

**Food vs. Food Stamps:
Evidence from an At-Scale Experiment in Indonesia***

Abhijit Banerjee, MIT
Rema Hanna, Harvard Kennedy School
Benjamin A. Olken, MIT
Elan Satriawan, Gadjah Mada University and TNP2K
Sudarno Sumarto, SMERU and TNP2K

March 2021

Abstract

Governments seeking to provide food assistance have a choice between providing in-kind food directly to beneficiaries, or providing vouchers that can be used to purchase food on the market. To understand the differences between these policies, the Government of Indonesia randomly phased in the transition from in-kind delivery of subsidized rice to approximately equivalent vouchers usable to buy rice and eggs across 105 districts comprising over 3.4 million beneficiary households. We find the transition led to substantial changes in the allocation of aid in practice. The vouchers provided concentrated assistance to targeted households, who received 45 percent more assistance in voucher areas than in in-kind districts. As a result, for households in the bottom 15 percent at baseline, poverty fell by 20 percent. Vouchers also allowed households to purchase higher-quality rice, and led to increased consumption of egg-based proteins. We find vouchers have little effect on aggregate rice prices, although we observe modest price increases in remote villages. Overall leakage from the program was not affected, but the administrative costs of benefits delivery substantially fell. In short, the results suggest that the change from in-kind food aid to vouchers led to substantial impacts on poverty through the way it changed how programs were implemented on the ground.

* 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, Isadora Frankenthal, and especially Robbie Dulin for their excellent research assistance. We thank our many Indonesian government colleagues, particularly Bambang Widiyanto, 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 and Jesse Shapiro for helpful comments. This project was financially supported by the Australian Government and the Development Innovation Ventures at USAID. This RCT was registered in the American Economic Association Registry for randomized control trials under trial number AEARCTR-0004675. All views expressed in the paper are those of the authors, and do not necessarily reflect the views any of the many institutions or individuals acknowledged here.

I. INTRODUCTION

Targeted food programs, which aim to provide nutritional assistance to the poor, are one of the most common forms of social welfare programs in the world (World Bank, 2018).¹ There are two broad approaches to these programs. The first approach is an in-kind program, which distributes a set amount of free or subsidized foods to poor households. In developing countries, India's public distribution system and Egypt's Baladi Bread program fall into this category (Alderman, Gentilini and Yemtsov, 2018); similarly, the United States recently launched and distributed more than 100 million food boxes under the "Farmers to Families Food Box" program.² An alternative approach is a voucher program, which provides poor families with a voucher, or increasingly a pre-filled electronic debit card, which can be used to purchase particular food items at participating shops. One of the most famous of these programs is the United States' "Food Stamps" program, now known as the Supplemental Nutrition Assistance Program (SNAP), which has had substantial effects since it was first rolled out in 1961 (Hoynes et al., 2016, Bailey et al., 2020), but these programs are also common in developing countries, such as, for example, Egypt's Tamween Ration Cards or Sri Lanka's Samurdhi program.

At some level, one might expect these programs to have similar outcomes: the government can distribute a free gallon of milk to a poor household or distribute a voucher with which a household can pick up a free gallon of milk, and in the end, a household will get a gallon of milk. However, there are a number of theoretical reasons, grounded in basic price-theory, why economists might expect that these programs could lead to different outcomes. 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, or even the specific brands or quality mix within those items, to best suit their needs. Thus, depending on the designs of the two programs, one could observe different consumption choices, and

¹ Acknowledging the importance of food assistance worldwide, the Nobel Peace Prize of 2020 was awarded to the World Food Programme (WFP) "for its efforts to combat hunger, for its contribution to bettering conditions for peace in conflict-affected areas and for acting as a driving force in efforts to prevent the use of hunger as a weapon of war and conflict."

² See, for example, <https://www.ams.usda.gov/selling-food-to-usda/farmers-to-families-food-box>

ultimately nutritional outcomes (see for example, Leroy et al., 2013; Hidrobo et al., 2014; Cunha, 2014; Aker, 2017; Gentilini, 2016).

Second, 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. However, with inelastic supply, this could make a difference, since in-kind programs may feature a supply shock due to the government's role in supplying food. In particular, one may imagine that food supplies may be less elastic in remote and isolated areas with limited transport capacity, in which case in-kind transfers could reduce prices relative to voucher programs if the government is sourcing the in-kind food elsewhere and bringing it in (Coate et al., 1994; Cunha et al., 2019; Jiménez-Hernández and Seira, 2021).³ Third, in-kind programs could have important targeting properties through self-selection into the programs (Nichols and Zeckhauser, 1982; Currie and Gahvari, 2008): often times 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, ensuring that limited program resources get to the poor.⁴ The flexibility of voucher-based programs may mean that this type of self-targeting is less effective.

However, in addition to these commonly cited reasons, these programs may differ for another important, but less frequently discussed reason: the ability on the ground to administer these programs as designed could differ dramatically. In low state-capacity environments, these administrative differences may be first-order. First, the logistical challenges of running voucher and in-kind programs are different. On the one hand, it may be easier and cheaper for governments to refill electronic vouchers each month rather than run a massive logistics operation, moving millions of tons of food to various parts of the country. On the other hand, administering debit cards raises its 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

³ Gadenne et al (2021) also discuss how in-kind programs can effectively provide insurance against price variation by fixing the quantities provided, rather than fixing the value of subsidy provided. As we discuss below, these types of year-to-year price fluctuations are comparatively small in our setting.

⁴ There is also substantially larger literature in economics studying the difference between *cash*, which is completely flexible, and in-kind programs, as opposed to the question we focus on here of whether the in-kind assistance is actually directly delivered or delivered via vouchers. See Currie and Gahvari (2008) for a review; some more recent papers include Lieber and Lockwood (2019) on self-targeting within Medicaid in the United States; Hastings and Shapiro (2018) on mental accounting and stickiness of in-kind transfers studying SNAP in the United States; and Hidrobo et al. (2014) examining cash versus in-kind programs in Ecuador.

ensuring that debit cards reach the correct households (Banerjee et al., 2020; Muralidharan et al., 2020; Muralidharan et al., 2016). Second, governments using 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 in whichever way that they choose in practice.⁵ Finally, overall leakage may differ in voucher programs: it may be easier to monitor the money as it is directly transferred to the debit cards, whereas food can “fall off the truck” at various points in the distribution process. On the other hand, it may be easier to fraudulently claim digital payments than to abscond with tons of physical rice. In short, it is ultimately an empirical question as to which of these programs can be more effective in redistributing to the poor, and to ultimately affecting poverty rates.

To study the differences between in-kind food programs and electronic voucher programs in a setting with imperfect state capacity, we conducted a unique policy experiment in cooperation with the Indonesian government. Starting in 2017, Indonesia began a national reform to replace its largest anti-poverty program—“*Rastra*” (an abbreviation of *beRAS sejahTeRA*, or Rice for Welfare), an in-kind food program that delivered 10kg of free rice per month to 15 million targeted households nationwide—with a voucher 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 assigned to receive the program in 2019.

The enormous 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 voucher

⁵ There are, of course, many reasons why beneficiaries would not necessarily get their full voucher payment—they do not receive the debit card, they get held up by local officials or stores to process their voucher, or other non-poor community members or family members ask for a share given that the targeted household got a transfer—but the technological nature of the voucher may make it easier ensure that targeted households receive their entitlements.

programs through the lens of both the price theory and administration 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 use three primary datasets. First, we wrote a special module that was included in three waves of the national social-economic survey (SUSENAS), which we use, as well as its detailed consumption module. 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. Third, we merged in pre-data from the village census (PODES) to provide additional baseline data and allow us to measure heterogeneity by village characteristics.

We find that switching from in-kind transfers to vouchers led to a substantial change in the allocation of aid. Whereas in-kind food aid was often subdivided and distributed to many households in villages, 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. The share of households receiving a transfer fell in voucher areas, but disproportionately among better-off households: poorer households (at baseline) were 16 percent less likely to receive benefits in voucher areas, whereas wealthier households (at baseline) were 49 percent less likely to receive benefits in voucher areas. Combined, this resulted in a dramatic concentration of benefits among the poor: poorer households (at baseline) as a group received 45 percent more assistance in voucher areas than in in-kind districts. We show that this represents substantially more aid received by those who are actually poor, rather than being an artifact of who is on the baseline eligibility list. Interestingly, the better targeting in voucher areas occurred despite the fact that the food was rated as higher quality in the voucher program, and aid was more fungible, both reasons that the vouchers could have been more appealing to richer households. The administrative benefits of being able to better enforce the targeting of the electronic transfers thus swamped the self-targeting benefits of the poor quality of food in this context.

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 voucher program reduced poverty.

In addition to these first-order changes from increased concentration of assistance, we also examine whether some of the more traditional price-theoretic mechanisms associated with the switch to vouchers are present in this context. First, we observe some changes in types of food consumed with the switch to vouchers. Almost all households consume more rice than what was provided under the in-kind transfers—97 percent of households in in-kind areas consume more than the 10kg rice/month, and in fact, the mean household consumes about 32kg rice/month.⁶ 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. We do not observe changes in any other types of food, including other proteins such as beef and chicken, suggesting that this is unlikely to represent just an income effect. Instead, it appears that there is some type of stickiness from the types of food included in the voucher programs. This suggests that governments seeking to improve nutrition can potentially do so through 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.⁷

One important concern is that voucher-based programs could lead to increases in prices relative to in-kind programs, particularly in remote areas (Cunha et al., 2019). On average, we do not find any change in rice prices associated with the transition. This is true even in areas where the transition represented a larger shock (i.e., subsidized rice comprises a larger share of rice consumed in the district). We observe modest price increases in remote areas—for example, in villages in the 75th percentile or more in terms of distance we observe a 3.6 percent increase in rice prices—but even taking the price increase into account, poor households are still substantially

⁶ The in-kind program actually provided a mean of only 3.9kg of rice per month because it was spread among so many more beneficiaries. More than 99 percent of eligible households consumed more than 3.9kg of rice per month, so the in-kind program was likely binding for almost no one.

⁷ 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.

better off under the voucher program because they receive so much more assistance than in the in-kind program.

Finally, we examine the administrative costs of the programs. The administrative costs are low even for the in-kind program: the administrative costs are 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. The voucher program was thus substantially less costly to administer, despite the fact that it was more effective in delivering aid to targeted beneficiaries.

In short, switching from in-kind food delivery to vouchers made a real difference in the lives of the poor, reducing poverty among the poorest households by about 20 percent. This change was due to relative gains in program administration, as the vouchers resulted in a de facto allocation that was much more concentrated among the poor, consistent with the de jure program rules in both programs. The differences that occur due to differences in how these two programs were implemented in practice—i.e., the fact that the voucher-based program delivered concentrated assistance to targeted beneficiaries, rather than being distributed much more widely in local communities—swamped the differences in the programs that stemmed from more traditional price theoretic explanations, such as the ability for households to have more choice in spending or changes in aggregate prices.

The rest of the paper proceeds as follows. Section II describes the setting and research design. Section III presents our findings on the effect of voucher versus in-kind subsidies on the delivery of assistance, consumption choices, prices and leakages. Section IV concludes.

II. SETTING, EXPERIMENTAL DESIGN, AND DATA

A. Setting

In 1998, during the Asian Financial Crisis, the Indonesian government launched a new program to provide a safety net to poor households in the form of heavily subsidized rice, delivered directly to poor beneficiary households through village and local governments.⁸ With an annual budget of

⁸ 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 4th 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.

US\$1.5 billion dollars, the subsidized rice program, called “Rastra,” aims to provide subsidized food assistance to 15 million beneficiary households per month.⁹ While the program has changed slightly over time, in its current form, households are entitled to receive 10kg of free rice per month, with a value of about Rp. 100,000 (US\$8)/month, depending on the market price of rice (~Rp. 9,700 in our period). This is a substantial subsidy: it is about 6.5 percent of the poverty line for family of 4.

In theory, to be eligible for Rastra, targeting is conducted through proxy-means testing conducted by the central government, with the recipient list provided by the Ministry of Social Affairs. The proxy-means test was computed by conducting a census of more than 28 million households in 2015, which collected a variety of indicators that are predictive of the household’s consumption level (assets, education, etc.; see Alatas et al. (2012) and Banerjee et al. (2020) for more details).¹⁰ The National Team for the Acceleration of Poverty Reduction (TNP2K) combines these indicators into a proxy-means test 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 to the Ministry of Social Affairs if they believe they have been excluded from the proxy-means testing (Ministry of Social Affairs, 2018). According to the 2018 official rules, villages are not officially allowed to subdivide the rice—benefit recipients must receive the full 10kg.

Rice is procured by the government through the central logistics agency (BULOG), typically stored at district-level warehouses, and then delivered to villages (often in 50kg sacks). BULOG seeks to procure rice at approximately the market price; in our period, BULOG procures rice at Rp. 8,600/kg, which corresponds to the average sale price at rice mills for non-quality rated rice in 2017. While the program is large as an anti-poverty program, it represents only about 6

⁹ 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.

¹⁰ 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.

percent of the total rice market.¹¹ 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—removing Raskin and replacing 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 non-poor households receiving rice and many beneficiaries only receiving a fraction of their entitlement (Banerjee et al., 2018). The redistribution to non-poor households in the in-kind program had been a perennial challenge that the government was unable to fully resolve despite repeated attempts, and it was believed that the voucher program could reduce the improper inclusion of non-beneficiary households. Second, and related to the first point, Raskin 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 Raskin 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.¹²

The new voucher program was called Bantuan Pangan Non-Tunai (BPNT), or “Non-Cash Food Aid.” In principle, the program eligibility for BPNT was the same as Raskin. Instead of receiving rice, households received a monthly voucher of Rp. 110,000 onto 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 Raskin rice.¹³

¹¹ To calculate this, we divide the total amount of Raskin 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 Raskin allocations represent 6.2 percent of the total rice consumed in in-kind districts in 2018.

¹² 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.

¹³ The Raskin rice was valued at 10kg * the market price of rice, which fluctuates around Rp. 10,000. Since Raskin rice is often delivered in neighborhoods or door-to-door, whereas BPNT is 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

Households could use the voucher for purchases of rice or eggs at a network for eligible small shops, i.e., shops that were registered as remote agents for the state-owned bank chosen to implement the program in the district.¹⁴ The bank provided each agent with a debit-card reader machine (EDC machine), which connected to the bank’s network over a cellular connection to debit the amount 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.¹⁵ The government required banks to increase the number of agents, with a goal of 1 agent per 250 beneficiaries in every village, with a minimum of 2 shops in every village. While the agent network did increase, and almost all villages had at least 1 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 the government logistics agency.

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 processes. 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 (e.g., 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 potentially ready to be converted next. However, given the government’s budget—they planned to convert the program for about 8.3 million beneficiary

6.4 percent higher and never fell more than 15.4 percent lower than the price we observe at the time of this study. 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 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 (e.g., Londoño-Vélez and Querubin, forthcoming).

¹⁴ The debit cards were issued and administered by one of the four state-owned banks, 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.

¹⁵ 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.

households total, including the ones that were definitely chosen—they could not convert all 105 districts to BPNT in 2018.

Therefore, the government randomized which of the 105 districts would be treated in 2018, and which ones would subsequently be treated in 2019 (see Appendix Figure 1). These 105 districts—spread across Indonesia’s diverse landscape—comprise our sample (see 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 sample gives us a unique opportunity to study both the general equilibrium effects and administration outcomes of moving from an in-kind to a cash transfer program.

C. Randomization Process and Compliance

We randomized at the district level, as it 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 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 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. 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 stratum. For this last 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 the probability a given district is treated is random, the treatment probability in the 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).

D. Data

In order to evaluate the experimental switch from in-kind transfers to the vouchers program, we worked with the Government of Indonesia to design a special 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 Appendix Figure 3). The special module collected questions on whether households received each program, 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 the 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 * 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 data set 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 (NIK), which is collected in both datasets; we then analyzed the de-identified 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.¹⁶

E. Experimental Validity

We examine potential differences between the in-kind and voucher districts in 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 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.

III. Results

Section III.A outlines the estimation strategy. We then begin by exploring the impact of switching from the in-kind to voucher program on the delivery of assistance and poverty in Section III.B. Section III.C explores the impact on consumption choices, while Section III.D explores the impact on overall rice prices. Finally, Section III.E explores the impact on leakages and the administrative program costs.

A. Estimation

¹⁶ The 2018 PODES was enumerated from May 2 – 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 these as well as other control variables; results are qualitatively similar.

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

$$y_{hvds} = \beta_0 + \beta_1 Voucher_{ds} + \mathbf{X}'_{hvds} \boldsymbol{\gamma} + \alpha_s + \epsilon_{hvds} \quad (1)$$

where y_{hvds} is the relevant outcome variable; $Voucher_{ds}$ is the randomization into the voucher conversion in 2018; \mathbf{X}'_{hvds} are control variables selected using a double LASSO (Belloni et al., 2014); and α_s are strata fixed effects. In the LASSO, we include baseline control variables such as household covariates from the UDB, village level covariates from the village census and district*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 Appendix B. We estimate the intent to treat using the original randomization results (only 3 of our control districts were treated during the study period, so treatment-on-treatment models would be very similar). Standard errors are clustered by district, as that is the level of randomization; we report randomization inference p-values in all tables (Young, 2019).

Following our pre-analysis plan, we differentiate results based on households that are likely to be eligible for subsidized assistance, and households that are unlikely to be eligible. To determine this, we use the baseline proxy-means test (PMT) score each household received in 2015 in the Unified Targeting Data Base (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 pre-analysis plan we divide households into two groups: those with a PMT score of 30 or below, and those who had either a PMT score above 30 or those who were not included in the UDB.¹⁷ 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.

¹⁷ A full pre-analysis plan can be found at <https://www.socialscienceregistry.org/trials/4675>. We pre-specified the regression in (1), as well as pre-specified the separation by baseline PMT score ≤ 30 or not. The primary outcome variables specified in the pre-analysis 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 Appendix C.

While imperfect, the 2015 PMT score cutoff of 30 has strong predictive power both for consumption/poverty levels and for program eligibility. 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 ≤ 30 (i.e., the approximate target group), and those with PMT > 30 / those with no PMT score. The data reveal that households with PMT scores ≤ 30 are substantially poorer on virtually all dimensions. In particular, for households with PMT scores ≤ 30 , average per-capita consumption is Rp. 689,000/month; for those with PMT score ≥ 30 , average per-capita consumption is Rp. 1,116,000 / month – i.e., 62 percent higher. 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 5). 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 0-5 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.

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 ≤ 30 are substantially more likely to have been officially eligible for the Rastra program in 2017 (75 percent eligible for those with PMT ≤ 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 ≤ 30 , compared to only 33.4 percent for those with PMT > 30). Given these results, we interpret the split based on household pre-period PMT score as indicative of both official program eligibility and overall poverty levels.

B. Delivery of Assistance and Poverty

Who gets assistance, and how much?

We begin by showing the impacts on the receipt of subsidized assistance. 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 4 months prior to the survey.¹⁸ 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.¹⁹ For the voucher program, we take the average reported receipt in Rupiah; if the monetary value is missing in the survey but quantities are available, we take the amount of rice and eggs reported, multiplied by the market rate of the item in each rural or urban area of a given island/island group.²⁰ Note that households who received nothing from either program are included with a value of zero.

Figure 1, Panel A begins by showing the treatment effect on value of the total subsidy received non-parametrically. We break the households into bins based on household pre-period 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 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 ≤ 30 (i.e., targeted

¹⁸ Note that 1 USD averaged Rp. 14,000 during the period of this study.

¹⁹ By island groups, we refer to the main regions of Indonesia, i.e., 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.

²⁰ 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 III.C below, we do not, however, find prices for rice change on average. Nevertheless, as a robustness, we consider an alternative definition of the subsidy variable where we use the quantities of rice and eggs received, which we then value the same fixed set of prices for both programs. The results, presented in Appendix Table 3 Panel A, are broadly similar to the main results in Table 1.

beneficiaries), and PMT score > 30 (i.e., those not targeted).²¹ The key result is that the voucher program reallocated aid towards those with low PMT scores, i.e., the poorer, targeted group.²² Specifically, as shown in Column 2, targeted households received about 45 percent more in subsidy in voucher districts (Rp. 13,235 additional per month on average, compared to Rp. 29,200 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 ($p = 0.001$). Overall, the mean amount received across all households was slightly higher in voucher areas, by about 9 percent (Column 1, $p=0.087$).²³

This 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 1, 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 smallest 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 ≤ 30 ; for those with PMT scores > 30 , the probability of receiving assistance falls by 49 percent (14.4 percentage points; $p < 0.001$).²⁴ Thus, while everyone across the income distribution has a lower probability of

²¹ Table 1 examines outcomes in the March 2019 SUSENAS. The September 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 43 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 Appendix Tables 4 and 5.

²² Our results are remarkably robust to specification choices, such as dropping the controls selected via LASSO (Appendix Table 6), dropping the holdout sample (Appendix Table 7), or winsorizing the subsidy outcome (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 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).

²³ This likely represents 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/kg nationwide, so 10kg of in-kind rice would be valued at Rp. 97,000, compared to Rp. 110,000 for the voucher. We explore more systematically whether this result represents a reduction in leakage from the program in Section III.D below. We find that it most likely does not.

²⁴ 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 this data.

receiving the voucher program than the in-kind program, the reductions were much more pronounced for those that 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, recipients in voucher areas received 85 percent more in voucher areas than in in-kind areas on average over the previous 4 months (Rp. 31,263 more; $p < 0.001$). In particular, conditional on receiving anything, recipients in voucher areas are substantially more likely to receive the full amount they are entitled to. This pattern can be seen in Figure 2, which plots histograms of the average amount received in the 4 months prior to the survey in both in-kind and voucher districts. Panel A in Figure 2 shows all respondents with PMT score ≤ 30 ; Panel B shows only respondents with PMT score ≤ 30 who received some transfers. What is remarkable are the 4 regularly spaced spikes in Panel B for voucher locations. These correspond to households who received the full Rp. 110,000 transfer in the previous 4 months, 3 of the previous 4 months, and so on. By contrast, recipients in in-kind areas not only tend to receive substantially less on average, but they also do not show the clear bunching we see in the voucher areas. In fact, in February 2019, 97 percent of BPNT recipients who reported their transfer amount received the full Rp. 110,000, compared to only 18 percent of Rastra recipients who reported receiving the full 10kg. Similar patterns can be seen in Panel A, which shows all respondents.

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, and usually the full amount they were entitled to. 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.²⁵

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.

²⁵ 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. Appendix Table 10 shows no observable differences in either of these outcomes.

We 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 re-targeting that communities were doing in the in-kind program is towards poor households excluded by the PMT scores (Alatas et al., 2012). To examine this issue, Appendix Figure 5 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. Appendix Figure 5 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 20th and 30th percentiles, and it does so by reducing the amount of subsidy received by wealthier households, particularly those in the 65th – 90th 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, we look within the voucher areas, and compare households who did and did not receive the program. We find that 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 19 percent poorer than households who do not, even holding PMT scores constant. 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.²⁶

There are several important differences in the mechanisms through which the two programs were administered that may explain these differences. 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

²⁶ These results are shown in Appendix Table 11. Columns 1-3 present results with no controls, with district fixed effects, and controlling for dummies for each fine PMT score level (i.e., a dummy for each integer of PMT scores, i.e., 1, 2, 3, etc.). Columns 4-6 present the same results, but with raw per-capita consumption as measured at endline (i.e., not making any adjustment for program receipt). Panel A considers the entire sample and Panel B restricts to households with PMT scores < 30. To investigate whether local officials target on other dimensions (as in Galasso and Ravallion (2005), Alatas et al. (2012) and Basurto et al. (2020)), we further explore differential targeting of the two programs on other household characteristics in Appendix Table 12. We find that the voucher program distributes more benefits than the in-kind program to poor 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.

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 pre-printed on the card (see 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 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 Appendix Table 13, 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.²⁷ 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 regulates 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.²⁸

Quality and targeting

In addition to how *much* assistance people received, there is also a question of the *quality* assistance people received. The poor quality of rice under the in-kind program was a frequent

²⁷ Unfortunately, the SUSENAS did not ask recipients of the 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.

²⁸ Data as of 2012; see <http://www.g-a-i.org/wp-content/uploads/2012/10/GAI-Report-ProfitsfromPoverty-FINAL.pdf>

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 rice received, on a 3-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.²⁹ This suggests that the quality-adjusted amount of subsidy received by targeted beneficiaries in voucher areas is even higher than the 45 percent increase discussed above. Putting a monetary value on this quality improvement is challenging, but a back-of-the-envelope calculation suggests that the quality-adjusted values of subsidy received by targeted households in voucher areas may have further increased the amount of the transfer received by about 20 percent.³⁰

The quality improvements in the voucher program are likely related to the fact that 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 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

²⁹ 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 vs. low quality rice over the period we studied here (e.g., it averaged Rp. 646/kg (7.4%) in 2016, before the transition began, and Rp. 518 (5.6%) in 2019).

³⁰ 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/kg, and the quality measure in Table 1 represents 0.25 standard deviations, we would assign a price equivalent of Rp. 250/kg). On average, this calculation suggests that the quality improvements in Table 1 would be worth an additional Rp. 1,060 per kg, which is anecdotally consistent with typical differences in market prices between low and average quality rice. 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 20 percent higher.

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.

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 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 in Figure 3 by presenting the results non-parametrically, 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 3 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 begin

with, and also because – as shown above – the impact of the transfer was the largest for these groups.³¹

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 poverty (6.5 percentage points, $p=0.012$).³² Appendix Table 18 shows similar patterns in the voucher program for both the poverty gap and the squared poverty gap.

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

³¹ Appendix Tables 14 and 15 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 Appendix Table 14, households with PMT scores 0-5 saw their mean subsidy increase by Rp. 19,401 in voucher areas compared to control ($p<0.001$), compared to Rp. 13,325 for the entire group with PMT ≤ 30 .

³² Since poverty is measured by per-capita consumption being below the relevant poverty line, these results are, mechanically, coming from increases in consumption for households near the poverty line. We also examine changes in *mean* consumption levels for each of these groups in Appendix Table 16, both in logs (Panel A) and levels (Panel B). While we do not see any changes in mean consumption level for either the entire PMT ≤ 30 or PMT > 30 samples, we do see increases in consumption when we zoom in on the poorest households. Households with PMT ≤ 10 have 3.8 percent higher mean per-capita consumption ($p=0.104$); households with PMT < 5 have 5 percent higher mean per-capita consumption ($p=0.049$). We find qualitatively similar results examining changes in CRRA consumption (assuming a coefficient of relative risk-aversion of 3), see Appendix Table 16 Panel C. We find no observable impact on the food insecurity index (Appendix Table 17).

x 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 Appendix Table 19, 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.

B. 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?

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. That is, under a standard consumption model, if a household was consuming more than 10kg of rice per month, it was doing so not because of the in-kind transfer of 10kg rice per month per se, but rather because its optimal consumption of rice was above 10kg. As such, one would expect households to consume the same consumption bundle regardless of whether they received the transfers of rice in-kind or the more flexible vouchers.

Figure 4, Panel A plots out rice consumption in the in-kind areas, and it shows that almost everyone is consuming more than 10kg of rice. Examining households with PMT scores below 30, we find similar patterns; in fact, only 3.3 percent of these households consume less than 10kg of rice. Moreover, in practice the in-kind program did not distribute 10kg 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 36kg of rice on average (Appendix Table 20). 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* consumption of rice and eggs (Table 3, Panel B). Note that the total consumption measures come from a completely separate module in the SUSENAS which ask about consumption of the food product in question (as well as over 200 other consumption items) over the previous week from all sources. Meanwhile, information on the subsidized consumption come from questions about the Rastra and BPNT programs specifically. 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).

Looking first at the bundle of subsidized food households receive, we see no change in consumption of subsidized rice, either all households (Panel A, Column 1) or those with PMT 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 ≤ 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, which we explore in Panel B. We find no impact on total rice consumption for either of these samples.³³ However, we do observe that eligible voucher households consume more total egg protein. Specifically, for households with PMT scores ≤ 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

³³ Appendix Table 20 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).

percent of the increase in subsidized egg protein consumed represents a net increase in consumption.

Appendix Table 21 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 labelling 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 (Appendix Table 22). 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.

C. Prices

One benefit of this unique policy experiment is that given that the unit of randomization was at a high geographic 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.

In Column 1, we first begin by looking at the overall impact of the voucher program on the price of rice. 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 0.296).

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 voucher represents a larger negative supply shock in these locations (Filmer et al., 2018). We measure the size of the supply shock by dividing the 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 75th percentile and higher in this supply shock measure; we observe a larger effect in magnitude than for those above median (Rp. 555 rather than Rp. 180), but it remains statistically insignificant (p-value of 0.145) and still relatively small in magnitude (representing a 5.8 percent price increase).³⁴

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 et al., 2019). We examine several measures of remoteness, which we can measure at the village level. We first examine whether your 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 0.579 and 0.355 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. The graphs show no impact on prices for those locations closest to the district capital, and 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 90th percentile in distance.

³⁴ We also examine these heterogeneity measures using alternative specifications, either examining continuous measures (Appendix Table 23). We observe similar conclusions.

Turning back to Table 4, we examine the price impacts on villages that are above the median (Column 6) and above the 75th 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 0.226). For those that are very remote (i.e., 75th percentile or higher), we observe a Rp. 338 (3.6 percent) increase in voucher areas compared to in-kind areas (p-value of 0.029), with no price change in the remaining 75 percent of locations. However, even the 3.6 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 19.5kg of rice from the private market per month from all sources, which implies a Rp. 6,591 increase in overall rice spending in these remote areas, compared to the 13,235 increase in benefits received by likely eligible households.³⁵

D. 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 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.

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 Column 2. We do this in case the prices paid at local agents are inflated. Finally, in Column 3,

³⁵ 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 Appendix Table 3, Panel A.

we adjust the market price of rice in BPNT to reflect the higher quality of rice bought to allow for quality weighted comparisons.³⁶

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 10kg 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 this measure 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 this leakage metric *changes* 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 5 percent level. Finally, adjusting for quality of rice received, we also observed negative, but insignificant impacts of the conversion on leakages.³⁷

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 Appendix Table 25. 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 EDC machines for the agents. Assuming that the agents only process

³⁶ 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.

³⁷ 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 Appendix Table 24, 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.

vouchers—and did not use the machines for other financial transactions—Appendix Table 25 shows that the total administrative costs would be 2.16 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 18 percent of administrative costs 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 a similar electronic voucher-based program. We find that, while the programs are broadly similar, they led to dramatically different differences in outcomes through the way the programs were administered in practice. Specifically, households with low baseline predicted poverty scores—i.e., those households who were the intended beneficiaries of both programs—received 45 percent more assistance in voucher areas compared to in-kind areas. By contrast, benefits in the in-kind program were diffused much more broadly throughout the population.

The voucher program’s substantially higher concentration of benefits among the poor led to a reduction in poverty. For those households in the bottom 15 percent at baseline, the poverty rate fell by 20 percent. This occurred despite the fact that the administrative costs of operating the voucher-based program were less than half that of the in-kind program.

We also find evidence that the flexibility of the voucher program led to an increase in consumption of egg-based proteins. This occurred despite the fact that households in the in-kind program were unconstrained; that is, since virtually all households consume more rice anyway than they receive in the program, basic price-theory suggests that these households always had the option of buying less rice at the margin and spending more on eggs. The fact that the vouchers allowed only rice and eggs, and it was specifically egg protein consumption that increased, suggests that there is ‘stickiness’ in the consumption bundles allowed in the voucher program. This

type of stickiness is reminiscent of the findings of Hastings and Shapiro (2018), who find the related result that the overall marginal propensity to spend SNAP benefits on food is higher than the marginal propensity to spend cash transfers on food. We find no changes in aggregate rice prices, although we find modest increases in villages that are very remote. Finally, we find no changes in in leakage.

On net, the results here suggest an important additional dimension for how to think about vouchers vs. 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 imperfect administrative capacity.

References:

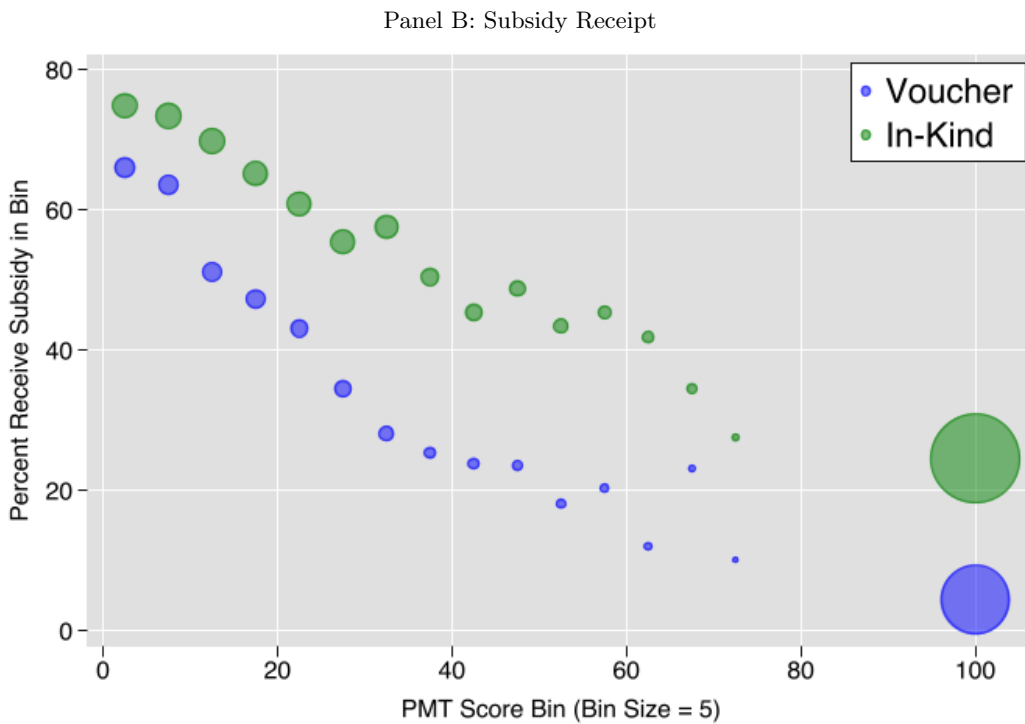
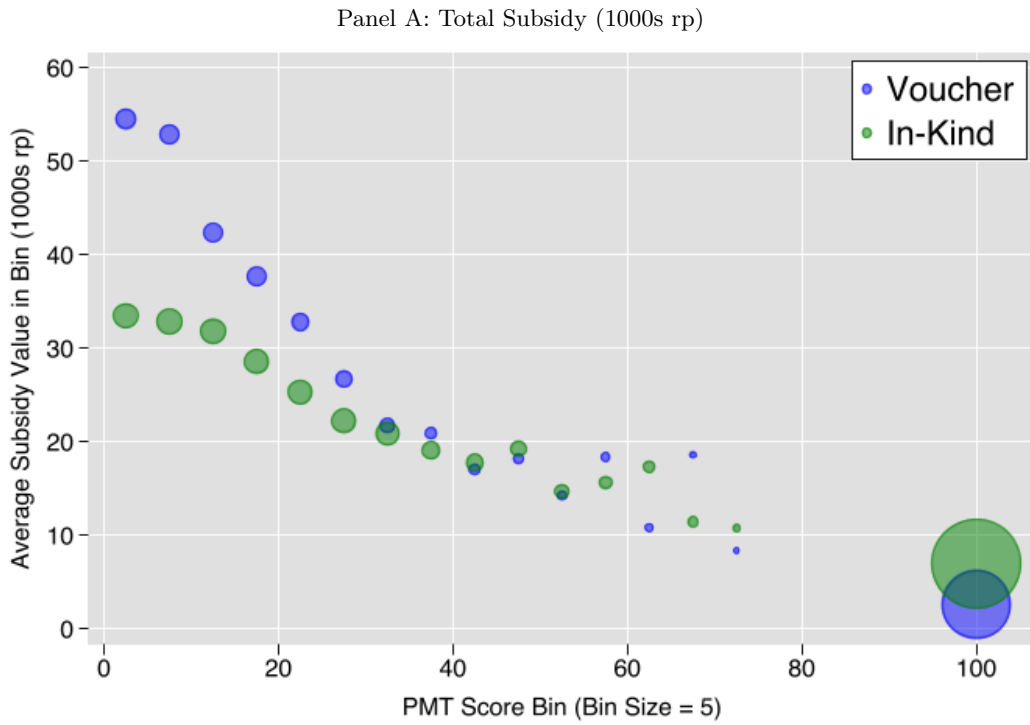
- Aker, Jenny C. 2017. "Comparing Cash and Voucher Transfers in a Humanitarian Context: Evidence from the Democratic Republic of Congo." *The World Bank Economic Review*, 31 (1): 44-70.
- Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, and Julia Tobias. 2012. "Targeting the Poor: Evidence from a Field Experiment in Indonesia." *American Economic Review*, 102 (4): 1206–1240.
- Alderman, Harold, Ugo Gentilini, Ruslan Yemtsov. 2018. *The 1.5 Billion People Question: Food, Vouchers, or Cash Transfers?* Washington, DC: World Bank.
- Bailey, Martha J., Hilary W. Hoynes, Maya Rossin-Slater, and Reed Walker. 2020. "Is the social safety net a long-term investment? Large-scale evidence from the food stamps program." NBER Working Paper #26942.
- 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, 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, 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: 451-491.
- 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): 101-137.
- 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." Working paper.
- 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, Christian Hansen. 2014 "Inference on Treatment Effects after Selection among High-Dimensional Controls." *Review of Economic Studies* 81 (2): 608–650.
- 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, Seema Jayachandran. 2019. "The Price Effects of Cash Versus In-Kind Transfers." *The Review of Economic Studies*, 86 (1): 240–281.

- 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): 705-727.
- 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: 144-156.
- Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-run impacts of childhood access to the safety net." *American Economic Review* 106, no. 4: 903-34.
- Jacoby, Hanan G. 1997. "Self-Selection and the Redistributive Impact of In-Kind Transfers: An Econometric Analysis." *The Journal of Human Resources*, 32 (2): 233-249.
- Jiménez-Hernández, Diego and Enrique Seira. 2021. "Should the Government Sell You Goods? Evidence from the Milk Market in Mexico." Working Paper.
- 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." *The Journal of Nutrition*, 143 (3): 378–383.
- Lieber, Ethan M. J., and Lee M. Lockwood. 2019. "Targeting with In-Kind Transfers: Evidence from Medicaid Home Care." *American Economic Review*, 109 (4): 1461-85.
- Londoño-Vélez, Juliana and Pablo Querubin, "The Impact of Emergency Cash Assistance in a Pandemic: Experimental Evidence from Colombia," *Review of Economics and Statistics*, forthcoming.
- 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, Sandip Sukhtankar, and Jeff Weaver. 2020. "Improving Last-Mile Service Delivery Using Phone-Based Monitoring." *American Economics Journal: Applied Economics*, forthcoming.
- Nichols, Albert L., and Richard J. Zeckhauser. 1982. "Targeting Transfers through Restrictions on Recipients." *The 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): 853-870.
- Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy*, 115 (2): 200-249.
- World Bank. 2018. The State of Social Safety Nets 2018. Washington, DC: World Bank. World Bank. <https://openknowledge.worldbank.org/handle/10986/29115> License: CC BY 3.0 IGO.
- Young, Alwyn. 2019. "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." *The Quarterly Journal of Economics*, 134 (2): 557–598.

Figures

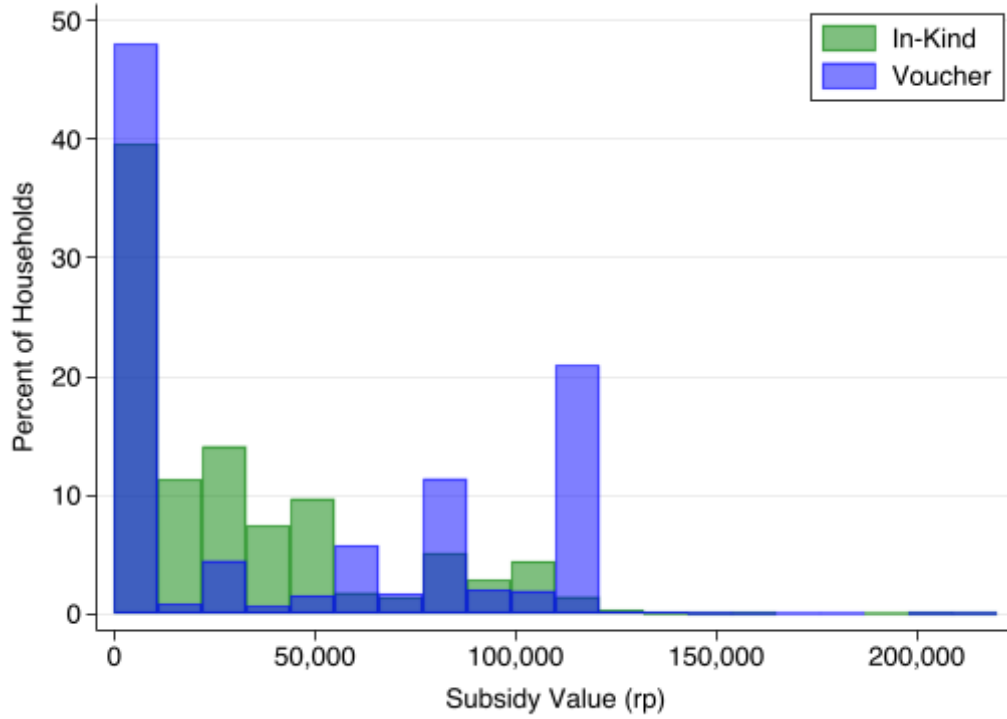
Figure 1: Relationship between PMT Score and Subsidy Received



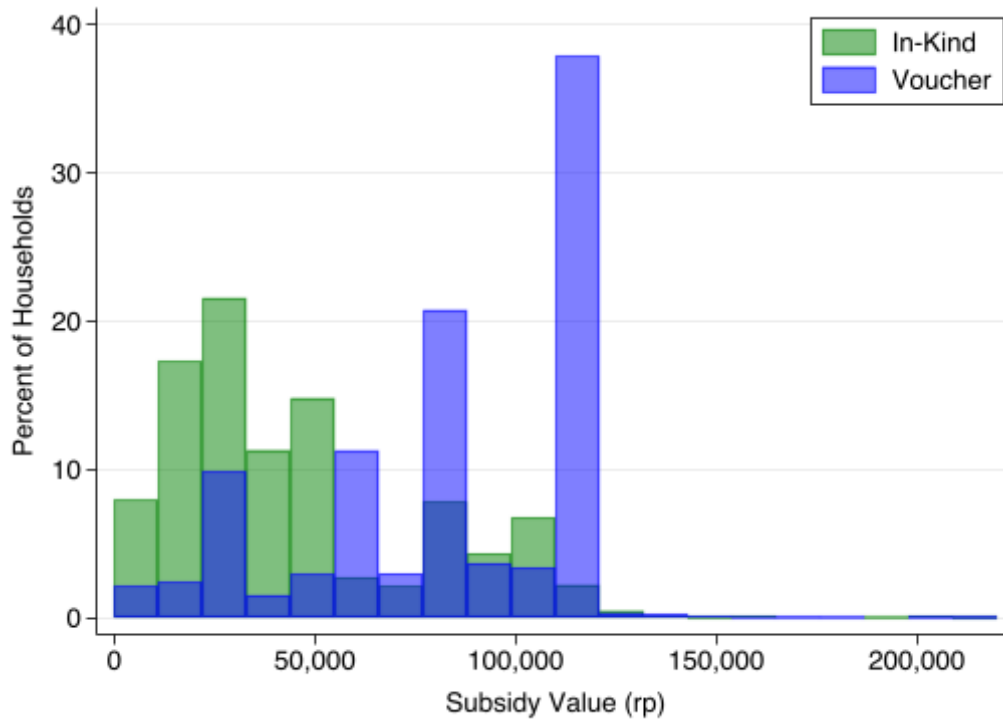
Note: These figures graph the relationship between a household's PMT score and, respectively, their total subsidy received (Panel A) and an indicator variable for received subsidy (Panel B), by treatment. PMT scores are binned in groups of 5, with those who have no PMT score grouped with those with a 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 Unified Targeting Data Base.

Figure 2: Concentration of BPNT Subsidy

Panel A: Households with PMT score ≤ 30

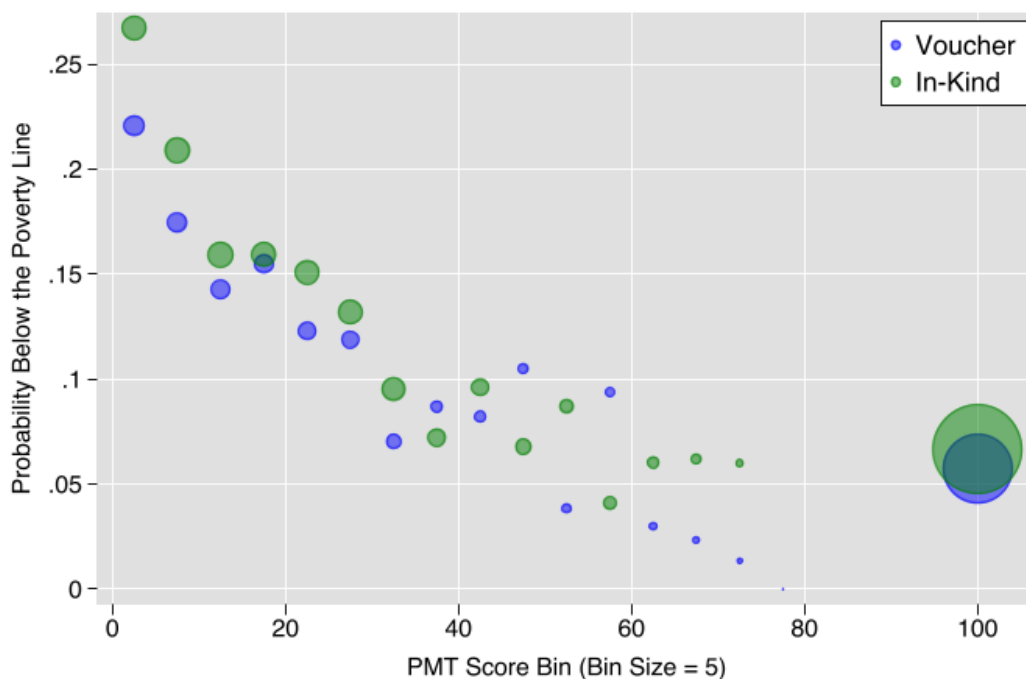


Panel B: Recipient households with PMT score ≤ 30



Note: These figures graph the distribution of subsidy received, by treatment and PMT score cutoff. The data are from March 2019 SUSENAS. The sample size in Panel A is 16,316, while it is 9,850 in Panel B.

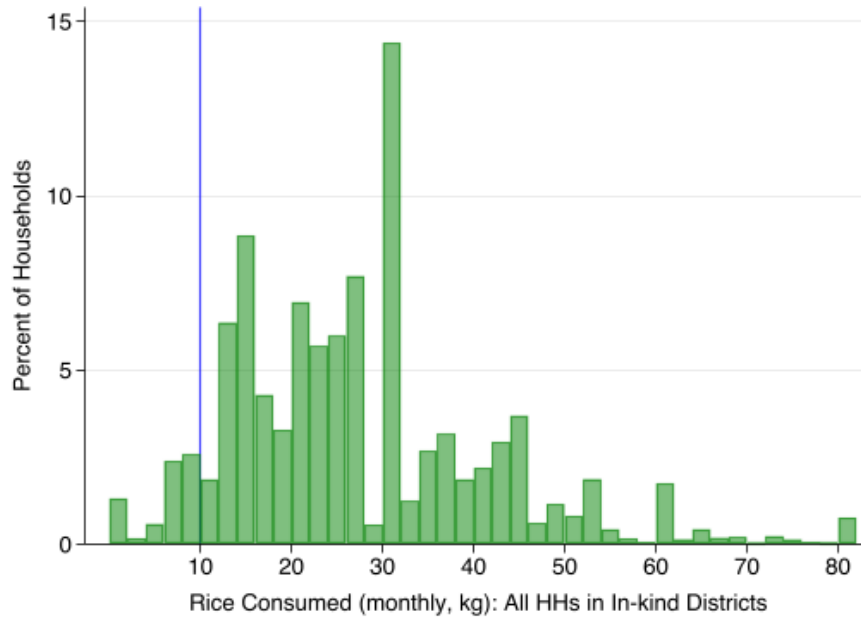
Figure 3: Relationship Between PMT Score and Poverty Status



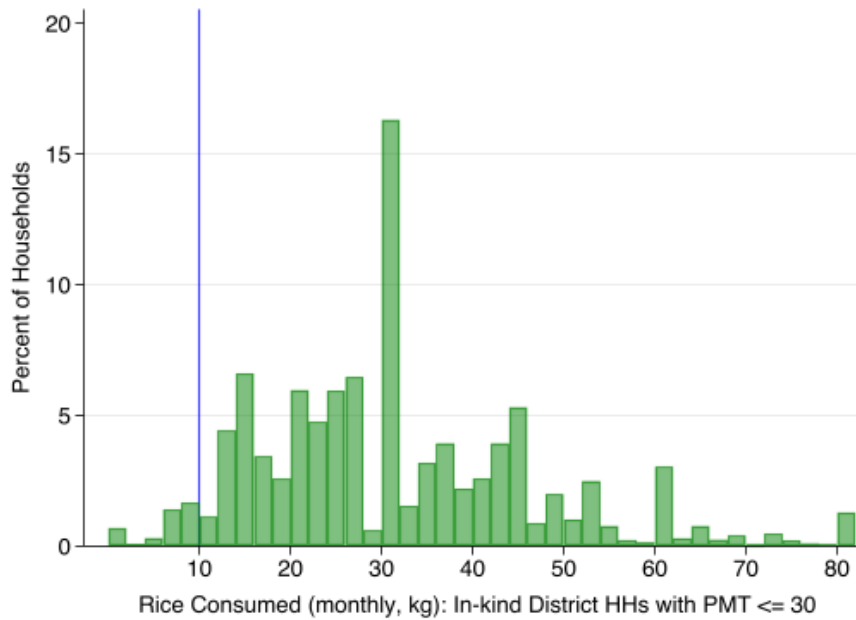
Note: This graph provides the relationship between whether the household is below the poverty line and PMT score, by treatment status. PMT scores are binned in groups of 5, with those who have no PMT score grouped with those with a score of 100. Markers are scaled by the number of households in each bin. Data on outcomes are from the March 2019 SUSENAS, while PMT data are from the Unified Targeting Data Base.

Figure 4: Total Monthly Rice Consumption (kg)

Panel A: All households in in-kind areas

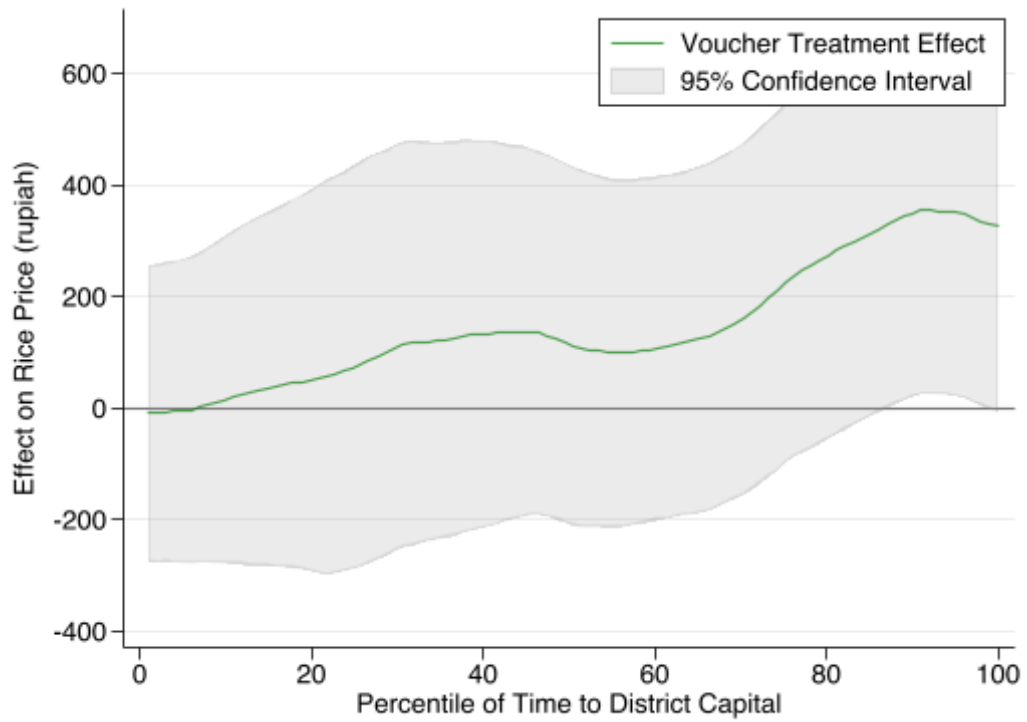


Panel B: Households in in-kind areas with PMT score ≤ 30



Note: This figure graphs the distribution of total rice consumed in kilograms in the in-kind districts for all households (Panel A) and those who have a PMT score that is less than or equal to 30 (Panel B). The data are from the March 2019 SUSENAS. The sample size in Panel A is 39,612, while it is 9,763 in Panel B.

Figure 5: Non-Parametric 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 Unified Targeting Data Base. Regressions are estimated using a triangular kernel and a bandwidth of 4. Data are from the March 2019 SUSENAS.

Tables

Table 1: Experimental Difference between Voucher and In-kind Districts on Subsidy Outcomes

	Total Subsidy (rp)						Recipients Only	
	Total Subsidy (rp)			Receive Subsidy			Total Subsidy (rp)	Rice Quality
	All (1)	PMT ≤ 30 (2)	PMT > 30 (3)	All (4)	PMT ≤ 30 (5)	PMT > 30 (6)	All (7)	All (8)
Voucher	1304.749 (617.738) [0.087]	13234.952 (1915.934) [0.000]	-2571.436 (563.267) [0.001]	-0.134 (0.019) [0.000]	-0.105 (0.021) [0.000]	-0.144 (0.020) [0.000]	31262.669 (3346.486) [0.000]	0.203 (0.020) [0.000]
Observations	66494	16327	49566	66496	16329	49566	19355	19260
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Double Lasso	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DV Mean (Control)	14456.314	29200.535	9161.727	0.393	0.669	0.293	36918.120	0.630

Note: 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 Unified Targeting Data Base. We used a double LASSO to choose the control variables (all potential variables used as inputs for the LASSO are listed in 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.

Table 2: Experimental Difference between Voucher and In-kind Districts Being Below the Poverty Line

	All (1)	PMT \leq 30 (2)	PMT \leq 25 (3)	PMT \leq 20 (4)	PMT \leq 15 (5)	PMT \leq 10 (6)	PMT \leq 5 (7)
Voucher	-0.009 (0.008) [0.205]	-0.023 (0.015) [0.134]	-0.025 (0.016) [0.166]	-0.034 (0.017) [0.078]	-0.043 (0.018) [0.028]	-0.052 (0.020) [0.020]	-0.065 (0.024) [0.012]
Observations	66496	16329	13707	11072	8307	5529	2788
Stratum FE	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

Note: 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 Unified Targeting Data Base. We used a double LASSO to choose the control variables (all potential variables used as inputs for the LASSO are listed in 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.

Table 3: Experimental Difference between Voucher and In-kind Districts on Food Consumption

<i>Panel A: Subsidized Food Consumption</i>						
	Subsidized Rice (kg)			Subsidized Egg Protein (g)		
	All (1)	PMT \leq 30 (2)	PMT $>$ 30 (3)	All (4)	PMT \leq 30 (5)	PMT $>$ 30 (6)
Voucher	-0.300 (0.066) [0.002]	0.062 (0.205) [0.773]	-0.424 (0.058) [0.000]	10.932 (1.534) [0.000]	32.719 (4.648) [0.000]	3.362 (0.463) [0.000]
Observations	66495	16328	49566	66423	16270	49552
Stratum FE	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

<i>Panel B: Total Food Consumption</i>						
	Total Rice (kg)			Total Egg Protein (g)		
	All (1)	PMT \leq 30 (2)	PMT $>$ 30 (3)	All (4)	PMT \leq 30 (5)	PMT $>$ 30 (6)
Voucher	-0.012 (0.314) [0.971]	-0.411 (0.478) [0.492]	0.143 (0.304) [0.704]	3.454 (3.110) [0.354]	9.279 (4.750) [0.100]	0.566 (3.781) [0.891]
Observations	66496	16329	49566	66483	16327	49555
Stratum FE	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

Note: 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 Unified Targeting Data Base. We used a double LASSO to choose the control variables (all potential variables used as inputs for the LASSO are listed in 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.

Table 4: Experimental Difference between Voucher and In-kind Districts on Price

	Main Effect Only (1)	Above Med. Supply Shock (2)	Above 75th Pct. Supply Shock (3)	Measures of Isolation			
				Non-asphalt Road (4)	Road Not Always Passable (5)	Above Med. Time to District Capital (6)	Above 75th Pct. Time to District Capital (7)
Voucher	140.536 (135.230) [0.296]	77.643 (180.112) [0.673]	69.657 (138.748) [0.623]	127.417 (145.451) [0.363]	135.296 (136.243) [0.322]	58.970 (150.040) [0.689]	56.963 (142.097) [0.655]
Voucher × [Variable]		180.141 (267.890) [0.541]	554.537 (487.630) [0.145]	71.061 (128.233) [0.579]	181.389 (179.736) [0.355]	155.648 (119.427) [0.226]	338.125 (140.447) [0.029]
Observations	32343	32343	32343	32334	32334	32334	32334
Stratum FE	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)	9478.508	9478.508	9478.508	9478.508	9478.508	9478.508	9478.508
[Variable] Mean		0.540	0.238	0.137	0.035	0.489	0.236

Note: 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 Unified Targeting Data Base. 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 75th percentile subsidized rice as a fraction of total rice consumption in the district, respectively. Non-asphalt 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 75th 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.

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	-0.020 (0.031) [0.533]	-0.059 (0.029) [0.055]	-0.013 (0.031) [0.708]
Observations	105	105	105
Stratum FE	Yes	Yes	Yes
Double Lasso	Yes	Yes	Yes
DV Mean (Control)	0.587	0.586	0.588

Note: 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 multiple the number of beneficiaries in the district by the 10kg rice disbursement and the official procurement price of Rastra rice; in the voucher districts, we multiple 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 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.