027]
Observations	32,343	32,343	32,343	32,334	32,334	32,334	32,334
Stratum fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main effect included	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Double lasso	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DV mean (control)	9,478.508	9,478.508	9,478.508	9,478.508	9,478.508	9,478.508	9,478.508
[Variable] mean		0.540	0.238	0.137	0.035	0.489	0.236

Notes: This table examines the impact of the vouchers on market rice prices. Data are from the March 2019 SUSENAS, taken from households that are not in the UDB. Measures of isolation data come from the 2018 PODES. Above median and above 75th pct. supply shock indicate whether the district has above median or seventy-fifth percentile subsidized rice as a fraction of total rice consumption in the district, respectively. Nonasphalt road indicates whether the roads connecting the village to others are unpaved. Road not always passable indicates whether these roads are impassable at some point during the year. Above median and above 75th pct. time to district capital indicate whether the village's travel time to the nearest district capital is above the median or seventy-fifth percentile, respectively. Standard errors are clustered at the district (*kabupaten*) level and displayed in parentheses. Randomization inference p-values are from 1,000 permutations of the treatment assignments and are displayed in brackets.

level—districts, with an average population of about 500,000—we can measure the general equilibrium effects of the switch to vouchers on overall prices in the district. Table 4 presents these results. We obtain rice prices from the SUSENAS consumption module; we focus on prices reported by households not in the UDB and thus not eligible for the subsidy programs.[29]

In column 1, we first begin by looking at the overall impact of the voucher program on the price of rice.[30] Standard price theory predicts that reducing the subsidized rice supply and instead increasing demand for private rice in shops could increase the price of rice more broadly, though of course the magnitude of this would depend on the elasticity of supply of rice. While the coefficient is indeed positive, it is small in magnitude and not significant (p-value of 0.309).

We then turn to two forms of heterogeneity as laid out by theory. First, theory may predict that in areas where the subsidized rice consumption is a large share of total rice consumed in the district, prices may adjust more, because the switch from in-kind to vouchers represents a larger negative supply shock in these locations (Filmer et al. 2018). We measure the size of the supply shock by dividing the

[29] We focus here on prices experienced by those not eligible in order to avoid compositional effects if *where* households purchase their rice is affected by the voucher treatment. Online Appendix Table 26 shows the results from the full sample instead. They are qualitatively similar, though the results in column 7 are no longer statistically significant on the full sample.

[30] We also examine egg prices in online Appendix Table 27. Voucher areas have about 2.0 percent higher egg prices (p-value of 0.106). We do not observe heterogeneity by the supply shock or by any of the measures of remoteness that we consider.

total amount of subsidized rice allocated in the in-kind program by an estimate of the total amount of rice consumed in the district from all sources, which we obtain from the SUSENAS household survey. In column 2, we find no observable differential effect of treatment in districts that are above the median in terms of this supply shock measure. In column 3, we examine districts that are at the seventy-fifth percentile and higher in this supply shock measure; we observe a larger effect in magnitude than for those above median (Rp 539 rather than Rp 173), but it remains statistically insignificant (p-value of 0.138) and still relatively small in magnitude (representing a 5.7 percent price increase).[31]

Second, theory would predict that areas that are more remote may have larger price adjustments, if these more isolated markets have less elastic supply (Cunha, De Giorgi, and Jayachandran 2019). We examine several measures of remoteness from PODES, which we can measure at the village level. We first examine whether a village lacks an asphalt road or whether the road is not always passable in columns 4 and 5, respectively. In both cases, the coefficient is positive, but small and statistically insignificant (p-values of 0.677 and 0.317, respectively).

Next, we examine the time it takes to reach the district capital. We first graph the treatment effects of being in a voucher district by a village's percentile time distance to the capital estimated using a series of locally weighted regressions in Figure 5; we also plot 95 percent confidence intervals.[32] The graphs show no impact on prices for those locations closest to the district capital, but a positive slope as one moves away—that is, in remoter areas, the market price of rice goes up more in voucher areas relative to in-kind areas. However, the magnitudes shown are relatively small, and the confidence intervals suggest that we cannot rule out differences from zero for those locations below the ninetieth percentile in distance.

Turning back to Table 4, we examine the price impacts on villages that are above the median (column 6) and above the seventy-fifth percentile (column 7) of time to reach the district capital. In above median districts, the effect is positive but small (1.6 percent) and insignificant (p-value of 0.233). For those that are very remote (i.e., seventy-fifth percentile or higher), we observe a Rp 334 (3.5 percent) increase in voucher areas compared to in-kind areas (p-value of 0.027), with no price change in the remaining 75 percent of locations. However, even the 3.5 percent price increase observed in remote locations is not enough to negate the benefits gained from the greater concentration of the program: households on average purchase about 19.5 kg of rice from the private market per month from all sources, which implies a Rp 6,512 increase in overall rice spending in these remote areas, compared to the Rp 13,496 increase in benefits received by likely eligible households.[33]

[31] We also examine these heterogeneity measures using alternative specifications, either examining continuous measures (online Appendix Table 28). We observe similar conclusions.

[32] In online Appendix Figure 8, we plot Figure 5 with time to distance capital rather than percentiles, and find similar results.

[33] Moreover, as discussed above, if we hold prices constant, and value the quantities received from both the in-kind and voucher program using fixed prices, our main results are qualitatively similar. See online Appendix Table 3, panel A.

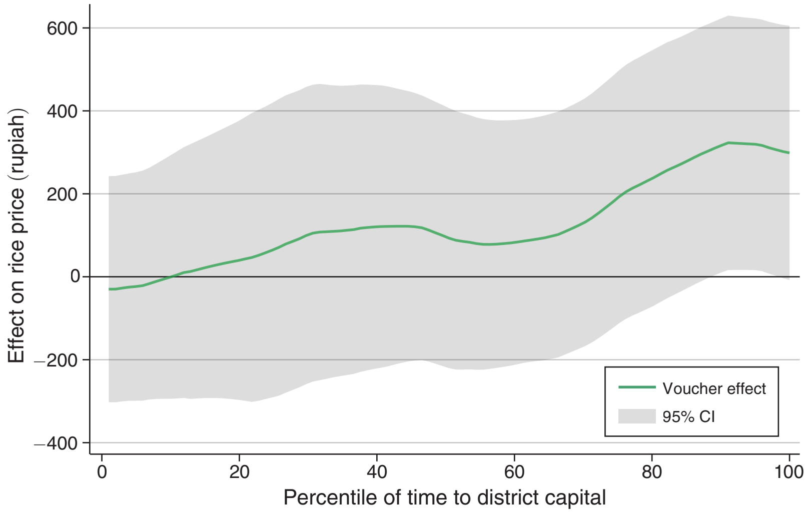

FIGURE 5. NONPARAMETRIC HETEROGENEOUS TREATMENT EFFECTS ON RICE PRICE BY TIME TO DISTRICT CAPITAL

Note: This graph investigates the relationship between the effect of the voucher on rice prices and village-level travel time to the nearest district capital. Village-level percentile of time to the nearest district capital is plotted on the *x*-axis, and treatment effects are plotted on the *y*-axis. Rice price is calculated from households not in the UDB. Regressions are estimated using a triangular kernel and a bandwidth of four. Data are from the March 2019 SUSENAS.

IV. Leakage and Administrative Program Costs

Finally, we explore the impact of the program on both leakage rates and the administrative costs of running the program.

A key criticism of the administration of in-kind programs is that the complexity of the food distribution makes it hard to monitor, and so much of the food can "fall off the truck" (Banerjee et al. 2018). In Table 5, we examine whether the vouchers reduced such leakage, by comparing the actual subsidy received by households as measured by the SUSENAS survey with the intended subsidy for each district. To be clear, in this section leakage is defined as the difference between total amount disbursed by the government and the total amount of subsidy received by *any* household, regardless of their official eligibility status. This measures the impact on the total amount missing, as opposed to changes in *who* receives the assistance, which was discussed in previous sections.

We compute the fraction of subsidy received by households in three ways. In column 1, we calculate subsidy as the sum of reported in-kind and voucher values received. In column 2, we compute the value of the voucher subsidy based on the amount of eggs or rice purchased times the market price of rice or eggs. If local agents were inflating prices charged, so that voucher households were receiving less "real" assistance than the nominal amount of subsidy would suggest, this would not be counted as leakage in column 1, but would show up as a form of leakage in

Table 5—Experimental Difference between Voucher and In-Kind Districts on Leakage

	Subsidy received/ intended subsidy (1)	Subsidy received (market prices)/ intended subsidy (2)	Subsidy received (quality-adjusted)/ intended subsidy (3)
Voucher	−0.018	−0.059	−0.013
	(0.031)	(0.029)	(0.031)
	[0.583]	[0.055]	[0.708]
Observations	105	105	105
Stratum fixed effects	Yes	Yes	Yes
Double lasso	Yes	Yes	Yes
DV mean (control)	0.587	0.586	0.588

Notes: In this table, we examine the subsidy received relative to the intended subsidy to disburse, by treatment status. The data come from the March 2019 SUSENAS and from administrative data. To compute the intended subsidy in an in-kind district, we multiply the number of beneficiaries in the district by the 10 kg rice disbursement and the official procurement price of Rastra rice; in the voucher districts, we multiply the BPNT beneficiaries in the district by the disbursement amount (Rp 110,000). We calculate subsidy received in three ways: in column 1, it is the sum of the value of any program received; in column 2, we adjust the voucher disbursement by the market price of rice in the area; in column 3, we adjust the voucher disbursement by the market price of higher quality rice. We used a double LASSO to choose the control variables (all potential variables used as inputs for the LASSO are listed in online Appendix B). Standard errors are clustered at the district (*kabupaten*) level and displayed in parentheses. Randomization inference *p*-values are from 1,000 permutations of the treatment assignments and are displayed in brackets.

column 2. Finally, in column 3, we adjust the market price of rice in BPNT to reflect the higher quality of rice bought to allow for quality weighted comparisons.[34]

In all cases, we compare this to the value of the intended subsidy received. We compute intended subsidy for the in-kind districts by multiplying the official number of targeted beneficiaries by 10 kg of rice and then by the official procurement price of the Rastra Rice. To compute the comparable measure for the voucher districts, we multiplied the number of official targeted beneficiaries by the Rp 110,000 payment.

Overall leakage by these measures is high, but since it is possible that the level of this metric is affected by aggregate reporting issues (Olken 2007), what is more relevant is whether these leakage metrics *change* in response to the switch, holding measurement constant. We find that it does not. As shown in column 1, the conversion to the voucher program has no impact on the share of the intended subsidy received by households. If anything, as shown in column 2, allowing prices to adjust by area, the voucher program led to a decrease in total subsidy received, significant at the 10 percent level. Finally, adjusting for quality of rice received, we also observed negative, but insignificant impacts of the conversion on leakages.[35] The lack of an impact on total leakage is in contrast to the results in Banerjee et al. (2018), which did find leakage impacts; here, the results are clearly driven by changes in *who* receives assistance, not the *total* amount of assistance received.

[34] To compute the quality adjusted price, we take the point estimate from the subsidized rice quality results and transform this into a z-score. Then, we take the same z-score in the rice price distribution of each district-urban area.

[35] It is a priori possible that overall leakages fall in districts that were better prepared to convert from in-kind to vouchers (e.g., had a more established agent network, etc.). In online Appendix Table 29, we therefore estimate the effect of treatment on leakages by whether the district scored higher or lower on the baseline readiness index. We do not find observable evidence of heterogeneity by baseline readiness.

We also examine the impact of the switch from in-kind transfers to vouchers on the administrative costs of running the program. We start by estimating the program costs for delivery of in-kind benefits. We obtain the operating costs for the in-kind program from BULOG's annual report. Note, however, that the national government only facilitates delivery to the district or subdistrict capital, and local governments are required to cover the costs of their own pick-up and delivery; we obtain an estimate of these costs through a survey of the distributors (Banerjee et al. 2019). These costs are summarized in online Appendix Table 30. On net, the administrative program costs for the in-kind delivery are about 4.1 percent of the total benefits.

The administrative costs of the voucher program are much less than the in-kind program. There are two main costs. First, there is a cost of printing the debit card (Rp 12,500 per person), which we assume lasts three years since targeting is conducted every three years. Second, there is the monthly rental cost of the debit-card reader machines for the agents. Assuming that the agents only process vouchers—and do not use the machines for other financial transactions—online Appendix Table 30 shows that the total administrative costs would be 2.1 percent of the benefits dispersed, or about half the size of the administrative costs incurred for in-kind benefits. Note, however, that a large fraction of the agents existed prior to the voucher program, and paid rent on the machines for other transitions, and so the marginal cost for them of processing the vouchers is minimal; in fact, evidence from a field experiment reported in Banerjee et al. (2021) shows that about 77 percent of agents already existed prior to the program. Under the assumption that all existing agents would have had the machines in any case, but all new agents use the machines only for the voucher program, the administrative costs would be 0.74 percent of the benefits disbursed or about 17 percent of administrative cost ratio for the in-kind benefits.

V. Conclusion

In this paper, we examine the results from an at-scale experiment to study the switch from an in-kind social assistance program to an electronic voucher-based program. The vouchers concentrated the aid toward poorer, targeted households: households with low baseline predicted poverty scores—i.e., those poor or near-poor households who were the intended beneficiaries of both programs—received, on net, 46 percent more assistance in voucher areas compared to in-kind areas. This led to a reduction in poverty. For those households in the bottom 15 percent at baseline, the poverty rate fell by 20 percent.

We posit that despite the fact that, in theory, both programs give about the same amount of aid, the administrative differences across these programs led to large changes in the lives of the poor. While we observe some differences in program outcomes based on more "classic" price-theory differences between in-kind and food stamps programs—e.g., consumption choices, food quality, and prices—these effects are small relative to the overall main result, namely, that the increased fidelity to program design from the vouchers delivers 46 percent more subsidized assistance to eligible households compared to the in-kind program. Finally, note that these gains in poverty reduction occurred despite the fact that the costs of operating the voucher-based program were less than half that of the in-kind program, conferring another administrative benefit.

On net, the results here suggest an important additional dimension for how to think about vouchers versus in-kind programs. While the economics literature has focused primarily on outcomes grounded in price-theory—such as changes in the consumption bundle caused by either the mechanical budget constraint or mental accounting, or aggregate price changes—our results here suggest that the differences in how these programs are *administered* may be a far more important determinant of their relative effectiveness, at least in settings with limited administrative capacity.

REFERENCES

Aker, Jenny C. 2017. "Comparing Cash and Voucher Transfers in a Humanitarian Context: Evidence from the Democratic Republic of Congo." *World Bank Economic Review* 31 (1): 44–70.

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

Alderman, Harold, Ugo Gentilini, and Ruslan Yemtsov. 2018. *The 1.5 Billion People Question: Food, Vouchers, or Cash Transfers?* Washington, DC: World Bank.

Banerjee, Abhijit, Esther Duflo, Amy Finkelstein, Lawrence F. Katz, Benjamin A. Olken, and Anja Sautmann. 2020. "In Praise of Moderation: Suggestions for the Scope and Use of Pre-analysis Plan for RCTs in Economics." NBER Working Paper 26993.

Banerjee, Abhijit, Rema Hanna, Jordan Kyle, Benjamin A. Olken, and Sudarno Sumarto. 2018. "Tangible Information and Citizen Empowerment: Identification Cards and Food Subsidy Programs in Indonesia." *Journal of Political Economy* 126 (2): 451–91.

Banerjee, Abhijit, Rema Hanna, Jordan Kyle, Benjamin A. Olken, and Sudarno Sumarto. 2019. "Private Outsourcing and Competition: Subsidized Food Distribution in Indonesia." *Journal of Political Economy* 127 (1): 101–37.

Banerjee, Abhijit, Rema Hanna, Benjamin A. Olken, and Sudarno Sumarto. 2020. "The (Lack of) Distortionary Effects of Proxy-Means Tests: Results from a Nationwide Experiment in Indonesia." *Journal of Public Economics Plus* 1: 1–9.

Banerjee, Abhijit, Rema Hanna, Benjamin A. Olken, Elan Satriawan, and Sudarno Sumarto. 2021. "Remote Banking and Financial Access: Evidence from a Randomized Experiment on Government-to-People Payments in Indonesia." Unpublished.

Banerjee, Abhijit, Rema Hanna, Benjamin A. Olken, Elan Satriawan, and Sudarno Sumarto. 2023. "Replication Data for: Electronic Food Vouchers: Evidence from an At-Scale Experiment in Indonesia." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E167262V1.

Basurto, Maria Pia, Pascaline Dupas, and Jonathan Robinson. 2020. "Decentralization and Efficiency of Subsidy Targeting: Evidence from Chiefs in Rural Malawi." *Journal of Public Economics* 185: 104047.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014 "Inference on Treatment Effects after Selection among High-Dimensional Controls." *Review of Economic Studies* 81 (2): 608–50.

Besley, Timothy, and Ravi Kanbur. 1988. "Food Subsidies and Poverty Alleviation." *Economic Journal* 98 (392): 701–19.

Cahyadi, Nur, Rema Hanna, Benjamin A. Olken, Rizal Adi Prima, Elan Satriawan, and Ekki Syamsulhakim. 2020. "Cumulative Impacts of Conditional Cash Transfer Programs: Experimental Evidence from Indonesia." *American Economic Journal: Economic Policy* 12 (4): 88–110.

Coate, Stephen, Stephen Johnson, and Richard Zeckhauser. 1994. "Pecuniary Redistribution through In-Kind Programs." *Journal of Public Economics* 55 (1): 19–40.

Cunha, Jesse M. 2014. "Testing Paternalism: Cash versus In-Kind Transfers." *American Economic Journal: Applied Economics* 6 (2): 195–230.

Cunha, Jesse M., Giacomo De Giorgi, and Seema Jayachandran. 2019. "The Price Effects of Cash Versus In-Kind Transfers." *Review of Economic Studies* 86 (1): 240–81.

Currie, Janet, and Firouz Gahvari. 2008. "Transfers in Cash and In-Kind: Theory Meets the Data." *Journal of Economic Literature* 46 (2): 333–83.

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

Filmer, Deon, Jed Friedman, Eeshani Kandpal, and Junko Onishi. 2018. "Cash Transfers, Food Prices, and Nutrition Impacts on Nonbeneficiary Children." World Bank Policy Research Working Paper 8377.

Gadenne, Lucie, Samuel Norris, Monica Singhal, and Sandip Sukhtankar. 2021. "In-Kind Transfers as Insurance." NBER Working Paper 28507.

Galasso, Emanuela, and Martin Ravallion. 2005. "Decentralized Targeting of an Antipoverty Program." *Journal of Public Economics* 89 (4): 705–27.

Gentilini, Ugo. 2016. "Revisiting the 'Cash versus Food' Debate: New Evidence for an Old Puzzle?" *World Bank Research Observer* 31 (1): 135–67

Hastings, Justine, and Jesse M. Shapiro. 2018. "How Are SNAP Benefits Spent? Evidence from a Retail Panel." *American Economic Review* 108 (12): 3493–3540.

Hidrobo, Melissa, John Hoddinott, Amber Peterman, Amy Margolies, and Vanessa Moreira. 2014. "Cash, Food, or Vouchers? Evidence from a Randomized Experiment in Northern Ecuador." *Journal of Development Economics* 107: 144–56.

Hirvonen, K., and J. Hoddinott. 2020. "Beneficiary Views on Cash and In-Kind Payments: Evidence from Ethiopia's Productive Safety." World Bank Policy Research Working Paper 9125.

Jacoby, Hanan G. 1997. "Self-Selection and the Redistributive Impact of In-Kind Transfers: An Econometric Analysis." *Journal of Human Resources* 32 (2): 233–49.

Jiménez-Hernández, Diego, and Enrique Seira. 2021. "Should the Government Sell You Goods? Evidence from the Milk Market in Mexico." Unpublished.

Leroy, Jef L., Paola Gadsden, Teresa González de Cossío, and Paul Gertler. 2013. "Cash and In-Kind Transfers Lead to Excess Weight Gain in a Population of Women with a High Prevalence of Overweight in Rural Mexico." *Journal of Nutrition* 143 (3): 378–83.

Londoño-Vélez, Juliana and Pablo Querubin. 2022. "The Impact of Emergency Cash Assistance in a Pandemic: Experimental Evidence from Colombia." *Review of Economics and Statistics* 104 (1): 157–65.

Ministry of Social Affairs. 2018. *Pedoman Umum Subsidi Rastra*. Jakarta, Indonesia: Kementerian Koordinator Bidang Pembangunan Manusia dan Kebudayaan.

Muralidharan, Karthik, and Paul Niehaus. 2017. "Experimentation at Scale." *Journal of Economic Perspectives* 31 (4): 103–24.

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

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2022. "Identity Verification Standards in Welfare Programs: Experimental Evidence from India." Unpublished.

Nichols, Albert L., and Richard J. Zeckhauser. 1982. "Targeting Transfers through Restrictions on Recipients." *American Economic Review* 72 (2): 372–77.

Olken, Benjamin A. 2006. "Corruption and the Costs of Redistribution: Micro Evidence from Indonesia." *Journal of Public Economics* 90 (4-5): 853–70.

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200–49.

World Bank. 2018. *The State of Social Safety Nets 2018*. Washington, DC: World Bank.

Young, Alwyn. 2019. "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." *Quarterly Journal of Economics* 134 (2): 557–98.