DEMS
Department of Economics, Management and Statistics

DEMS WORKING PAPER SERIES

When Humanitarian Assistance Fades: Experimental Evidence from a Nationwide Program in Lebanon

Ghassan Baliki, Micah Cruz, Felipe Dunsch
Jonathan Garcia, Andrea Guariso
Jonas Heirman, Dorothee Weiffen

No. 584 – November 2026

Department of Economics, Management and Statistics
University of Milano – Bicocca
Piazza Ateneo Nuovo 1 – 2016 Milan, Italy
http://dems.unimib.it/

When Humanitarian Assistance Fades: Experimental Evidence from a Nationwide Program in Lebanon

Ghassan Baliki[1] Micah Cruz[2] Felipe Dunsch[2]
Jonathan Garcia[1,3,4] Andrea Guariso[1,2,5]
Jonas Heirman[2] Dorothee Weiffen[1]

September 2026

Abstract:

What happens when humanitarian assistance ends? We study a 57 percent contraction of Lebanon's nationwide food assistance program. Random assignment of recipient households to four prioritization rules generates variation in assistance continuation, while a second randomization varies the length of the transition out of assistance. Households losing assistance replace only one-fifth of the lost transfer through higher food spending and experience persistent declines in food security, psychological well-being, and trust in aid agencies. A longer transition reduces trust losses but does not cushion other welfare losses. No prioritization rule clearly dominates, but households eligible under more rules are less able to absorb withdrawal.

* We thank Vojtěch Bartoš, Marianna Battaglia, Tilman Brück, Johannes Haushofer, Macartan Humphreys, Eliana La Ferrara, Rocco Macchiavello, Karen Macours, Owen Ozier, Nancy Qian, Jakob Svensson, David Yanagizawa-Drott and seminar participants at the University of Alicante, KU Leuven, LEAP, University of Milan, University of Milan-Bicocca, and WFP headquarters for helpful comments and suggestions. We also thank Isabella Monteiro for invaluable research assistance. Finally, we thank Marco Principi and the WFP Lebanon Country Office for insightful discussions about the program and for their collaboration throughout the study, as well as for sharing administrative data. All mistakes are our own. Financial support from USAID Bureau for Humanitarian Assistance (#720BHA23IO00036) is greatly appreciated. The study was registered with the AEA RCT registry (11831) and received ethical approval from the ethics committees of Humboldt University (2023-08-W) and from HML International Review Board (2766).
[1]International Security and Development Center (ISDC), Berlin, Germany
[2]World Food Programme, Rome, Italy
[3]Humboldt University, Berlin, Germany
[4]Leibniz Institute for Horticultural Sciences (IGZ)
[5]University of Milan-Bicocca, Milan, Italy

1. Introduction

Humanitarian agencies routinely downscale or terminate programs in response to funding shortfalls and shifting priorities. The pressure to contract programs has become increasingly acute in recent years, as humanitarian needs have risen faster than available resources, resulting in large and widening funding gaps: in 2024, the United Nations raised only 49 percent of the $49.6 billion required to meet global humanitarian needs (UNOCHA, 2026).[1] Funding is also volatile, with resources abruptly reallocated in response to emerging crises or political shifts, as illustrated by recent large-scale cuts by major donor governments.[2] Agencies thus face a recurring allocation problem, as they must decide which households should continue to receive assistance and how support should be withdrawn from the others. Yet, while a large literature experimentally studies how assistance programs should be introduced and designed, causal evidence on program exit remains scarce (Deserranno et al., 2026).

The withdrawal of humanitarian assistance raises three key questions. First, what are the consequences of losing assistance? Humanitarian programs are typically introduced in response to acute shocks and intended to be temporary, with exit ideally occurring as the emergency subsides and the marginal value of continued support declines. In practice, however, programs often contract because of budget constraints, changing donor priorities, or competing crises rather than improvements in recipients' circumstances (Fink and Redaelli, 2011). When temporary transfers are repeatedly renewed, households may also come to expect their continuation and incorporate them into consumption plans. Withdrawal can therefore reduce households' available resources while also generating a loss relative to expected consumption, potentially magnifying its welfare cost (Kőszegi and Rabin, 2006, 2009). Transfers also shape attitudes towards the institutions that provide them (Manacorda et al., 2011), so withdrawal may impose institutional costs beyond its effects on material welfare. Second, how do households prepare for the end of assistance? Households may try to smooth the transition by adjusting consumption, saving, or labor supply. But liquidity and credit constraints may impede such adjustment, while present bias, myopia, or limited attention may delay responses until benefits are withdrawn (Deaton, 1991; O'Donoghue and Rabin, 1999; Reis, 2006; Ganong and Noel, 2019). Third, how should programs phase

[1] Figure A.1 in the Appendix combines annual estimates from the Global Humanitarian Overview with data from UN OCHA's Financial Tracking Service to document the evolution of the number of people in need and the share of humanitarian funding requirements left unmet between 2016 and 2025.

[2] See, for example, "Diminished UK aid budget is 'new normal', says development minister" (The Guardian, 17 May 2025) and "Trump administration says it's cutting 90% of USAID foreign aid contracts" (AP, 27 February 2025)

out assistance to limit welfare losses while making effective use of their scarce resources? Different exit strategies may ease the transition, for instance by giving households more time to adjust or by providing additional support during the phase-out. A central decision is which households should retain assistance. Ideally, this choice would reflect households' needs and vulnerability, but both are latent and multidimensional. Agencies may therefore rely on alternative rules that operationalize different notions of need and, as a result, select different beneficiaries (Ligon and Schechter, 2003; Brown et al, 2018).

We study these questions in the context of the 2024 scale-down of Lebanon's nationwide food assistance program. The program, which is currently still running, is managed by the World Food Programme (WFP) in partnership with the Lebanese government. Launched in 2020 in response to Lebanon's severe socio-economic crisis, the program expanded to 58,501 Lebanese households by 2023, covering more than 5% of the population. Beneficiary households receive monthly food parcels valued at $32 each (January 2023 market prices). The transfer corresponds to roughly 42% of the median household's monthly food expenditure and 18% of its total expenditure. Funding constraints and shifting priorities forced WFP to reduce the caseload to about 25,000 households, or 43% of its previous level, by January 2024.

In collaboration with WFP's Lebanon Country Office, we designed a randomized controlled trial across the entire beneficiary population. First, we randomly divided the population into four equal-sized groups. Within each group, we ranked households using one of four alternative vulnerability rules, and selected the 43 percent ranked as most vulnerable to continue receiving assistance. The rules represent established targeting practices and reflect distinct approaches (data- vs. human- driven) and notions of need (low economic resources vs. food deprivation vs. multidimensional vulnerability). Importantly, the rules select different households: when simulated over the full beneficiary population, they disagreed on eligibility for 61 percent of households. Among households with the same eligibility pattern across the four rules, random assignment to one group rather than another therefore generates exogenous variation in assistance continuation.

Second, we randomly assigned a subset of households excluded from the assistance program to a more gradual exit. Under WFP's standard procedure, excluded households were notified in December 2023 that assistance would end in that same month. About 1,000 randomly selected households instead received a six-month transition period: they were informed in December that assistance would continue through June 2024 and then end. This "transition package" combines advance notice with six additional months of support. We timed our first follow-up survey round to occur shortly before this group exited the program, allowing us to

observe anticipatory responses while households were still receiving assistance. The second survey round followed approximately seven months after program exit.[3]

This two-step design allows us to address the three questions outlined above. First, we estimate the consequences of assistance withdrawal for households whose eligibility for continued support depends on the prioritization rule to which they are randomly assigned. Second, we study whether and how households adjust in anticipation of the end of assistance, by comparing households that received the "transition package" while they were still receiving assistance, with households whose assistance continued. Third, we evaluate two design margins that may mitigate the adverse effects of withdrawal: the transition package, by comparing excluded households with and without the package after program exit, and the choice of prioritization rule.

Our analysis combines administrative records covering the full beneficiary population (58,501 households) with two survey rounds from a stratified random sample of approximately 6,400 households, conducted in May–June 2024 and January–February 2025. Our main outcomes capture different dimensions of households' well-being, and include measures of food expenditure, food security, mental health, and trust.

Balance and attrition checks support the integrity of the original design, although an implementation error reduced the transition package (TP) sample at midline and therefore the precision of our estimates of anticipatory adjustment.[4] Between the two survey rounds, however, the escalation of the conflict with Israel led WFP to introduce temporary emergency cash assistance in affected areas, for households not already receiving WFP support. This programmatic response illustrates a broader challenge of conducting research in humanitarian settings, where new emergencies can rapidly reshape program priorities. It also weakened the endline treatment contrast by extending WFP support to many households whose food assistance had been discontinued, likely attenuating the estimated effects of program withdrawal. We therefore complement our main analysis with a specification that allows the experimental effect of assistance discontinuation to vary by priority-area status.

[3]An alternative design to isolate the impact of a longer notice window, would have held the date of assistance withdrawal fixed while randomly varying the timing of the advance notice. This was not feasible for two reasons. First, it would have required withholding information from some households after their discontinuation status was known, which raised ethical concerns: in our design, all households were informed as soon as a decision had been made. Second, given the tight program timeline, it would have required postponing exit for households otherwise scheduled to leave earlier, financing additional months of assistance for all households assigned to discontinuation, which was not feasible due to budget limitations.

[4]As discussed in section 6, due to an error in constructing the survey target list, only 31 percent of the intended TP sample was surveyed at midline. Importantly, this subsample remained randomly selected and well balanced across treatment arms. The error was corrected at endline, when 93 percent of the TP sample was surveyed.

Our results can be summarised in three main points.

First, households losing assistance partially offset the transfer through higher food expenditure, partly financed by purchases on credit, but still experience significant welfare losses. Food expenditure increases by about USD 12.5 at midline and USD 10 at endline, with increases concentrated in food categories previously provided through the assistance basket. These increases replace roughly 19 percent of the market value of the lost transfer and are insufficient to prevent a deterioration in food security. At midline, food consumption declines and households increase their use of risky coping strategies, such as borrowing food and reducing meal size or frequency, resulting in a 0.10 SD decline in our food-security index. By endline, food consumption partially recovers, but reliance on risky coping strategies increases further, leaving the index 0.08 SD below that of households whose assistance continued. Psychological well-being falls by 0.10 SD at midline and 0.12 SD at endline. Trust also declines sharply: the overall trust index falls by 0.15 SD at midline, driven by a 0.25 SD decline in trust in aid agencies. The trust effects fall by roughly half by endline, but remain statistically significant at the 5% level. By contrast, we find little systematic evidence of effects on employment, social cohesion, or broader financial outcomes. Overall, the observed effects are sizable relative to a transfer worth about 18 percent of median household expenditure and they appear even larger, by endline, in areas that were not covered by the cash assistance.[5]

Second, the six-month transition generated only limited evidence of adjustment before exit and no clear reduction in welfare losses. At midline, while transition-package households were still receiving assistance, we find suggestive evidence of increased spending relative to continued-assistance households, with higher expenditure on food outside the WFP basket and on non-food items. We also find some weak evidence of labor-supply adjustment, driven by a drop in female labor supply in households with both working-age men and women. Taken together, however, these responses do not point to broad anticipatory adjustments. Once assistance ended, food security and psychological well-being deteriorated by magnitudes similar to those experienced by households exiting under the standard procedure. The main exception is trust, which declined less among transition-package households. A longer transition may therefore have improved perceptions of the exit process, but did not enable

[5]These effects correspond to approximately one-third of the cross-sectional food security gap and one-half of the psychological well-being gap between households in the top and bottom quintiles of the baseline wealth distribution, as measured among households in our sample assigned to continued assistance. For additional context, Crosta et al. (2026), in a meta-analysis of randomized evaluations of unconditional cash transfers in low- and middle-income countries, estimate average improvements of approximately 0.20 SD in both food security (95 percent credible interval: 0.2–0.3) and psychological well-being (95 percent credible interval: 0.1–0.2). This clearly remains an imperfect benchmark, since these studies examine program introduction rather than withdrawal and differ in program modality, context, time horizon, and measurement.

households to cushion its material consequences. This pattern is consistent with consumption-smoothing models in which households anticipate a loss but face liquidity or savings constraints that limit adjustment in advance. It is also consistent with behavioral frictions, such as present bias or limited attention, that delay costly adjustment until the loss of assistance becomes immediate.

Third, the choice among prioritization rules changed who received assistance without producing clear differences in average welfare. The four methods selected different beneficiary profiles and elicited different reactions, with the human-driven method generating more appeals against exclusion. Yet no method consistently outperformed the others in terms of average food security, psychological well-being, or other outcomes, including when focusing on the outcomes explicitly targeted by each method or on marginal households whose continuation status differed across rules. Agreement across methods, however, is informative. Households classified as eligible by relatively more rules had lower baseline asset wealth, were more likely to appeal when cut off from assistance, were less able to replace the lost transfer through higher food expenditure, and experienced larger declines in food security. Combining prioritization methods may therefore better identify households least able to absorb assistance withdrawal. This pattern could arise because agreement captures overlapping dimensions of vulnerability or because combining multiple noisy signals reduces classification error. Although we cannot pin down the exact mechanisms, we discuss suggestive evidence that the latter mechanism plays an important role.

This paper speaks to three strands of work. First, it contributes to the large literature on cash and in-kind transfers. A rich body of experimental studies provides extensive evidence on the effects of introducing transfers and on the consequences of different design choices such as transfer modality and delivery (e.g., Hidrobo et al., 2014; Hoddinott et al., 2018; Banerjee et al., 2023).[6] We study the opposite margin: what happens when a recurrent transfer is withdrawn from existing recipients, and how alternative scale-down decisions shape these effects? Causal evidence on this margin remains scarce and relies largely on naturally occurring program exits or cuts. Sterck and Bruni (2025) combine high-frequency data with randomized survey timing around a 20 percent cut in humanitarian assistance to refugees in Kenya and find sharp declines in food consumption, food security, and expenditure. Márquez-Padilla et al. (2025) exploit the nationwide rollback of Mexico's Prospera program with a difference-in-differences design and find declines in school enrollment accompanied by increased work, especially among male high-school students. Finally, Iqbal et al. (2026)

[6]See Crosta et al. (2024) for a meta-analysis of unconditional cash transfers, Jeong and Trako (2022) for a review of cash and in-kind transfers in humanitarian settings, and Gentilini (2016) for a review of studies comparing cash, vouchers, and food transfers.

use a regression discontinuity-in-differences design around a revised eligibility threshold for Pakistan's largest cash transfer program and find no evidence of adverse effects on women's empowerment, intimate-partner violence, or well-being one year after exit.[7] We complement this evidence by providing, to our knowledge, the first randomized evidence on the welfare consequences of withdrawing a recurrent humanitarian transfer from existing recipients.

Second, our "transition package" experiment contributes to the literature on household responses to predictable income changes. Under standard consumption-smoothing models, households with access to savings and credit should adjust when a future income loss becomes known rather than wait until it is realized (Friedman, 1957; Hall, 1978). Evidence from low-income settings, where liquidity and credit constraints are pervasive, is mixed. Some studies document substantial anticipatory responses to expected transfers or future income changes (Bianchi and Bobba, 2013; Banerjee et al., 2023) and successful smoothing of predictable income fluctuations (Angelucci et al., 2024), while others find delayed adjustments and incomplete smoothing (Edmonds, 2006; Aguila et al., 2017).[8] Most of this literature, however, concerns anticipated gains or predictable payment timing, rather than preparation for a known income loss.[9] Our experiment studies this latter margin directly. We find that extending the transition before assistance withdrawal induces only limited anticipatory adjustment and does not prevent subsequent losses in food security and well-being. This pattern is consistent with liquidity constraints and behavioral frictions that may delay adjustment (Duflo, et al., 2011; Karlan et al., 2016).

Finally, we contribute to the literature on targeting social assistance. In settings where income is difficult to observe, program targeting relies on approaches such as proxy means tests (PMTs), categorical rules, community-based selection, and self-targeting. Experimental comparisons reveal that PMTs tend to better identify households with low consumption,

[7]Rigorous evidence on the consequences of broader development program exit is similarly rare. A notable exception is Fishman et al (2022), who study the randomized phaseout of an agricultural extension and subsidy program in Uganda and find no decline in promoted agricultural practices or inputs, as farmers shift toward market-based alternatives. A separate literature follows households after temporary transfer programs end (Baird et al, 2019; Altındağ and O'Connell, 2023; Ahmed et al., 2025; Handa et al., 2025). These studies estimate whether initial program effects persist after support ends, rather than the causal effect of withdrawal relative to continued provision. For broader reviews of long-run program impacts, including persistence after interventions end, see Bouguen et al (2019), Molina Millán et al. (2019), and Grisolia (2024).

[8]Delayed payments provide complementary evidence of households' limited financial buffers: transfer delays reduce expenditure in Indonesia and food consumption and food security among refugees in Kenya (Bazzi et al., 2015; Bruni and Sterck, 2025).

[9]Related evidence suggests that liquidity can facilitate preparation for anticipated adverse shocks. Evidence from unemployment benefits in Brazil shows that households may fail to smooth consumption ahead of predictable benefit expiration (Gerard and Naritomi, 2021), while evidence from Bangladesh shows that providing resources before an anticipated climatic shock can induce preemptive action and improve subsequent welfare (Pople et al., 2026).

while alternative approaches may capture different notions of need or enjoy greater legitimacy among target populations (Alatas et al., 2012, 2016; Premand and Schnitzer, 2021; Sumarto et al., 2025). Recent work uses machine learning and novel data to improve poverty measurement and targeting and reduce screening costs (Blumenstock et al., 2015; McBride and Nichols, 2018; Aiken et al., 2022, 2025), and explores alternative welfare concepts or targeting objectives (Henderson and Follett, 2022; Haushofer et al., 2025; Aiken et al., 2026). Yet relatively little is known about how the choice of welfare measure used to construct a data-driven targeting rule translates into differences in realized program impacts.[10] We experimentally compare three data-driven prioritization rules trained on different definitions of need with an expert-designed categorical rule and evaluate them using realized outcomes. Our setting also shifts the targeting problem from entry to exit, allowing us to assess which rules best preserve assistance for households most vulnerable to withdrawal. We find little evidence that any one rule dominates in terms of realized welfare, but also show that agreement across distinct rules provides additional information about households' vulnerability to assistance withdrawal.

2. Setting and Program

2.1 Lebanon's Crisis and the Food Assistance Program

In recent years, Lebanon suffered the most devastating crisis in its modern history, with overlapping economic, political, and social dimensions (World Bank, 2022). The economic crisis started in August 2019, when Lebanese banks withheld deposits in US dollar accounts, and was compounded in the following months by two other major shocks: the Beirut Port explosion and the COVID-19 pandemic. The port explosion in August 2020 killed nearly 200 people, injured over 6,500, and displaced 300,000 individuals, while destroying critical infrastructures, including the country's primary grain silos. At the same time, the COVID-19 pandemic and the associated containment measures further deepened the economic downturn, resulting in widespread business closures, job losses, and a severe strain on Lebanon's already fragile healthcare system. Together, these shocks led to a dramatic collapse in purchasing power, which, in turn, led to a sharp rise in food insecurity across the country, increasing social tensions and exacerbating existing inequalities. Lebanon's nominal GDP per capita

[10] A relevant exception is Premand and Schnitzer (2021), who compare community-based targeting with a consumption-based PMT and a proxy formula designed to predict lean-season food insecurity in a randomized experiment in Niger, and assess welfare outcomes following the introduction of cash transfers. They find mostly no differences across welfare outcomes, with the exception of higher food stocks under the consumption-based PMT.

fell by more than 50% between 2019 and 2021 (World Bank, 2026), resulting in the country's downgrading in July 2022 to lower-middle-income status after decades as an upper-middle-income country (World Bank, 2022).

In 2020, as a response to these shocks, WFP, in collaboration with the Lebanese government, started providing emergency assistance in the form of monthly in-kind food parcels to Lebanese households in need. The initial beneficiary identification combined information from government agencies, partner organizations, and a self-enrollment portal, followed by WFP validation.[11] Beneficiaries of the program (which is currently still ongoing) are guaranteed the provision of monthly food assistance for at least the duration of an "assistance cycle" (12 months), which might then be renewed. Assistance is provided in the form of one parcel of food for every two household members, and the parcels are designed to provide the equivalent of at least 1,300 kcal/person/day. Each parcel includes about 20 kg of commodities and is worth $32 (at January 2023 market prices).[12] The median household in the program receives two parcels per month, which corresponds to approximately 42% of the median Lebanese household's monthly food expenditure and 18% of its total expenditure.[13] Households are informed each month of the arrival of their parcels through a message on their phone, which enables them to collect the assistance from one of the many collection points scattered across the country. WFP administrative data show that 89% of beneficiaries collect their parcels every month, and virtually everyone (99%) collects them at least half of the time within a cycle.

Since 2020, the program has been repeatedly renewed and expanded, reaching 58,501 households by 2023 (Figure A.2 shows the geographic distribution of beneficiaries across the country). However, due to limited funding and shifts in priorities, in 2024 WFP had to reduce the beneficiary caseload to about 25,000 households, i.e. only 43% of the existing caseload.[14]

[11] More specifically, WFP validated the inclusion of each household by conducting a short survey. In 2020 data collection and validation were based on phone surveys, while for later program expansion WFP relied on in-person surveys.

[12] Each parcel contains 10 items: rice (3 kg), bulgur (3 kg), pasta (4 kg), lentils (4 kg), beans (1 kg), chickpeas (1 kg), tomato paste (400 g), oil (1.8 L), salt (500 g), and sugar (1 kg).

[13] Authors' estimates based on the nationally representative Lebanon Household Survey (LHS) dataset, collected between January and May 2023.

[14] The 43% target was based on WFP's estimation of the number of beneficiaries with consumption levels below the minimum expenditure basket, which was set at USD 71 per capita per month (at January 2023 prices). This estimate was based on extrapolation from the nationally representative LHS dataset, collected between January and May 2023 and on a post-distribution monitoring survey that WFP conducted as part of its regular operations in the first half of 2023 among a representative subset of program recipients.

2.2 Communication and Appeals

Between August and October 2023, WFP collected census data from all 58,501 program recipients. The census provided the key input for the prioritization exercise, whose details are described below. Once prioritization decisions were made, in December 2023, WFP sent a message to all program recipients updating them on their status. Following WFP standard practice, communication was conducted entirely via SMS to ensure timely, direct, and consistent messaging across all program beneficiaries. The messages contained short, generic information on the inclusion or exclusion from the assistance program for the 2024 cycle. Households selected for continued assistance were informed that assistance would continue for them (at least) until 2025.[15] The remaining households were instead informed that their assistance would end as a consequence of a prioritisation exercise motivated by limited funding. Within this latter group, most households were informed that assistance would end in December 2023, while randomly selected transition-package households were informed that their assistance would continue until June 2024 (more details on the selection of these households below).

For households excluded from the program, the message also included information on how to submit an appeal either through WFP's dedicated phone line or through an online form. Households had a month from the notification to submit the appeal. Following standard practice, WFP reviewed each appeal to determine whether the household should be reassessed and/or readmitted. Households whose appeals were successful were informed of any resulting changes to their status and were readmitted into the programme. In practice, this happened very rarely.[16]

2.3 The Conflict with Israel and the Emergency Cash Program

During the study period, the country experienced yet another dramatic shock with the escalation of tensions with Israel along the southern border. Between September and November of 2024, the conflict caused thousands of fatalities and injuries, and led more than a quarter of the population to evacuate, with 1 million individuals internally displaced. A ceasefire agreement took effect on November 27, and more than 80% of the displaced population returned to their homes by December (IOM, 2024). In response to this new

[15] In practice, all selected households kept receiving assistance at least until June 2025, when a new prioritization exercise took place.

[16] The review process was entirely managed by WFP, without any intervention or assistance from the research team. Overall, 52% of households excluded from assistance submitted an appeal. By the time of our midline, in June 2024, less than 1% of the appeals had resulted in a change in status. An additional 2% of appeals were accepted, and households readmitted into the program, after July 2024.

crisis, in October 2024 WFP decided to provide temporary emergency cash transfers to all households that were: i) registered in the WFP system; ii) located in "high priority areas", as defined based on conflict and displacement intensity (Figure A.3)[17]; iii) not already supported through other WFP programs. This last condition meant that beneficiaries of the food assistance program were excluded from the emergency cash assistance.[18] Overall, most of the country (733 cadasters out of 1,236) fell into the "high priority" definition and as a consequence, most households assigned to stop receiving WFP food assistance in our study had received at least three months of emergency cash support by the time of our endline survey in January-February 2025. The progressive phase-out of the emergency cash assistance program began after the conclusion of our study, in February 2025. In the following sections, we will discuss in more detail how the emergency cash program might affect our endline analysis and how we address it.

3. Study Design

We structure the study design description into three parts. First, we focus on the random assignment of beneficiaries to the four prioritization methods. Second, we discuss how to move from experimental variation in the prioritization method to experimental variation in the discontinuation of assistance. Third, we discuss the randomization of the subset of discontinued households into the "transition package" arm.

Random Allocation into Different Prioritization Methods

The first and key step in our design is the random allocation of the beneficiary households into four equal-sized groups, each assigned to a different prioritization method. The four methods reflect distinct approaches (data- vs human- driven) and distinct concepts of household vulnerability, and are selected among approaches effectively used in the field by aid and humanitarian organizations to target their programs.

Three of these approaches are data-driven and based on the proxy mean tests (PMTs) methodology, a widely used approach for approximating latent household vulnerability in contexts where direct measurement is not feasible or is likely to be inaccurate. PMTs can be

[17]A given administrative area (cadastre) was deemed high priority area if it fulfilled any of the following three criteria: (i) it had a value of 3 or higher on the WFP Conflict Impact Index (which took into account the number of strikes, the number of fatalities, whether civilians were targeted, whether citizens had market access, whether the majority of the population had been displaced, and whether an evacuation order had been issued by the IDF in an area); (ii) it hosted at least 500 internally displaced persons; or (iii) more than 20 percent of the local population was internally displaced.

[18]The monthly cash transfer was equal to $25 per household plus $20 per household member and was thus somewhat more generous than the food support.

viewed as prediction functions that map observable household characteristics into proxies of underlying vulnerability. The PMT models must first be developed using a training dataset, ideally drawn from a population as similar as possible to the target population, that contains both the target vulnerability measure of interest and a set of socioeconomic variables expected to be predictive of it. Modern applications of this methodology rely on machine learning algorithms to select the most relevant variables associated with the outcome of interest and estimate their weights. The model is then applied to the target population to generate a predicted score for each potential beneficiary. Inclusion in the assistance program is then determined by ranking these scores and selecting households up to a cutoff defined either by the desired number of beneficiaries (as in our case) or by a specific predicted score.

In our setting, the three PMT models were trained on the Lebanon Household Survey (LHS) dataset, which was collected between January and May 2023, across a nationally representative sample of 3,597 households. The LHS dataset records multiple welfare measures and more than one hundred socio-economic variables. The three PMT models differ in the vulnerability concept they seek to predict. The first model (*PMT-Exp*) targets per capita expenditure, a common proxy for poverty. The second model (*PMT-FCS*) targets food consumption, as measured through the Food Consumption Score (FCS), which captures short-run dietary diversity and frequency of consumption, and can thus be interpreted as a proxy for food needs. The third model (*PMT-CARI*) targets a multidimensional vulnerability index based on the "CARI framework" developed by WFP, which combines expenditure, food consumption, and coping capacity, to capture broader constraints on household welfare (WFP, 2021). The comparison across these PMT specifications thus provides a test of how alternative definitions of vulnerability (poverty-based, food needs-based, or multidimensional) affect targeting and subsequent outcomes. Our setting is particularly well suited to comparing these alternative approaches, as we can rely on a recently collected training dataset from a similar population that includes multiple measures of vulnerability and a rich set of household characteristics. Our PMT models are estimated from 126 distinct candidate predictors, augmented with interactions and nonlinear terms.[19]

The fourth targeting method (*Categorical*) is instead a score-based categorical approach based on expert opinions. Rather than estimating a statistical mapping from observables to vulnerability measures, categorical targeting usually defines eligibility based on a set of household characteristics ("categories") identified by the decision-makers. The central design

[19]These set of 126 distinct socioeconomic variables include all variables collected in the LHS that were also available in the WFP census conducted a few months later among the beneficiary population. Restricting the predictor set to variables observed in both datasets was needed to ensure that the resulting measures could be replicated in the prioritization exercise.

choice in categorical targeting lies in selecting the variables that define eligibility. In our case, we used a score-based approach based on expert judgement. A survey was administered by WFP in November 2023 to a pool of 28 experts with operational or research experience in food security and humanitarian assistance in Lebanon. After receiving brief background information, the experts were asked to identify up to 10 household characteristics that should make households eligible (or ineligible) for continued food assistance. Their answers were then pooled to define a scoring system, with weights reflecting both the frequency with which each characteristic was selected and its rank in the responses.

While the four different vulnerability definitions are correlated, their overlap is far from perfect. When ranking households according to the three PMT outcomes, within the LHS sample, the average pairwise correlation in the ranking is only 0.39. And when comparing those rankings to the one based on the categorical score, the average pairwise correlation drops to 0.24. More details on the definition of the four methods and on their overall performances are reported in Appendix B.

In November 2023, the 58,501 assistance recipient households were individually randomly assigned to four prioritization groups of equal size (Figure A.4).[20] Within each group, households were ranked based on the specific prioritization method assigned to that group, and the 43% most vulnerable households were selected to keep receiving assistance (*continued* assistance households), while the remaining 57% were excluded from the program (*discontinued* assistance households). Table A.1 in the Appendix shows that the initial randomisation created four observationally equivalent groups.[21] Importantly, the balance in the composition of the four groups does not translate into balance in the profiles of households selected for continued assistance (Table A.2), reflecting the fact that different targeting approaches select different household profiles. When considering the full set of baseline characteristics, statistical tests reject the null hypothesis of equality of means across the selection in the four groups for 87 percent of the variables at the 5-percent significance level. Even when excluding the categorical approach and restricting the analysis to only the three PMT approaches, 78 percent of the variables show statistically significant differences at the 5-percent level. These differences can be further confirmed by simulating the application of each method to the

[20] In practice, the four methods were first simulated across the full population and the randomization was then stratified by eligibility status under the four methods, to ensure balance along this dimension. In other words, the randomization was stratified by the 16 "eligibility spaces" that will be better discussed in the following section.

[21] This table, like all other balance check tables, reports only a selected subset of 20 variables. When considering the entire set of variables included in the baseline census and the four targeting score variables, there are a total of 160 possible comparisons. The notes below each balance table report the statistics when considering all comparisons.

entire population and comparing the selection outcomes (i.e., who is included in the 43% most vulnerable). Although there is a positive correlation in the selection across the four methods, it is low: on average, just 0.32. As one might expect, the three PMT-based methods are relatively more similar, but even in their case, correlation in their final selection is only 0.55, on average.

Experimental Variation in the Discontinuation of Assistance

Differences in selection across prioritization methods imply that random assignment to a method generates exogenous variation in assistance continuation. To characterize this variation, we apply all four methods to the full population and classify each of the 58,501 households into one of the 16 "eligibility spaces" shown in Figure 1. Each space groups households with the same pattern of eligibility across methods, ranging from E0, for households selected by none of the four methods, to E1E2E3E4, for households selected by all four. Intermediate spaces capture disagreement across methods. For example, E1E2 includes households selected for continued assistance under methods 1 and 2 but not 3 and 4. Table A.3 in the Appendix shows that households are distributed across the full eligibility spectrum, with at least 458 households in every space. For 39 percent of the sample, the methods agree: 26 percent of households are excluded by all four methods (space E0) and 13 percent are selected by all four (space E1E2E3E4). For the remaining 61 percent, eligibility depends on the method applied. Random assignment to a prioritization arm therefore generates random variation in assistance continuation within these eligibility spaces. Table A.4 confirms that, conditional on eligibility-space fixed effects, households assigned to continued and discontinued assistance are well balanced. We can thus identify the causal effect of assistance withdrawal among households eligible under some, but not all, prioritization methods by comparing continued and discontinued households within eligibility spaces.

Random Allocation into the Transition Package Arm

As a final step, we randomly selected a subset of households assigned to discontinued assistance to receive an additional six months of support. Within the 14 eligibility spaces in which the prioritization methods disagree, we identified the 28 combinations of eligibility space and prioritization method that led to the discontinuation of assistance (indicated in red in Table A.3), and we randomly selected 35 households from each of these cells, yielding a total of 980 households assigned to the transition package.

This step therefore identifies three groups of households: households assigned to *continued* assistance, which received food assistance throughout the study period; households assigned to *regular discontinuation*, whose assistance ended in December 2023; and households assigned to the *transition package*, which continued receiving assistance through June 2024. Table A.4

confirms that, conditional on the inclusion of eligibility space fixed effects, groups are well balanced in the characteristics of households assigned to the transition package (TP) and households in the other two groups. As mentioned earlier, our midline data collection took place while the TP households were still receiving their assistance. By comparing continued assistance and TP households, we can thus learn whether and how households were preparing for the upcoming end of their assistance. By endline, instead, TP households had stopped receiving their food assistance. The comparison with the regular discontinuation group at endline can thus reveal the effect of the full "transition package", which includes the longer notice window and the six extra months of assistance.

Ethical Considerations

Although part of the analysis of this study builds on experimental variation in the discontinuation of food assistance, the inclusion in the discontinued assistance group was always based on the assessment of a household's vulnerability, as defined by the different prioritization methods. The four methods were developed in collaboration with WFP, to reflect distinct approaches and vulnerability concepts, based on actual methodologies that WFP and other aid agencies already apply for regular targeting activities across their programs. Households ranked in the 43% most vulnerable group by each targeting method continued to receive assistance, while the remaining 57%, deemed relatively less vulnerable, stopped receiving it. The random variation in the discontinuation of assistance arises solely from differences in targeting decisions among the four methods. In line with WFP's regular practices, households excluded from assistance always had the option to appeal the decision via the dedicated hotline or an online form. Appeals were handled directly by WFP in accordance with its standard procedures, without any involvement from the research team. Finally, the definition of a random transition-package sample implied that a subset of households whose assistance should have ended in December 2023 continued to receive it for six additional months. The study design thus never accelerated assistance discontinuation. The study received ethical clearance from the ethics committees of Humboldt University of Berlin (#2023-08-W) and from HML International Review Board (#2766).

4. Sample, Timeline, and Data

4.1 Sampling

The study sample is a stratified random sample from the overall beneficiary population. We randomly selected 100 households from each of the 64 cells defined by the combination of the 16 eligibility spaces and the four targeting methods (Table A.3). By construction, out

of the total sample of 6,400 households, half were selected for continued assistance, and the other half was instead assigned to stop receiving assistance. This latter group includes 2,220 regular-notice households and the 980 (randomly selected) households that received the transition package. By giving equal representation to each cell, this sample maximises power for the comparison between households assigned to continued and discontinued assistance, though it does not provide a representative distribution of the population across eligibility spaces. By applying the corresponding population weights, it is however straightforward to retrieve population-wide estimates (see below).

4.2 Timeline

Figure 2 summarizes the study timeline. In preparation for the prioritisation exercise, between August and October 2023, WFP collected a census among all 58,501 beneficiaries. In December 2023, following the completion of the targeting exercises, all households received a message from WFP to update them on their status. The message was sent to the same mobile phone number used to inform households about their food parcels. The message either confirmed that assistance would continue (at least) until 2025 or that assistance would end. Within the latter group, most households were informed that assistance would end in that same month, while the TP households were informed that their assistance would continue until June 2024. From January 2024, regular-notice households stopped receiving WFP assistance. The TP households were then reminded of the approaching end of assistance through a new message in April 2024.

The first round of data collection from the study sample (midline) took place in May-June 2024, just before the end of assistance for the TP households. The second round of data collection (endline) took place in January-February 2025. All continued assistance households kept receiving food assistance beyond the study period, at least until June 2025, when WFP implemented a new program adjustment. In October 2024, in response to the escalation of the conflict between Lebanon and Israel, WFP started providing the emergency cash transfers. The second round of data collection took place while this WFP cash emergency support was ongoing.

4.3 Data Sources

We make use of four main data sources.

LHS Database – This database (previously called Lebanon Vulnerability Assessment Panel or LVAP database) is the result of a collaborative initiative between the World Bank, UNHCR, and WFP, with support from the Joint Data Center on Forced Displacement, that aims to monitor and analyse the socioeconomic conditions of the Lebanese population. It is

structured as a nationally representative household survey of approximately 3,500 households. The dataset is very rich and records detailed information on poverty levels, food security, employment status, access to essential services, and coping mechanisms. A data collection round took place between January and May 2023, just a few months before the start of our study. We rely on this dataset to train the three PMT prioritization models.

Household Census – The census was collected by WFP between August and October 2023, to inform its forthcoming prioritization exercise. The census covered the full set of 58,501 beneficiaries and contained more than 150 variables, capturing household characteristics, demographics, and assets. It did not contain any direct vulnerability measures. This decision was driven by both the need to limit the length of the survey and the high risk of strategic misreporting on direct vulnerability-related questions. The census serves two key purposes. First, it provides up-to-date information on the entire population of beneficiaries, which was essential for implementing the prioritisation. Second, it provides the baseline information for our study sample.

Administrative Data – WFP maintains detailed administrative records of its operations, collected independently of this study. We use two main sets of records in our analysis: (1) monthly data on food parcel collection, which allow us to assess adherence to the study protocol; and (2) records of appeals submitted by phone or online, which we use to measure beneficiaries' responses to exclusion decisions. For the endline analysis, WFP also provided information on areas prioritized for emergency support during the conflict with Israel and records of households that received emergency cash transfers.

Survey Data – Our main analysis draws on the two rounds of survey data collected from the study sample (6,400 households). The questionnaires were developed by the research team to cover a broad range of household characteristics and welfare measures, detailed in the following section. Data collection was implemented by the Central Arab Statistical Network (CASN), an independent survey firm with extensive experience in large-scale data collection across Lebanon. Overall, over 60 enumerators were deployed across the country for both rounds of data collection. Survey teams were familiar with local customs and were masked to the assignment status of the households they were interviewing. Surveys were always conducted in-person, using SurveyCTO, a secure data collection tool built on Open Data Kit (ODK).

4.4 Outcome Measures

We group our main survey outcomes into four families: expenditure, food security, psychological well-being, and trust. We describe these measures below, along with additional outcomes,

related to household finances, employment, and social cohesion. Appendix C provides the corresponding survey questions and full details on variable construction.

Expenditure – We measure per capita monthly food and non-food expenditure. Food expenditure is based on a 30-day recall across nine food categories and includes both amounts paid at the time of purchase and purchases made on credit, as well as the estimated market value of own-produced food.[22] Non-food expenditure covers ten regularly purchased items using a 30-day recall and six less frequent categories using a six-month recall. For the analysis, we convert the latter category to monthly equivalents. All expenditure measures are transformed into US dollars.

Food Security – We construct our food security index by combining two standard measures widely employed by WFP: the Food Consumption Score (FCS) and the reduced Coping Strategies Index (rCSI). FCS combines information on the frequency and nutritional value of foods consumed by the household over the previous seven days, with higher values indicating better food consumption (WFP, 2026). The reduced Coping Strategies Index (rCSI) measures the frequency and severity of five behaviours used to cope with food shortages, such as reducing meal size or frequency, with higher values indicating greater reliance on adverse coping strategies (Maxwell and Caldwell, 2008). We construct our food security index by reverse-coding the rCSI and combining it with the FCS using inverse-covariance weights, following Anderson (2008).

Psychological Well-Being – We construct our psychological well-being index by combining three standard measures: the life-satisfaction index, the worry index, and the WHO-5 Well-Being Index. The life-satisfaction index combines five questions on satisfaction with food security, health, household finances, community relations, and life overall. The worry index, adapted from Hidrobo et al. (2023), combines ten questions covering concerns about basic needs, debt, employment, family, conflict, and other sources of insecurity. Finally, the WHO-5 Well-Being Index measures mental well-being through five questions on positive affect and functioning over the previous week (WHO, 2024). Higher raw values indicate worse outcomes for life satisfaction and worry and better outcomes for the WHO-5. We reverse-code the first two measures before constructing the summary psychological well-being index using inverse-covariance weights, following Anderson (2008).

Trust – Respondents report their level of trust in seven groups and institutions: family, neighbors, religious organizations, international organizations and NGOs, local government,

[22]The value of each type of home-grown food consumed in the household was estimated based on local market prices, as reported by the households. These assessments are expected to be noisier than market purchases, but are unlikely to play a major role in our analysis, as the population mostly lives in urban areas and only 3.2% reports consuming home-grown food.

national government, and migrants. Responses are recorded on a four-point scale and coded so that higher values indicate greater trust. We construct the trust index by combining the seven items using inverse-covariance weights, following Anderson (2008).

Additional Outcomes – We examine two additional dimensions of household financial circumstances: whether the household has outstanding debt with an institution or individual and whether it received cash remittances from abroad during the survey reference period.[23] We also construct two employment outcomes from the household roster. We classify individuals as employed if they engaged in at least one hour of wage or own-account work during the previous seven days. Our main household-level employment measures are the percentage of household members aged 15 or above who are employed and weekly hours worked per household member aged 15 or above. Finally, we build a social-cohesion index based on Bossuroy et al. (2022), combining three measures: household membership in a community group or association, participation in community or neighborhood volunteering, and perceived tension within the community. We reverse-code community tension so that higher values consistently indicate greater cohesion and combine the three components using inverse-covariance weights, following Anderson (2008).

5. Empirical Strategy

5.1 Assistance Discontinuation and Transition Package

Our main analysis of the effects of assistance withdrawal compares households assigned to discontinued food assistance with households assigned to continued assistance, within the same eligibility space. Because the conflict-related emergency cash assistance may affect our endline results, we estimate the effects separately at midline and endline using the following specification:

$$\begin{aligned} Y_{h,e,t} = \beta &Discontinuation_{h,e} \\ &+ \gamma TransitionPackage_{h,e} + \eta_e + \mu_{h,e,t} \end{aligned} \tag{1}$$

where $Y_{h,e,t}$ is the outcome of interest (e.g., food security) for household h, belonging to eligibility space e, at time t with $t \in \{Midline, Endline\}$. $Discontinuation_{h,e}$ is an indicator variable for households assigned to stop receiving food assistance, while $TransitionPackage_{h,e}$ is an indicator variable for the subset of discontinued households randomly assigned to receive the transition package. Importantly, the inclusion of eligibility space fixed effects η_e ensures

[23] The survey initially also included a question on savings, which was a pre-specified outcome, but we dropped it after piloting because virtually no respondent reported positive savings.

that the estimation relies on variation in assistance *within* eligibility spaces.[24] This also implies that the estimated effects hold for households that belong to eligibility spaces where there is variation in assistance status across prioritization methods. We restrict the estimation sample accordingly: households in E0 and E1E2E3E4 are excluded because assistance status does not vary within those spaces and they therefore contribute no identifying variation. The coefficient β captures the average impact of being assigned to stop receiving food assistance under the regular (short-notice) protocol on the outcome Y.

The same specification is used to study the impact of providing the transition package. Given that TP households are always a subset of the households assigned to stop receiving assistance, the coefficient γ captures the *differential* impact of the provision of the package on outcome Y, compared to regular discontinued households. To learn about the comparison between TP and continued assistance households, we can sum $\beta$and γ. The interpretation of the transition package effect differs by wave. At midline, TP households still receive assistance, so $\beta + \gamma$ tells us whether knowing assistance will end, while still receiving it, changes outcomes relative to households assigned to continue receiving assistance. At endline, TP households have exited the program, so $\beta + \gamma$ tells us the (short-term) consequences of the end of assistance for households that had six months to prepare for it. The comparison with regular discontinued households (γ) will instead reveal the impact of the transition package, which combines the provision of six additional months of food assistance and a longer notice window. Randomization was conducted at the household level and we report robust standard errors to account for heteroskedasticity.

As discussed above, interpretation of the endline discontinuation effects is complicated by the emergency cash assistance introduced during the escalation of the conflict with Israel. Emergency cash was provided to households in WFP-defined high-priority areas that were no longer receiving food assistance. This means that households assigned to assistance discontinuation located in non-priority areas remained without any WFP support, while those in priority areas subsequently received the emergency cash. We therefore also estimate a fully interacted specification of the form:

[24] As a robustness check, we also re-estimate all our specifications using post-double-selection LASSO, following Belloni et al (2014) and drawing on a pool of 156 baseline covariates. The estimated treatment coefficients remain virtually unchanged, while median standard errors decline by approximately 7 percent. Because the use of data-driven covariate selection was not prespecified, we retain the design-based specifications without LASSO selected covariates as our preferred estimates. This is also in line with recent evidence that post-double-selection generally delivers limited gains in field experiments and is most useful as a robustness check (Cilliers et al, 2024; Burlig et al, 2026).

$$
\begin{aligned}
Y_{h,e,t} = \beta_P(&Discontinuation_{h,e} \times Priority_h) \\
&+ \beta_N \left(Discontinuation_{h,e} \times (1 - Priority_h)\right) \\
&+ \varphi Priority_h + X_{h,e}\tau + \mu_{h,e,t}
\end{aligned} \tag{2}
$$

where $Priority_h$ indicates residence in a high-priority area, and $X_{h,e}$ includes the usual eligibility-space fixed effects and transition package indicator, plus their interaction with the priority-area indicator. The coefficient β_N captures the effect of food assistance discontinuation among households in non-priority areas, which received no subsequent emergency cash, while $\beta_P$captures the effect among households in priority areas, which subsequently received emergency cash. Both effects are measured relative to households assigned to continued food assistance within the corresponding areas. We estimate equation (2) separately by survey round. At midline, before emergency cash was introduced, comparing β_N and β_P provides a diagnostic of whether discontinuation effects already differed across areas that would later receive different emergency responses. At endline, the same comparison shows whether the effects subsequently diverged across the two groups.[25]

5.2 Prioritization Approaches

To compare prioritization methods, we make use of our full sample and estimate the following basic equation:

$$
\begin{aligned}
Y_{h,e} = &\alpha_1 PMT_FCS_{h,e} + \alpha_2 PMT_CARI_{h,e} \\
&+ \alpha_3 Categorical_{h,e} + \theta_e + \varepsilon_{h,e}
\end{aligned} \tag{3}
$$

where $PMT_FCS_{h,e}$ and $PMT_CARI_{h,e}$ are indicator variables for households assigned to the PMT arms that target FCS or CARI, respectively, while $Categorical_{h,e}$ is an indicator variable for households assigned to the expert-driven categorical arm. The PMT model that targets per-capita expenditure is the excluded reference category. We always include eligibility-space fixed effects (θ_e) to account for the stratified randomization and improve precision.

The coefficients α_i capture the average difference in the outcome Y between households assigned to a given prioritization method (PMT targeting FCS, PMT targeting CARI, or categorical) and those assigned to the reference prioritization method (PMT targeting

[25]We cannot directly interpret any endline divergence as the causal effect of emergency cash, given that priority areas were differentially exposed to the conflict and high-priority status was not randomly assigned. The presence of differential effects emerging at endline would therefore be consistent with emergency cash affecting (attenuating) the consequences of withdrawal, but also with differential exposure to the conflict.

expenditure). By comparing $\alpha_i$coefficients, it is also possible to assess the relative performance across all methods.

When considering the number of appeals, equation 3 can be run on the entire population of 33,341 recipients who were excluded from assistance. For all other outcomes, instead, information is only available for the study sample of 6,400 households.

5.3 Weights and Multiple Hypotheses

For transparency and ease of interpretation, our main specifications report unweighted estimates. Two sets of weights can be used to translate these estimates into average effects for the broader population. First, calibrated population weights can be used to account for the stratified sampling design and to reweight the survey sample to the full 2023 recipient population.[26] Second, in the discontinuation analysis, treatment-probability weights can be applied to account for differences across eligibility spaces in the probability of assistance discontinuation.[27] Our results and conclusions are substantively unchanged under either weighting scheme (Appendix A).

Given the multiple outcomes and to reduce concerns about multiple hypothesis testing, as explained above, we follow Anderson (2008) and construct an inverse-covariance-weighted index for each family of closely related outcomes. We additionally report false-discovery-rate-adjusted q-values using the Benjamini and Hochberg (1995) procedure. For each survey wave and treatment comparison, adjustments are applied across the four primary outcomes for our main results and within outcome families for the analysis reported in the Appendix.

6. Descriptive Statistics and Internal Validity

Table 1 reports baseline characteristics for households in our sample, separately by prioritization arm. Households have fewer than four members on average (median: 4) and fewer than one school-age child (median: 0). About one-third of households are headed by a woman. The socioeconomic variables point to substantial vulnerability: about 60% of household heads have at most primary education, 74% of households include a member with a chronic illness,

[26]For population-weighted estimates, we construct wave-specific weights in two steps. First, within each targeting-arm-by-eligibility sampling cell, respondents receive the ratio of the number of households in that cell in the full beneficiary population to the number of surveyed households from that cell in the relevant wave (midline or endline). Second, we calibrate these design weights using raking so that the weighted respondent sample matches known population margins from the administrative data. The calibration margins include the 64 sampling cells, governorate, household-size bins, wealth quintiles, appeal status by targeting arm, and selected baseline characteristics (Deville and Särndal, 1992).

[27]To implement this correction, we follow the fully interacted regression-adjustment logic of Lin (2013), as suggested by Humphreys (2025).

and about one-third have no regularly employed household member. Asset ownership and access to basic services are also limited: only 30% of households own a car and 24% have piped water, although smartphone ownership and access to gas are more widespread (81% and 73%, respectively). Overall, Table 1 depicts a vulnerable population in a middle-income setting.[28] The bottom part of the table summarises the scores associated with the four targeting methods. The three PMT-related scores are predictions from the respective models, while the categorical score is the result of applying the weights derived from the expert-based surveys.

Importantly, Table 1 shows that household characteristics are very similar across the four prioritization groups. Table A.6 in the Appendix provides a similar comparison across households assigned to continued assistance, and to regular discontinuation or transition package (while controlling for eligibility space). Taken together, these tables support our study design and confirm that the randomization created observationally equivalent groups across the different comparisons we are interested in.

In terms of attrition, when considering the sample of continued assistance and regular discontinued households, we successfully surveyed 99% of the target sample at midline and 94% at endline, with no differential attrition across prioritization arms or assistance status (Table 2, Panels A and B). Whenever a target household could not be surveyed, we replaced it with another household randomly selected from the same eligibility cell. Our main analysis focuses on the full sample, but all our results are robust to excluding replacements. When it comes to the transition package households, however, at midline we surveyed only 31% of them due to an error in the list of households used by the survey firm.[29] Although substantially smaller, this subsample was still randomly selected and remains well balanced across arms (Table A.7). At the endline, the mistake was corrected, and we surveyed 93% of the TP sample.

We verify program implementation and compliance using both administrative and survey data. WFP systematically records the collection of each food parcel. Administrative records show that 97% of households assigned to continue receiving assistance collected their parcels during the study period, while 99% of households assigned to discontinue assistance did

[28]Table A.5 in the Appendix compares these characteristics to the nationally representative LHS dataset. The table shows that, as expected, program beneficiaries are more vulnerable than the average Lebanese household (e.g. they have more household members, are more likely to be headed by a woman, more likely to suffer from health issues, and have lower wealth).

[29]TP status was assigned through a stratified randomization in December 2023, when WFP communicated households' assistance status. The survey sample was selected through a separate stratified randomization. By mistake, this second randomization did not preserve the previously selected TP households and instead drew them again. As a result, only 31 percent of TP households were also selected into the survey sample at midline.

not. We also collected information from the households themselves during our midline and endline surveys. While such data are noisier, potentially due to recollection error or strategic misreporting, they mirror the expected pattern, with 83% (80%) of the households expected to keep receiving food assistance reporting it at midline (endline), and 89% (76%) of households expected to stop receiving food assistance reporting no assistance by midline (endline).

Overall, balance, attrition, and implementation checks reassure us on the integrity of the original experimental design, with two important caveats. First, the error in the midline data collection means that our midline analysis for TP households is underpowered compared to the original design. In discussing those results we will therefore place relatively more emphasis on the point estimates than on statistical significance. Second, and as discussed above, the emergency cash provided by WFP is expected to influence our endline estimates. To address this challenge, we will discuss and report in the Appendix the alternative specification that distinguishes between priority and non-priority areas.

7. Results

7.1 The Consequences of Assistance Withdrawal

We begin by examining how households adjust their food expenditure after losing assistance. This analysis is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance (i.e., excluding E0 and E1E2E3E4). Figure 3 reports the estimated coefficient of interest from equation (1), with the associated 95% confidence intervals, separately for midline (blue) and endline (red). The underlying regression results are reported in the Appendix (Table A.8). In the short run, households increase monthly food expenditure by USD 12.5 on average. This amounts to about 39% of the market value of one food parcel and, since the median household received two parcels per month, replaces about 19% of the value of the lost assistance. By endline, the increase in total food expenditure remains at just over USD 10 per month. The figure shows that these increases are concentrated in food categories previously provided by WFP. In the short run, households seem to partly offset the higher spending by reducing expenditure on other foods. In the longer run, expenditure on other foods recovers, while spending on WFP-related food categories declines somewhat, particularly for pulses.

How do households cover their increased food expenditure? The decomposition by payment mode shows that households increase food purchases on credit by approximately 4 USD at midline and 3 USD at endline (Table A.9). Purchasing food on credit is a relatively widespread practice in Lebanon, which became more common during the economic crisis, when about half of the population reported purchasing at least some food on credit (WFP,

2022). Within our continued assistance group, households report purchasing about 19% of their food on credit, on average. There is also some evidence of increased non-food purchases on credit at endline for the discontinued group, while the corresponding midline estimate is imprecisely estimated. Overall, credit appears to have financed about one-third of the adjustment. For the remaining part, our data do not permit a complete accounting, as we do not observe detectable changes across non-food expenditures, other financial dimensions, or employment (Table A.10).

We next assess whether the increase in food expenditure was sufficient to preserve household food security, considering both food consumption and reliance on food-related coping strategies. Our food consumption data are coarser than our food expenditure data, as they only record whether and how frequently households consumed different food categories over the previous seven days, not the quantities consumed. Panel A of Figure 4 shows that in the short run, food consumption falls by 0.10 SD, driven by lower consumption of items previously provided through WFP assistance (the corresponding regression estimates are reported in Table A.11).[30] By endline, aggregate food consumption largely recovers, although a large and statistically significant gap remains for pulses (-0.13 SD). Panel B examines reliance on food-related risky coping strategies (the corresponding regression estimates are reported in Table A.12). Households losing assistance increase their reliance on these strategies across the board. The short-run effects are generally small and imprecisely estimated, but by endline households are more likely to borrow food (+0.11 SD), reduce meal frequency (+0.07 SD), and reduce adults' food consumption (+0.07 SD), resulting in a 0.11 SD deterioration in their coping-strategies index.

Overall, the results indicate that assistance withdrawal substantially worsened household food security. Figure 5 summarizes the effects on our main outcomes in standardized units, including the food security index that combines food consumption and adoption of risky coping strategies (the corresponding regression estimates are reported in Table A.13). The top two estimates confirm the patterns documented above: food expenditure increases by 0.16 SD at midline and 0.14 SD at endline, while food security declines by 0.10 SD and 0.08 SD, respectively.

Figure 5 also shows a deterioration in psychological well-being of similar magnitude at midline (-0.10 SD). When disaggregating the index (Table A.10), the estimates indicate lower life satisfaction and mental well-being and greater worry. These effects persist at endline,

[30]For this analysis, standardized outcomes are standardized within each survey wave using the mean and standard deviation of the continued assistance group, which serves as the reference group.

when psychological well-being declines by 0.12 SD, driven primarily by lower life satisfaction and mental well-being.

Finally, Figure 5 shows that assistance withdrawal significantly reduces trust at midline (-0.15 SD), with the effect halving by endline (-0.06 SD). Figure 6 and Table A.14 disaggregate the trust index and show that the decline is driven primarily by a large and persistent reduction in trust in aid organizations (international organizations and NGOs).[31] At midline, we also observe smaller declines in trust in local government and migrants. The former may reflect some uncertainty about responsibility for assistance withdrawal, while the latter may reflect some resentment toward groups that continued to receive support through other assistance programs unaffected by the cuts. By endline, however, these effects fade and the persistent decline is concentrated in trust in aid organizations, which remains 0.13 SD lower. This persistence is notable given that by endline, most households excluded from the food assistance program in our sample had subsequently received emergency cash assistance from WFP. This result suggest that the loss of trust associated with assistance withdrawal is difficult to recover and can persists even after support is subsequently restored in another form.

Taken together, our results show that households losing assistance replaced only a limited share of the lost transfer through higher food expenditure, partly financed through credit. This response was insufficient to prevent a deterioration in food security and may have created additional payment obligations that could increase longer-run vulnerability. Assistance withdrawal also worsened psychological well-being and reduced trust in aid institutions, with both effects persisting over time.[32]

Heterogeneity

We test for heterogeneity along three pre-specified dimensions capturing distinct forms of vulnerability: baseline poverty (measured by the asset index), household size, and the gender of the household head. The results, reported in the Appendix (Tables A.15–A.17), suggest that wealthier and larger households increase food expenditure more following assistance withdrawal, consistent with differences in spending capacity and household needs. Effects on food security and other dimensions of well-being, however, are broadly similar across groups. The point estimates also suggest that female-headed households increase food expenditure less and experience somewhat larger declines in well-being than male-headed households,

[31] International Organizations (IOs) and NGOs were grouped together in the survey question, as piloting revealed that households often struggle to distinguish the two. This is not surprising, given that WFP and other IOs often rely on local NGOs to deliver their programs.

[32] Figure A.5 in the Appendix reports the full distributions of the primary outcomes by assistance status group, separately at midline and endline.

consistent with greater vulnerability. These differences, however, are generally small and imprecisely estimated.[33]

Additional Checks and Channels

Weights – Results remain similar when using standard population weights to account for the stratified sampling design (Tables A.18-A.19) and when accounting for the fact that the likelihood of discontinuation varies across eligibility spaces (Tables A.20-A.21).[34]

Local spillovers – We next examine whether the consequences of assistance discontinuation depend on the assistance status of nearby beneficiary households. This exercise can be interpreted as a test for local spillovers and speaks to an important feature of program scale-downs: unlike complete program termination, they withdraw assistance from some households while maintaining it for others. Nearby households that continue to receive assistance could attenuate the consequences of discontinuation if families share food, credit, or other resources. Conversely, continued support to nearby households could make exclusion more salient, generate unfavourable comparisons or perceptions of unfairness, and thereby amplify its adverse effects. Using WFP administrative data on beneficiaries' locations, we calculate, for each household, the total number of other program households located within a 200-meter radius, as well as the number that continue to receive assistance.[35] Conditional on the total number of nearby beneficiaries, experimental variation in household-level assistance status generates variation in the local prevalence of assistance. Table A.22 in the Appendix reports the results. The first column considers the full sample of discontinued households and uses administrative data to assess whether the probability of appealing the discontinuation decision varies with the number of neighbouring households that continue to receive assistance.The specifications always control for the total number of program households in the same radius,

[33]As an additional test of heterogeneity, we apply the generic machine-learning framework of Chernozhukov et al. (2025) separately to each primary outcome and survey wave. Using 156 baseline covariates from the census data, we construct out-of-sample treatment effect scores using five repetitions of five-fold cross-fitting, with a random forest T-learner as our primary method and a lasso T-learner as a robustness check. We assess heterogeneity using best linear predictors and sorted group average treatment effects across quintiles of the predicted treatment effect distribution. We find no robust evidence that the effects of assistance discontinuation vary systematically with observed baseline characteristics.

[34]The tables show estimates using a centered-strata interaction adjustment (Lin, 2013; Humphreys, 2025). Rather than including only eligibility-stratum fixed effects, this specification interacts assigned discontinuation with each stratum indicator, after centering the indicators at their sample means within each outcome-wave sample. This allows discontinuation effects to vary flexibly across eligibility strata. The reported discontinuation coefficient corresponds to the sample-share-weighted average effect across strata.

[35]After cleaning the location data, valid GPS coordinates are available for 86 percent of the full beneficiary sample. For the remaining 14 percent, only the administrative unit (cadaster) of residence is known, and these households are therefore excluded from the spatial analysis. Results remain very similar when using alternative radii of 100 or 500 meters.

thereby accounting for differences in local beneficiary density. The estimates indicate that each additional nearby household that continues receiving assistance (the average is 7.4 households) increases the probability of appealing the exclusion decision by about 0.2 percentage points. The effect is modest, but precisely estimated. The remaining columns consider the study sample and show instead no evidence that the effects of discontinuation on the main welfare outcomes systematically vary with the number of nearby beneficiaries who remain on the program.[36]

Emergency cash assistance at endline – As discussed above, interpretation of the endline discontinuation effects is complicated by the emergency cash assistance introduced by WFP during the escalation of the conflict with Israel. Discontinued households in WFP-designated priority areas subsequently received temporary cash support, whereas those in non-priority areas remained without WFP assistance. We therefore allow the experimental effect of food-assistance discontinuation to vary by priority-area status, using the fully interacted specification (2) described above. The estimates capture the effect of discontinuation under two distinct treatment histories: withdrawal without subsequent WFP support and withdrawal followed by emergency cash assistance. Table A.24 in the Appendix reports the results. At midline, before emergency cash was introduced, the effects of discontinuation are similar across future priority and non-priority areas. This means that households in the two sets of areas were not responding systematically differently to assistance withdrawal before the emergency response. By endline, differences emerge for most outcomes. Among discontinued households in non-priority areas, psychological well-being declines by 0.24 SD and trust by 0.16 SD, compared with substantially smaller declines among discontinued households in priority areas (that subsequently received emergency cash). By contrast, the effects on food security remain similar across the two groups. These patterns are consistent with emergency cash attenuating some consequences of assistance withdrawal, particularly for psychological well-being and trust.[37] They also indicate that the average endline effects reported above mask larger losses among households that remained without WFP support during the new

[36]When we restrict the focus to *similar* households, i.e., households that belong to the same eligibility space, we find somewhat stronger evidence of local treatment composition. The likelihood to appeal, and the declines in food security and psychological well-being are larger when similar nearby households remain in the program (Table A.23 in the Appendix). The coefficient on the main discontinuation effect remains large, indicating that the adverse effects of withdrawal arise regardless of local assistance composition, but the results provides suggestive evidence that visible differences in treatment can amplify the consequences of exclusion.

[37]Although the similar midline effects attenuate concerns related to differential responses across priority and non-priority areas, one should be cautious in interpreting the difference as the causal effect of emergency cash. Priority-area status was not randomly assigned and priority areas were differentially exposed to the conflict, which might have affected some of the reactions to assistance withdrawal.

emergency, suggesting that withdrawing assistance in fragile settings exposed to recurrent shocks generates substantial and persistent welfare losses.

7.2 Transition Package Experiment

We next examine households assigned to the transition package arm, which received an additional six months of assistance before exiting the program. Figure 7 summarizes the effects on food expenditure and our main outcome indices (the underlying regression results are reported in Table A.13). Panel A compares TP households with households assigned to continued assistance, while Panel B compares them with households that exited the program under the regular protocol. We use these comparisons to address two distinct questions: whether households adjust in anticipation of program exit and whether the transition package mitigates the consequences of withdrawal.

Anticipatory Adjustment Before Exit

At midline, TP households were still receiving food assistance but had known for approximately six months that their support would soon end. The key comparison is therefore with continued-assistance households, which were receiving the same support but did not face a forthcoming exit. This comparison enables us to isolate anticipatory responses to the expected withdrawal of support from the direct effect of assistance withdrawal itself.

At midline, TP households spend 7.3 USD more on food than continued-assistance households. The estimate is economically meaningful (corresponding to 0.10 SD) but imprecisely estimated, likely reflecting the limited size of the midline TP sample. Disaggregating expenditure, Table A.8 shows that the clearest increase occurs for protein-rich foods, which were not included in the WFP basket. Beyond this suggestive expenditure response, however, we find little evidence of anticipatory effects: the estimated effects on food security, psychological well-being, and trust are all close to zero. Consistent with this pattern, Panel B shows that TP households at midline have substantially better food security, psychological well-being, and trust than regular-notice discontinued households, whose assistance had already ended.

We similarly find little evidence of anticipatory adjustment through other financial or labor-market margins (Table A.10). Point estimates suggest an 8% increase in non-food expenditure (concentrated in housing, health care, and education) and a 12% decline in the share of household members employed, relative to continued-assistance households. Neither estimate, however, is statistically significant after FDR adjustment. More disaggregated labor-market outcomes suggest that the change in employment is mostly driven by a large drop in female employment (-33%) among households that contained at least one working-age man and one working-age woman at baseline (Table A.25). Point estimates also suggest

higher job quality among employed household members.[38] We treat these disaggregated results as exploratory, however, because they were not prespecified.

Effects of the Transition Package

We next ask whether the transition package mitigated the adverse consequences of assistance withdrawal. At midline, TP households were still receiving assistance and, as shown above, had better outcomes than regular-notice households, whose support had already ended. By endline, both groups had exited the program. Differences in outcomes between TP and regular-notice households at that point therefore capture the effect of having received the transition package.

Estimates reported in Figure 7 show that, by bendline, the transition package provides little protection against the material consequences of withdrawal. TP households show expenditure patterns similar to those of regular-notice households after exit, driven by increases in spending on items previously provided through the WFP basket (Table A.8). Food security and psychological well-being also deteriorate to similar levels once assistance ends. Trust represents the only exception, as it remains closer to the continued-assistance level for TP households, although the difference from regular-notice households is also small and imprecisely estimated.

The endline comparison captures the overall effect of the transition package at a common point in time, but TP and regular-notice households had then been without assistance for different lengths of time. With the data at our disposal, we can also perform a suggestive comparison between the two groups at similar post-exit horizons: TP households at endline, approximately seven months after program exit, and regular-notice households at midline, approximately six months after program exit. The estimated effects of assistance withdrawal are strikingly similar for food expenditure (+0.17 SD vs. +0.16 SD), food security (-0.08 SD vs -0.08 SD), and psychological well-being (-0.15 SD and -0.10 SD), while trust declines substantially less among TP households (-0.03 SD vs. -0.15 SD). Because this comparison spans different periods, which include the escalation of the conflict with Israel and the introduction of emergency WFP support, it should be interpreted only as suggestive. Nevertheless, together with the lack of clear evidence of anticipatory adjustment, it reinforces the conclusion that the transition package did little to cushion the material consequences of assistance withdrawal.

[38]The job-quality index ranges from 0 to 2 among employed household members aged 15 or older, assigning one point for salaried employment and one point for working between 1 and 48 hours during the previous week, summed across all jobs.

Overall, the results suggest that a longer transition may have improved the perceived legitimacy of program exit, but did little to induce anticipatory adjustment or cushion the subsequent welfare losses from assistance withdrawal. The limited anticipatory response is consistent with liquidity constraints: even when income losses are anticipated, households may be unable to smooth consumption if they lack savings or sufficient margins of adjustment. This interpretation is supported by the fact that virtually no respondent in our survey reports any savings and that more than 90 percent of assistance recipients report consuming each monthly parcel within the month it is received. The suggestive increase in expenditure and decline in employment appear, however, more behavioral frictions (present bias, myopia, or limited attention) that may induce households to delay adjustment until withdrawal materializes.

7.3 Prioritisation Methods

We next explore whether the choice of prioritization method influences the consequences of program scale-down. Our goal is to assess whether some methods select beneficiaries in ways that more effectively limit the adverse consequences of withdrawal across the target population. Random assignment to the four prioritization arms allows us to compare the aggregate consequences of adopting each method across the beneficiary population, while also examining outcomes among "marginal" households whose continuation status depends on the method applied. We begin by studying whom the different methods select and how excluded households respond to the exclusion decision, before turning to realized welfare outcomes.

Household Selection

As we have seen above, the four prioritization methods select substantially different sets of households (Table A.2). Assessing whether one method is "better" than another in identifying the most vulnerable households within the target population, however, requires a benchmark against which the quality of the selection can be evaluated.[39] This is challenging both conceptually and empirically. Conceptually, because vulnerability is multidimensional and each prioritization method was explicitly designed to capture a different notion of it. Empirically, because the baseline census provides limited information on household vulnerability, against which the methods can be compared. The main available benchmark

[39]Appendix B provides standard assessments of targeting performance using the LHS dataset, which contains direct measures of household welfare. We evaluate each PMT-based rule out of sample against the welfare measure it is designed to predict, using standard targeting metrics, and then compare all four rules against common benchmarks based on expenditure, food consumption, and asset wealth. This validation exercise is, however, different from the analysis discussed here, which aims to compare the methods within the actual program population.

is baseline asset wealth, constructed from 36 indicators of asset ownership and connectivity using principal component analysis (PCA). Under all four methods, households with greater asset wealth are more likely to be excluded from assistance (Table A.26 and Figure A.6). This relationship is somewhat weaker under the categorical rule compared to the PMT methods, although the difference is small.

Appeal behavior provides a complementary indirect signal of targeting performance. Although the process of appealing the exclusion decision was very simple, as households could call the dedicated WFP hotline or submit the online form, it still required an active step, and only about half of the discontinued households appealed. More vulnerable households may be more likely to take this step if the value of having assistance restored is greater.[40] Consistent with this interpretation, appeals are more common among households with lower baseline asset wealth and among those closer to the 43 percent eligibility cutoff (Table A.27). At the same time, access also matters: households owning a phone or smartphone are more likely to appeal, likely because contacting WFP is less costly for them in terms of time and resources. Figure A.7 focuses on the full discontinued sample and shows modest differences in terms of appeal rates across prioritization methods, with somewhat higher rates under the categorical approach (56 percent of the discontinued households) than under the three PMT-based methods (50–52 percent).

Taken together, these results provide little evidence that any one method clearly dominates in targeting performance. The human-driven categorical approach performs relatively worse in identifying asset-based deprivation and generates more appeals against the exclusion decisions, but the differences are modest. We find virtually no differences across the three PMT-based methods, even though they target different outcomes and select distinct beneficiary profiles.

Well-being

Differences in beneficiary selection could translate into differences in overall well-being outcomes even if the methods appear similar along the dimensions considered above. Table 3 tests this directly by comparing outcomes across the four randomly assigned prioritization arms, separately at midline (Panel A) and endline (Panel B).[41] Given that the analysis includes both continued and discontinued households, these estimates capture the aggregate consequences of adopting one prioritization rule rather than another across the target population.

[40]This interpretation is consistent with evidence that even small application costs can generate substantial self-selection into transfer programs, with relatively better-off households less likely to apply (Alatas et al., 2016).

[41]For this analysis, standardized outcomes are standardized within each survey wave using the mean and standard deviation of households in the PMT-Exp arm, which serves as the reference group.

We find little evidence that any method systematically outperforms the others. Food expenditure, food security, psychological well-being, and trust are similar across prioritization arms at both survey rounds.[42] The same conclusion holds for the individual components of the main indices and for secondary outcomes related to social cohesion, employment, and household finances (Table A.28). Remarkably, we also find no advantage for a method when focusing specifically on the outcome it was designed to target, such as FCS under PMT-FCS or food expenditure under PMT-Expenditure. Population-weighted estimates yield the same conclusions (Tables A.29–A.30).

The analysis so far considers the consequences of adopting one prioritization method over another across the target population. While this is arguably the primary parameter of interest from a policy standpoint, the lack of average effect on the target population may mask differences in performance at the margin, among households whose continuation status depends on the prioritization method. We therefore repeat the comparison among marginal households in the 14 eligibility spaces where the four methods disagree. Within each eligibility space, we estimate the effect of discontinuation and then compare effects across spaces selected vs. not selected by each method. This provides a summary measure of whether a given method systematically selects households that would suffer larger losses if assistance ended. The results, detailed in Appendix D again reveal no clear ranking: no prioritization method consistently retains "marginal" households that would otherwise experience larger losses in food security, psychological well-being, or trust.[43]

Finally, we examine heterogeneity along the three prespecified dimensions: baseline wealth, household size, and the gender of the household head (Tables A.31–A.33). The results broadly reflect differences in whom the methods select. For example, female-headed households fare somewhat better under the categorical approach, which is more likely to retain them in the assistance program, while larger households fare somewhat worse under that method,

[42]Figure A.8 in the Appendix reports the full distributions of the primary outcomes by prioritization arm, separately at midline and endline, showing that the full distributions overlap closely across arms.

[43]The main exception is food expenditure: PMT-Expenditure tends to select households that would otherwise increase food expenditure more following withdrawal, while the opposite pattern appears under categorical targeting. These differences, however, do not translate into systematic differences in food security or other welfare outcomes.

which tends to select smaller households. These differences, however, are generally small and imprecisely estimated.[44]

Taken together, the results show that the differences in beneficiary selection across prioritization methods do not translate into systematic differences in realized welfare outcomes. This result is consistent with previous evidence from the targeting of new programs, which shows that differences in targeting precision do not generally translate into differences in overall program impacts (e.g. Alatas et al., 2012).

Agreement Across Methods

One possible explanation for the previous results is that targeting is inherently noisy, with performance depending on the precision of both the underlying model (especially for PMT-based approaches) and the data used to implement it. Such noise may lead to widespread targeting errors and reduce contrast across methods. If different methods provide distinct but noisy signals of vulnerability, agreement across them may, however, contain relevant information to identify households most in need of assistance. Our design allows us to test this possibility by asking whether households classified as eligible by more methods are systematically more vulnerable to assistance withdrawal.

To test this, we compare the effects of assistance discontinuation across households deemed eligible by relatively more versus fewer methods. We start by focusing on appeal behaviour among discontinued households. Figure 8 shows a clear positive relationship between the number of methods classifying a household as eligible for continued assistance and the probability of appealing when assistance is discontinued. This pattern suggests that agreement across methods might indeed identify households that are more vulnerable along dimensions not fully captured by any single rule.

Figure 9 examines households' responses to assistance withdrawal, with corresponding estimates reported in Table A.34. Households classified as eligible by three methods appear less able to replace the lost transfer through higher food expenditure than those classified as eligible by only one method: at midline, households eligible under three methods show no increase in food expenditure, whereas those eligible under only one method increase spending by 0.27 SD, or approximately USD 21.5. The gap narrows over time but remains economically

[44]We also explored heterogeneity in the relative performance of the targeting methods using the generic machine-learning framework of Chernozhukov et al. (2025), separately by primary outcome, survey wave, and pairwise comparison with PMT-Expenditure. The clearest evidence concerns food security, with some methods (PMT-FCS and Categorical) performing relatively better for specific household types (e.g. smaller households with fewer working-age members and weaker labor-market attachment). Evidence for psychological well-being and trust is instead weaker across all comparisons. Given the large number of outcome-wave-comparison combinations, however, we treat these patterns as exploratory and suggesting no large or systematic heterogeneity across observable characteristics.

meaningful at endline. A similar pattern emerges for food security: households classified as eligible by more methods experience larger declines following discontinuation, although differences are imprecisely estimated. By contrast, psychological well-being and trust display no clear ranking across eligibility groups.

The observed pattern could arise because agreement across methods captures multiple dimensions of vulnerability, or because combining distinct but noisy signals reduces classification error. Although we cannot definitively distinguish between these mechanisms, two pieces of evidence suggest that the latter may play an important role. First, PMT-CARI was itself designed to target a multidimensional measure of vulnerability, combining information on expenditure, food consumption, and coping capacity. Yet it does not systematically outperform the other prioritization rules, suggesting that multidimensionality alone is unlikely to explain the gains from agreement across methods. Second, Figure A.9 shows that households classified as eligible by more methods are increasingly concentrated at the lower end of the baseline wealth distribution, while relatively wealthy households become less common. This pattern is consistent with combining multiple signals filtering out some classification errors that arise under individual rules.

Taken together, these results suggest that, although no single prioritization method clearly dominates, agreement across methods is informative. Households identified as eligible by multiple methods are poorer in assets, more likely to contest exclusion, and less able to compensate for the loss of assistance. Combining information across targeting methods may therefore improve the identification of households least able to absorb assistance withdrawal.

8. Conclusion

In this paper, we study the scale-down process of a nationwide humanitarian assistance program and provide experimental evidence on three related dimensions: the consequences of assistance withdrawal, how households prepare for it, and whether alternative transition and prioritization strategies can mitigate any adverse effects. Although the study takes place in a specific setting, the underlying challenge is widespread across humanitarian contexts: once the acute phase of a crisis has subsided, funding constraints and shifting priorities may require programs to reduce coverage.

Our analysis shows that discontinuing assistance has substantial adverse effects on household food security and psychological well-being, despite the scale-down occurring after the initial emergency and being designed to retain the most vulnerable recipients. The estimated effects are sizable given that the transfer represented about 18 percent of median household expenditure and correspond to approximately one-third of the cross-sectional food

security gap and one-half of the psychological well-being gap between households in the top and bottom quintiles of the baseline wealth distribution, as measured among households in our sample assigned to continued assistance.[45]

Assistance withdrawal also reduced trust, particularly in aid organizations, and this effect persisted after the initial loss of support. This finding highlights a dimension of program exit that extends beyond material welfare. Humanitarian organizations operate repeatedly in the same populations, particularly in settings exposed to recurrent shocks, and their effectiveness may depend on households' support and willingness to engage with them in future emergencies.[46] The persistence of the trust effect suggests that the consequences of withdrawal may thus extend beyond the immediate loss of resources.

Providing a longer transition period did little to mitigate the material consequences of withdrawal. TP households showed little anticipatory adjustment while assistance was still being provided and, after exit, experienced losses in food security and psychological well-being comparable to those observed under the regular exit procedure. The main exception was trust, suggesting that greater predictability or a more gradual transition can improve the perceived legitimacy of exit, but that households facing chronic vulnerability may lack the resources or flexibility needed to prepare for and cushion the consequences of the loss of assistance. Future work should test transition strategies that more directly try to relax the constraints households might face, such as temporary liquidity or savings interventions, links to employment or livelihood opportunities, progressive reductions in support, and behavioral interventions that help households plan for life after assistance.

The four prioritization approaches that we considered for guiding the scaling down process, drawn from methods commonly used by aid agencies, generated little systematic variation in realized welfare across the target population, despite selecting substantially different households. Agreement across methods, however, proved to be informative. Households classified as eligible by several approaches were poorer in assets, more likely to appeal program exclusion, and less able to compensate for the loss of assistance. Rather than searching for a

[45]These effects are also roughly half the pooled average improvements in food security and psychological well-being reported by Crosta et al. (2026) in their meta-analysis of randomized evaluations of unconditional cash transfers in low- and middle-income countries. This comparison should be interpreted cautiously, because the underlying studies differ in program design, context, populations, follow-up horizon, and outcome measurement.

[46]Table A.35 provides descriptive supporting evidence from the Arab Barometer, showing that trust in the United Nations and civil-society organizations is positively associated with willingness to accept aid and with support for international organizations' involvement in socioeconomic and governance issues. The analysis is reported both for the full set of available countries and for Lebanon separately. All specifications control for age, gender, education, employment status, settlement type, household size, income hardship, and location fixed effects. Further details are provided in the table note.

universally superior targeting rule, agencies may therefore benefit from combining information across approaches and placing greater weight on cases where different measures of need point in the same direction. This insight has already informed WFP programming in Lebanon, where intersection-based targeting has now been incorporated into multiple programs.

Our study has a number of limitations. First, the discontinuation effect is identified for households whose assistance continuation status differs across prioritization methods. These households are particularly relevant to actual scale-down decisions, but the estimates do not directly identify the effects of withdrawing assistance from households classified identically (vulnerable or non-vulnerable) by all methods. Second, as with any empirical evaluation, the findings are tied to the specific population and context studied. The national spread of our sample strengthens external validity, but the study took place during a volatile period in Lebanon. In particular, the conflict with Israel and the emergency WFP response altered the assistance environment between the two survey rounds. Our analysis shows how the effects of assistance withdrawal differ depending on subsequent emergency support, but also illustrates the difficulty of evaluating program exit in settings exposed to recurrent crises. Third, and relatedly, the consequences of withdrawal may differ with the generosity and duration of the original program and the availability of alternative support tools. Evidence from other programs and contexts will therefore be important for assessing the external validity of our findings.

More broadly, our findings highlight a relatively neglected stage of the humanitarian and social-protection program lifecycle. Substantial analytical and operational attention is devoted to program entry (whom to target, what support to provide, and how to deliver it), while much less is known about when and how assistance should end. In practice, program contractions are often driven by funding constraints and competing priorities rather than by evidence that recipients can sustain their welfare without support. Our results show that these exit decisions have important consequences: withdrawal can generate persistent welfare and institutional costs, while the way households are prioritized and transitioned out of support can shape those consequences. As organizations increasingly face volatile funding and overlapping emergencies, developing and testing more deliberate exit strategies should become an integral part of program design. Future research should continue to build the evidence base on how assistance can be withdrawn in a way that best preserves program benefits.

9. References

Aguila, E., Kapteyn, A., & Perez-Arce, F. (2017). Consumption smoothing and frequency of benefit payments of cash transfer programs. *American Economic Review*, *107*(5), 430–435. https://doi.org/10.1257/aer.p20171147

Ahmed, A., Hidrobo, M., Hoddinott, J., Kolt, B., Roy, S., & Tauseef, S. (2025). Sustainable poverty reduction through social assistance: Modality, context, and complementary programming in Bangladesh. *American Economic Journal: Applied Economics*, *17*(2), 102–126. https://doi.org/10.1257/app.20230108

Aiken, E., Bellue, S., Karlan, D., Udry, C., & Blumenstock, J. E. (2022). Machine learning and phone data can improve targeting of humanitarian aid. *Nature*, *603*(7903), 864–870. https://doi.org/10.1038/s41586-022-04484-9

Aiken, E., Ashraf, A., Blumenstock, J. E., Guiteras, R. P., & Mobarak, A. M. (2025). *Scalable targeting of social protection: When do algorithms out-perform surveys and community knowledge?* (NBER Working Paper No. 33919). National Bureau of Economic Research. https://doi.org/10.3386/w33919

Aiken, E., Blumenstock, J. E., Milusheva, S., & Smith, M. M. (2026). Predicting well-being with mobile phone data: Evidence from four countries. *AEA Papers and Proceedings*, *116*, 178–183. https://doi.org/10.1257/pandp.20261092

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., Purnamasari, R., & Wai-Poi, M. (2016). Self-targeting: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, *124*(2), 371–427. https://doi.org/10.1086/685299

Alatas, V., Purnamasari, R., Wai-Poi, M., Banerjee, A., Olken, B. A., & Hanna, R. (2016). Self-targeting: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, *124*(2), 371–427. https://doi.org/10.1086/685299

Altındağ, O., & O'Connell, S. D. (2023). The short-lived effects of unconditional cash transfers to refugees. *Journal of Development Economics*, *160*, Article 102942. https://doi.org/10.1016/j.jdeveco.2022.102942

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, *103*(484), 1481–1495. https://doi.org/10.1198/016214508000000841

Angelucci, M., Chiapa, C., Prina, S., & Rojas, I. (2024). Transitory income changes and consumption smoothing: Evidence from Mexico. *Journal of Public Economics*, *230*, Article 105013. https://doi.org/10.1016/j.jpubeco.2023.105013

Baird, S., McIntosh, C., & Özler, B. (2019). When the money runs out: Do cash transfers have sustained effects on human capital accumulation? *Journal of Development Economics*, *140*, 169–185. https://doi.org/10.1016/j.jdeveco.2019.04.004

Banerjee, A., Faye, M., Krueger, A., Niehaus, P., & Suri, T. (2023). *Universal basic income: Short-term results from a long-term experiment in Kenya* [Working paper, September 15]. https://www.povertyactionlab.org/media/file-research-paper/universal-basic-income-short-term-results-long-term-experiment-kenya

Bazzi, S., Sumarto, S., & Suryahadi, A. (2015). It's all in the timing: Cash transfers and consumption smoothing in a developing country. *Journal of Economic Behavior & Organization*, *119*, 267–288. https://doi.org/10.1016/j.jebo.2015.08.010

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, *81*(2), 608–650. https://doi.org/10.1093/restud/rdt044

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bianchi, M., & Bobba, M. (2013). Liquidity, risk, and occupational choices. *The Review of Economic Studies*, *80*(2), 491–511. https://doi.org/10.1093/restud/rds031

Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, *350*(6264), 1073–1076. https://doi.org/10.1126/science.aac4420

Bossuroy, T., Goldstein, M., Karimou, B., Karlan, D., Kazianga, H., Parienté, W., Premand, P., Thomas, C. C., Udry, C., Vaillant, J., & Wright, K. A. (2022). Tackling psychosocial and capital constraints to alleviate poverty. *Nature*, *605*(7909), 291–297. https://doi.org/10.1038/s41586-022-04647-8

Bouguen, A., Huang, Y., Kremer, M., & Miguel, E. (2019). Using randomized controlled trials to estimate long-run impacts in development economics. *Annual Review of Economics*, *11*, 523–561. https://doi.org/10.1146/annurev-economics-080218-030333

Brown, C., Ravallion, M., & van de Walle, D. (2018). A poor means test? Econometric targeting in Africa. *Journal of Development Economics*, *134*, 109–124. https://doi.org/10.1016/j.jdeveco.2018.05.004

Bruni, V., & Sterck, O. (2025). *The welfare and market effects of delays in humanitarian assistance* (Working paper No. 5190878). SSRN. https://doi.org/10.2139/ssrn.5190878

Burlig, F., Jina, A., & Sudarshan, A. (2026). The value of clean water: Experimental evidence from rural India. *American Economic Review*, *116*(3), 1148–1187. https://doi.org/10.1257/aer.20250278

Chernozhukov, V., Demirer, M., Duflo, E., & Fernández-Val, I. (2025). Fisher–Schultz Lecture: Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India. *Econometrica*, *93*(4), 1121–1164. https://doi.org/10.3982/ECTA19303

Cilliers, J., Elashmawy, N., & McKenzie, D. (2024). *Using post-double selection lasso in field experiments* (Policy Research Working Paper No. 10931). World Bank. https://documents1.worldbank.org/curated/en/099721209262431433/pdf/IDU-9a34656a-9be7-4ca2-ac1f-c8c30a8b071d.pdf

Crosta, T., Karlan, D., Ong, F., Rüschenpöhler, J., & Udry, C. R. (2026). *Unconditional cash transfers: A Bayesian meta-analysis of randomized evaluations in low and middle income countries* (NBER Working Paper No. 32779). National Bureau of Economic Research. https://doi.org/10.3386/w32779

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica*, *59*(5), 1221–1248. https://doi.org/10.2307/2938366

Deserranno, E., Guariso, A., & Qian, N. (2026). *Delivering aid at the last mile*. Working paper prepared for the *Annual Review of Economics*. https://sites.google.com/site/erikadeserranno/research

Deville, J.-C., & Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, *87*(418), 376–382. https://doi.org/10.1080/01621459.1992.10475217

Duflo, E., Kremer, M., & Robinson, J. (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. *American Economic Review*, *101*(6), 2350–2390. https://doi.org/10.1257/aer.101.6.2350

Edmonds, E. V. (2006). Child labor and schooling responses to anticipated income in South Africa. *Journal of Development Economics*, *81*(2), 386–414. https://doi.org/10.1016/j.jdeveco.2005.05.001

Fink, G., & Redaelli, S. (2011). Determinants of international emergency aid—Humanitarian need only? *World Development*, *39*(5), 741–757. https://doi.org/10.1016/j.worlddev.2010.09.004

Fishman, R., Smith, S. C., Bobić, V., & Sulaiman, M. (2022). Can agricultural extension and input support be discontinued? Evidence from a randomized phaseout in Uganda. *The Review of Economics and Statistics*, *104*(6), 1273–1288. https://doi.org/10.1162/rest_a_01029

Friedman, M. (1957). The permanent income hypothesis. In *A theory of the consumption function* (pp. 20–37). Princeton University Press. https://www.nber.org/books-and-chapters/theory-consumption-function/permanent-income-hypothesis

Ganong, P., & Noel, P. (2019). Consumer spending during unemployment: Positive and normative implications. *American Economic Review*, *109*(7), 2383–2424. https://doi.org/10.1257/aer.20170537

Gentilini, U. (2016). Revisiting the "cash versus food" debate: New evidence for an old puzzle? *The World Bank Research Observer*, *31*(1), 135–167. https://doi.org/10.1093/wbro/lkv012

Gerard, F., & Naritomi, J. (2021). Job displacement insurance and (the lack of) consumption-smoothing. *American Economic Review*, *111*(3), 899–942. https://doi.org/10.1257/aer.20190388

Grisolia, F. (2024). *Can cash transfers really be transformative? A literature review of the sustainability of their impacts* (IOB Discussion Paper No. 2024.02). University of Antwerp, Institute of Development Policy. https://medialibrary.uantwerpen.be/files/2137/36430341-e22a-4e66-bb6b-ab280aed867b.pdf

Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy*, *86*(6), 971–987. https://doi.org/10.1086/260724

Handa, S., Tembo, G., Natali, L., & Chakrabarti, A. (2025). In search of the holy grail: Post-intervention effects of an unconditional cash transfer program in Zambia. *Journal of Development Economics*, *174*, Article 103454. https://doi.org/10.1016/j.jdeveco.2025.103454

Haushofer, J., Niehaus, P., Paramo, C., Miguel, E., & Walker, M. (2025). Targeting impact versus deprivation. *American Economic Review*, *115*(6), 1936–1974. https://doi.org/10.1257/aer.20221650

Henderson, H., & Follett, L. (2022). Targeting social safety net programs on human capabilities. *World Development*, *151*, Article 105741. https://doi.org/10.1016/j.worlddev.2021.105741

Hidrobo, M., Hoddinott, J., Peterman, A., Margolies, A., & Moreira, V. (2014). Cash, food, or vouchers? Evidence from a randomized experiment in northern Ecuador. *Journal of Development Economics*, *107*, 144–156. https://doi.org/10.1016/j.jdeveco.2013.11.009

Hidrobo, M., Karachiwalla, N., & Roy, S. (2023). The impacts of cash transfers on mental health and investments: Experimental evidence from Mali. *Journal of Economic Behavior & Organization*, *216*, 608–630. https://doi.org/10.1016/j.jebo.2023.10.016

Hoddinott, J., Sandström, S., & Upton, J. (2018). The impact of cash and food transfers: Evidence from a randomized intervention in Niger. *American Journal of Agricultural Economics*, *100*(4), 1032–1049. https://doi.org/10.1093/ajae/aay019

Humphreys, M. (2025). Bounds on the fixed effects estimand in the presence of heterogeneous assignment propensities. *Journal of Causal Inference*, *13*(1), Article 20240040. https://doi.org/10.1515/jci-2024-0040

International Organization for Migration (IOM). (2024). *Lebanon Emergency Response: Activity Update #5*, 29 November–12 December 2024. Beirut: IOM.

Iqbal, N., Jalal, A., Mahmud, M., & Vyborny, K. (2026). What happens when women suddenly stop receiving cash transfers? *Journal of Public Economics*, *257*, Article 105618. https://doi.org/10.1016/j.jpubeco.2026.105618

Jeong, D., & Trako, I. (2022). *Cash and in-kind transfers in humanitarian settings: A review of evidence and knowledge gaps* (Policy Research Working Paper No. 10026). World Bank. https://documents1.worldbank.org/curated/en/099445304272240978/pdf/IDU01240088a0d11904ded09c4b0aa5fc374c8d8.pdf

Karlan, D., McConnell, M., Mullainathan, S., & Zinman, J. (2016). Getting to the top of mind: How reminders increase saving. *Management Science*, *62*(12), 3393–3411. https://doi.org/10.1287/mnsc.2015.2296

Knickmeyer, E., Sherman, M., & Lee, M. (2025, February 27). Trump administration says it's cutting 90% of USAID foreign aid contracts. *Associated Press*. Republished by Boston.com. https://www.boston.com/news/national-news/2025/02/27/trump-administration-says-its-cutting-90-of-usaid-foreign-aid-contracts/

Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, *121*(4), 1133–1165. https://doi.org/10.1093/qje/121.4.1133

Kőszegi, B., & Rabin, M. (2009). Reference-dependent consumption plans. *American Economic Review*, *99*(3), 909–936. https://doi.org/10.1257/aer.99.3.909

Ligon, E., & Schechter, L. (2003). Measuring vulnerability. *The Economic Journal*, *113*(486), C95–C102. https://doi.org/10.1111/1468-0297.00117

Lin, W. (2013). Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique. *The Annals of Applied Statistics*, *7*(1), 295–318. https://doi.org/10.1214/12-AOAS583

Manacorda, M., Miguel, E., & Vigorito, A. (2011). Government transfers and political support. *American Economic Journal: Applied Economics*, *3*(3), 1–28. https://doi.org/10.1257/app.3.3.1

Márquez-Padilla, F., Parker, S. W., & Vogl, T. S. (2025). *Rolling back Progresa: School and work after the end of a landmark anti-poverty program* (NBER Working Paper No. 33527). National Bureau of Economic Research. https://doi.org/10.3386/w33527

Maxwell, D., & Caldwell, R. (2008). *The coping strategies index: Field methods manual* (2nd ed.). CARE.

McBride, L., & Nichols, A. (2018). Retooling poverty targeting using out-of-sample validation and machine learning. *The World Bank Economic Review*, *32*(3), 531–550. https://doi.org/10.1093/wber/lhw056

Molina Millán, T., Barham, T., Macours, K., Maluccio, J. A., & Stampini, M. (2019). Long-term impacts of conditional cash transfers: Review of the evidence. *The World Bank Research Observer*, *34*(1), 119–159. https://doi.org/10.1093/wbro/lky005

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, *89*(1), 103–124. https://doi.org/10.1257/aer.89.1.103

Pople, A., Hill, R. V., Dercon, S., & Brunckhorst, B. (2026). The importance of being early? Anticipatory cash transfers for flood-affected households. *The World Bank Economic Review*. Advance online publication. https://doi.org/10.1093/wber/lhag016

Premand, P., & Schnitzer, P. (2021). Efficiency, legitimacy, and impacts of targeting methods: Evidence from an experiment in Niger. *The World Bank Economic Review*, *35*(4), 892–920. https://doi.org/10.1093/wber/lhaa019

Reis, R. (2006). Inattentive consumers. *Journal of Monetary Economics*, *53*(8), 1761–1800. https://doi.org/10.1016/j.jmoneco.2006.03.001

Sterck, O., & Bruni, V. (2025). *The welfare and general-equilibrium impacts of aid cuts* (Working paper No. 5270846). SSRN. https://doi.org/10.2139/ssrn.5270846

Stewart, H. (2025, May 17). Diminished UK aid budget is 'new normal', says development minister. *The Guardian*. https://www.theguardian.com/politics/2025/may/17/diminished-uk-aid-budget-is-new-normal-says-development-minister

Sumarto, S., Satriawan, E., Olken, B. A., Banerjee, A., Tohari, A., Alatas, V., & Hanna, R. (2025). *Community targeting at scale* (NBER Working Paper No. 33322). National Bureau of Economic Research. https://doi.org/10.3386/w33322

United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA). (2026). *Financial Tracking Service: 2024 plan-level funding data* [Data set]. Financial Tracking Service. Retrieved July 16, 2026, from https://api.hpc.tools/v1/public/fts/flow?year=2024&groupby=Plan

World Bank. (2022). *Lebanon economic monitor, fall 2022: Time for an equitable banking resolution*. https://www.worldbank.org/en/country/lebanon/publication/lebanon-economic-monitor-fall-2022-time-for-an-equitable-banking-resolution

World Bank. (2026). *GDP per capita (current US$)—Lebanon* [Data set]. World Development Indicators. Retrieved July 14, 2026, from https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=LB

World Food Programme (WFP). (2021). *Technical guidance for WFP consolidated approach for reporting indicators of food security (CARI)* (3rd ed.). https://wfp.tind.io/record/130402

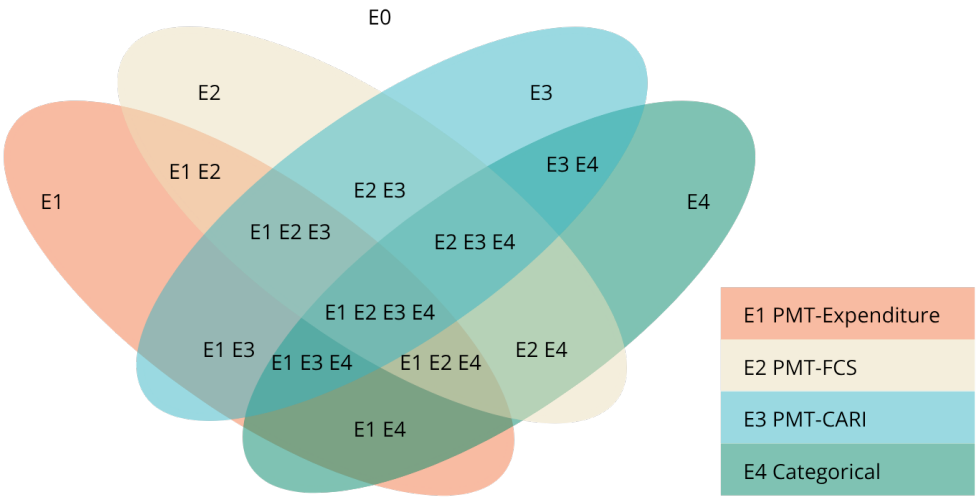
World Food Programme (WFP). (2022, July). *Food security and vulnerability analysis of Lebanese residents*. https://www.3rpsyriacrisis.org/wp-content/uploads/2022/10/WFP-0000142737.pdf

World Food Programme (WFP). (2026, May 28). *Food consumption score*. VAM Resource Centre. https://vamresources.manuals.wfp.org/docs/food-consumption-score

World Health Organization (WHO). (2024). *The World Health Organization-Five Well-Being Index (WHO-5)* (Publication No. WHO-UCN-MSD-MHE-2024.01). https://www.who.int/publications/m/item/WHO-UCN-MSD-MHE-2024.01

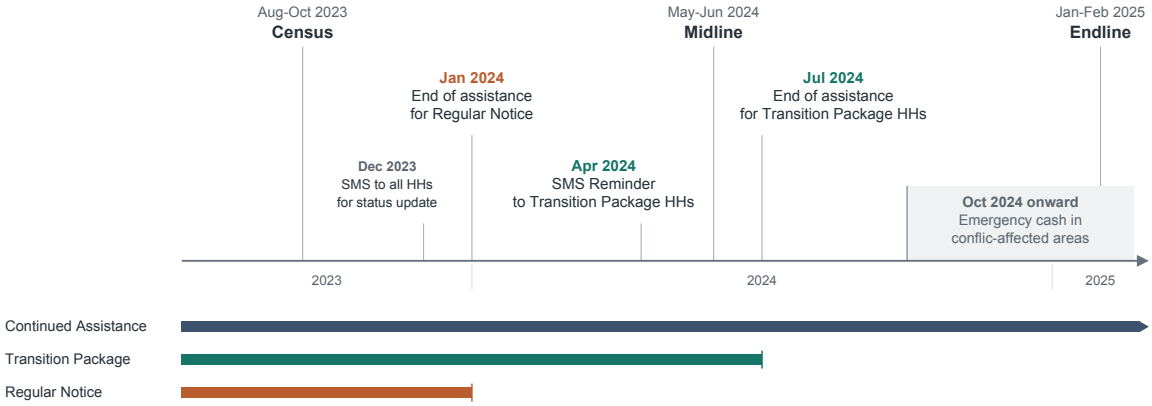
Figures and Tables

Figure 1. Eligibility Spaces.



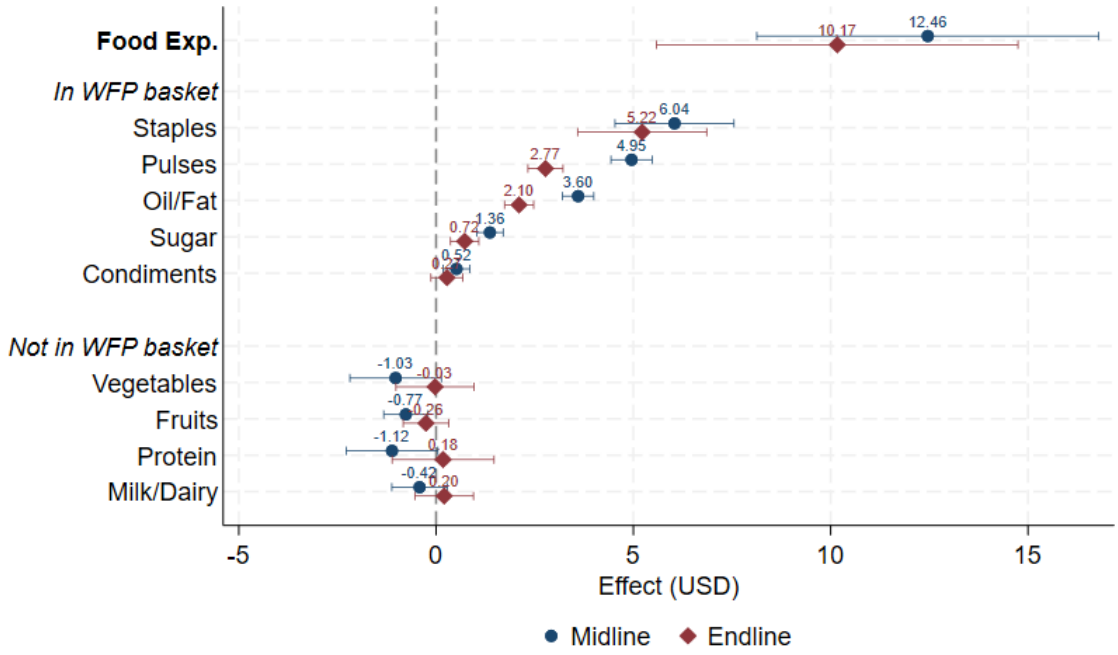
Notes: The figure illustrates the 16 "eligibility spaces" generated by applying all four prioritization methods to the 2023 beneficiary population. Each space groups households that share the same eligibility pattern across the four methods, meaning that they would be selected for continued assistance under the same method or combination of methods. Method 1 is PMT-Exp, Method 2 is PMT-FCS, Method 3 is PMT-CARI, and Method 4 is Categorical. E0 denotes the set of households that would not be selected by any method, while Ei indicates that households in that group would be selected for continued assistance under method i.

Figure 2. Timeline.



Notes: The figure summarizes the sequence of events. In December 2023, all households received an SMS informing them of their assistance status. Continued-assistance households were informed that they would continue receiving assistance throughout the next program cycle (at least until 2025). Regular-notice discontinued assistance households were informed that, due to funding constraints, their assistance would end that month. Households assigned to the transition package were informed that, due to funding constraints, their assistance would continue only through June 2024. This latter group received a reminder in April 2024 about the forthcoming end of assistance.

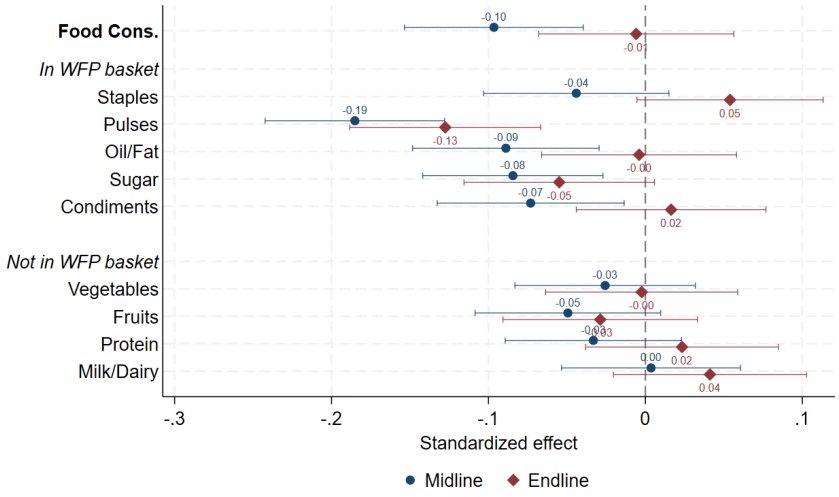
Figure 3. Impact of Assistance Discontinuation on Food Expenditure.



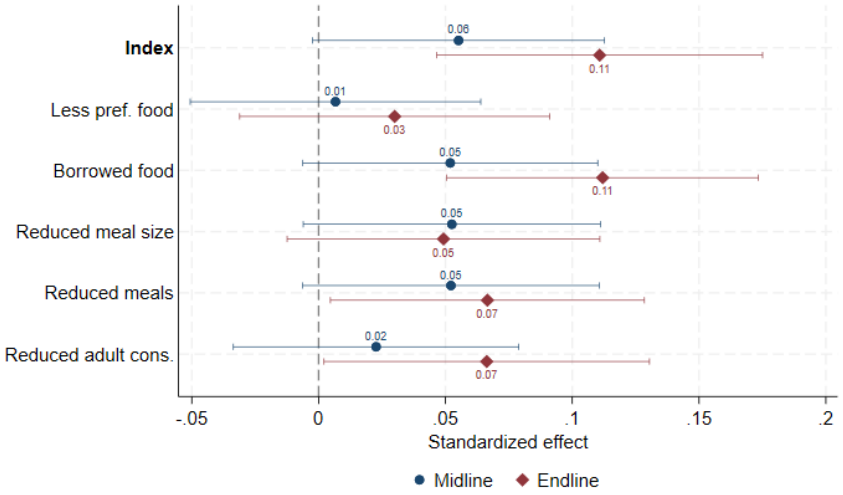
Notes: The figure plots the estimated impact of assistance discontinuation on total food expenditure and expenditure across food categories. The figure distinguishes between food categories included in the WFP assistance basket and other food categories. Midline (blue) and endline (red) estimates are reported separately. Estimates are obtained from regressions that include eligibility-space fixed effects and a control for random assignment to the transition package. Outcomes are measured in USD and include food purchased on credit as well as the estimated value of home production. Lines indicate 95 percent confidence intervals. The underlying regression results are reported in Table A.8 in the Appendix.

Figure 4. Impact of Assistance Discontinuation on Food Security.

Panel A: Food Consumption

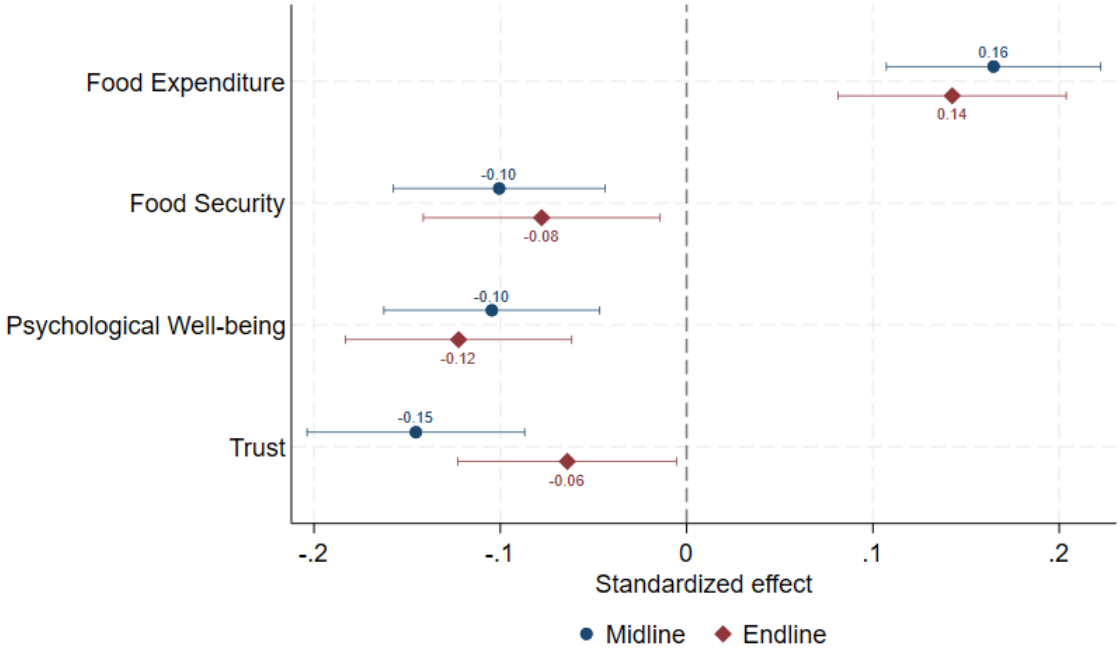


Panel B: Risky Coping Strategies



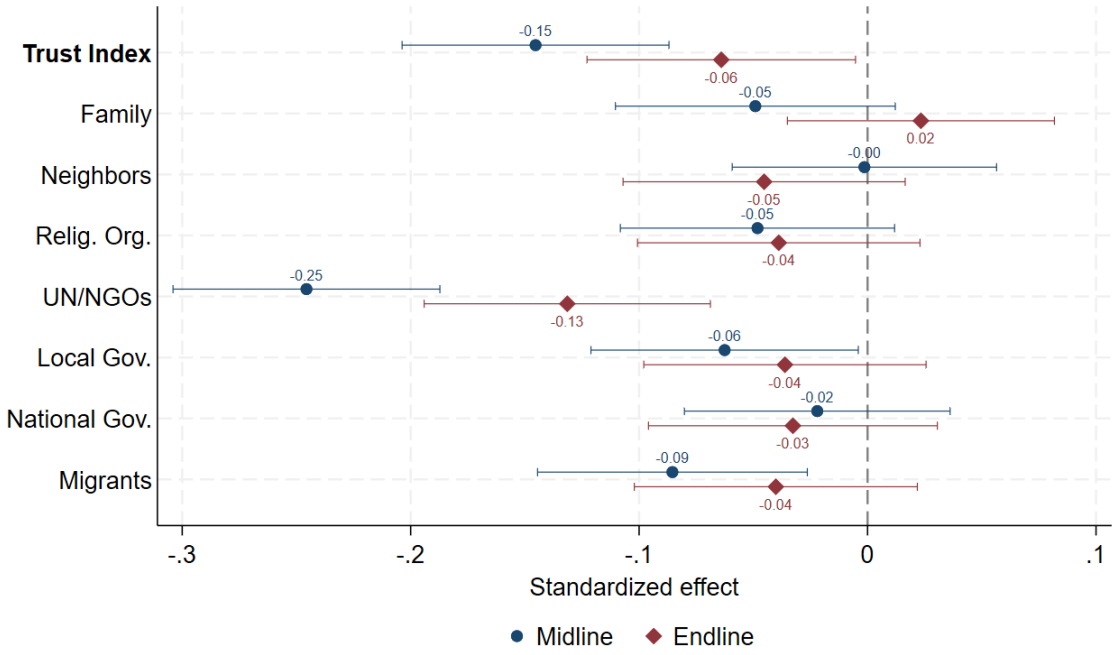
Notes: Panel A plots the estimated impact of assistance discontinuation on the Food Consumption Score (FCS) and its components. The figure distinguishes between food categories included in the WFP assistance basket and other food categories. Panel B plots the estimated impact of assistance discontinuation on the reduced Coping Strategies Index (rCSI) and its components. For the rCSI outcomes, positive values indicate more frequent use of coping strategies. Midline (blue) and endline (red) estimates are reported separately. Estimates are obtained from regressions that include eligibility-space fixed effects and a control for random assignment to the transition package. Outcomes in both panels are standardized within each survey wave using the mean and standard deviation of continued-assistance households. Lines indicate 95 percent confidence intervals. Full regression results (in the original outcome units) are reported in Tables A.11 and A.12 in the Appendix.

Figure 5. Impact of Assistance Discontinuation on Main Outcomes.



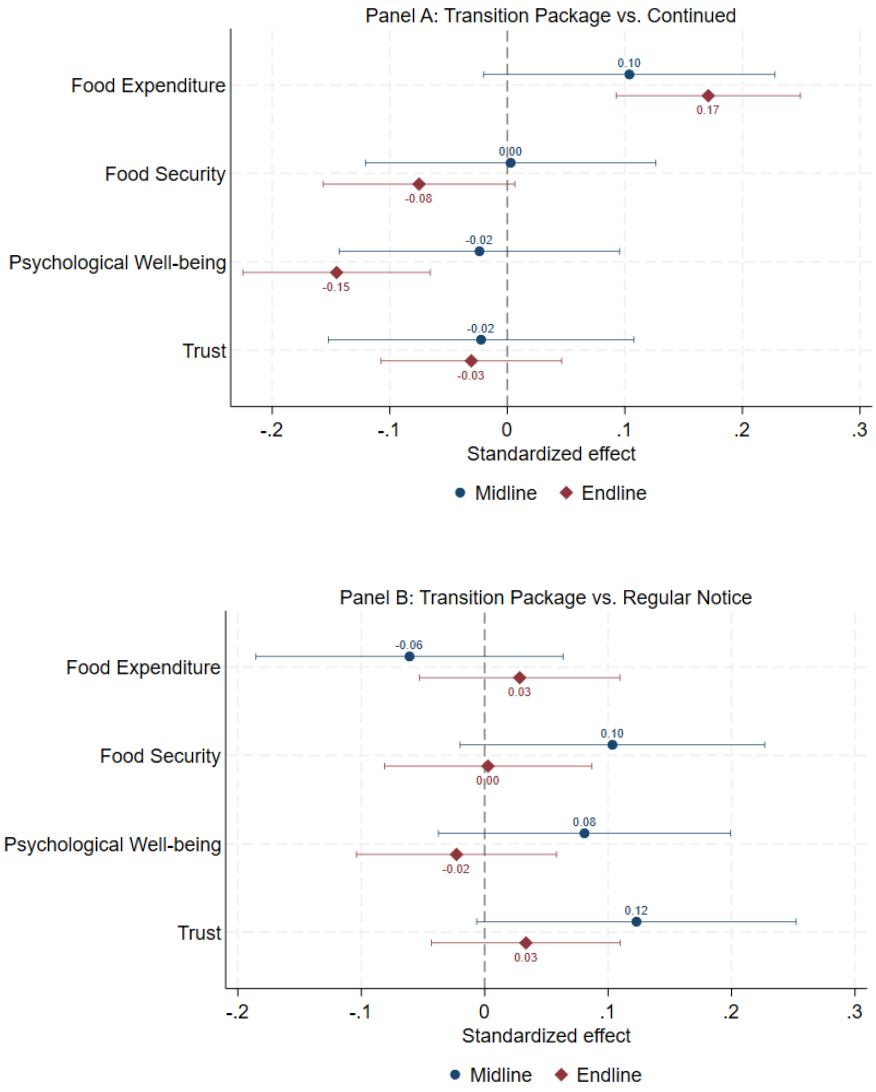
Notes: The figure plots the estimated impact of assistance discontinuation on food expenditure, food security, psychological well-being, and trust. Food expenditure includes food purchased on credit as well as the estimated value of home production. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, reverse-coded worriedness, and mental health. The trust index combines seven measures of trust in different groups and institutions. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights. All outcomes, including food expenditure, are standardized within each survey wave using the mean and standard deviation of continued-assistance households. Midline (blue) and endline (red) estimates are reported separately. Estimates are obtained from regressions that include eligibility-space fixed effects and a control for random assignment to the transition package. Lines indicate 95 percent confidence intervals. The underlying regression results are reported in Table A.13 in the Appendix.

Figure 6. Impact of Assistance Discontinuation on Trust.



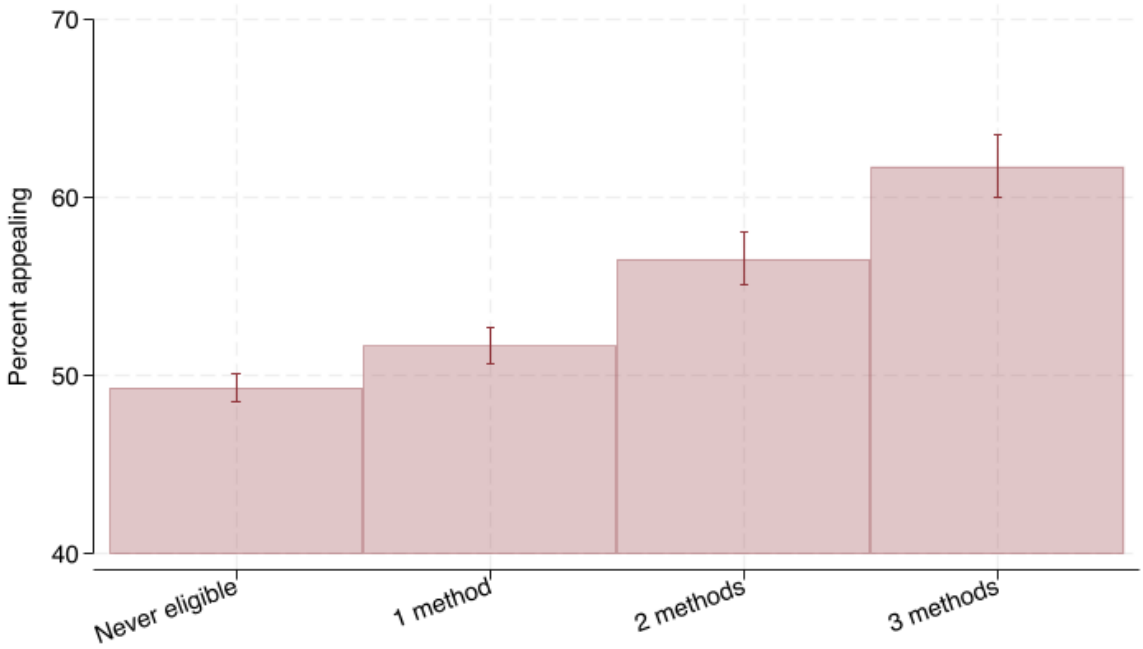
Notes: The figure plots the estimated impact of assistance discontinuation on the trust index and its seven components. The index combines the seven components using inverse-covariance weights. Higher values indicate greater trust.Outcomes are standardized within each survey wave using the mean and standard deviation of continued-assistance households. Midline (blue) and endline (red) estimates are reported separately. Estimates are obtained from regressions that include eligibility-space fixed effects and a control for random assignment to the transition package. Lines indicate 95 percent confidence intervals. The underlying regression results are reported in Table A.14 in the Appendix.

Figure 7. The Impact of the Transition Package.



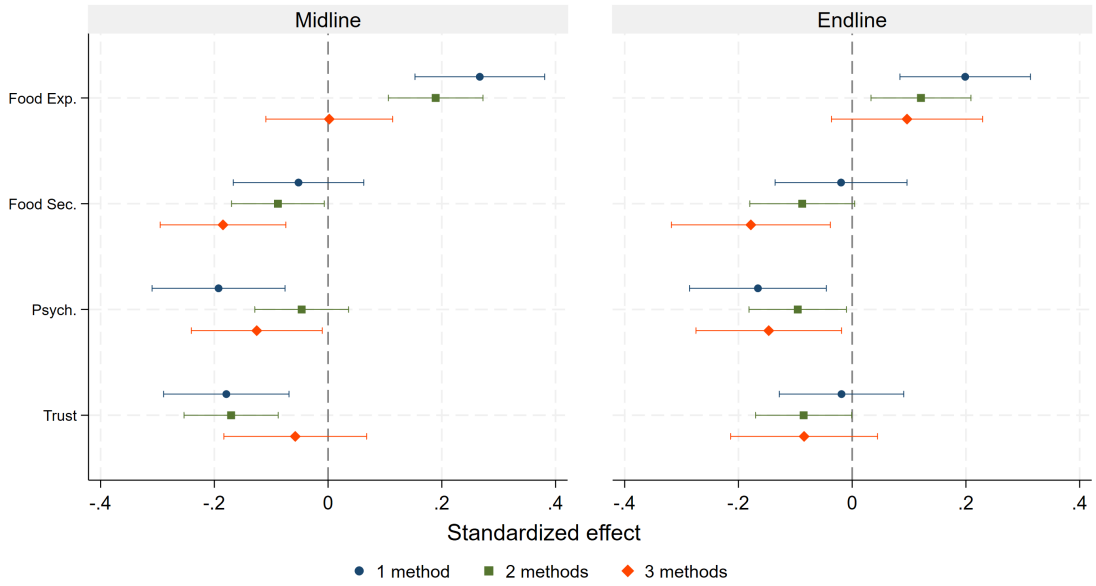
Notes: Panel A plots the estimated impact of assignment to the transition package relative to continued assistance. Panel B plots the estimated impact of assignment to the transition package relative to regular-notice discontinuation. The transition package provided households with six additional months of assistance and a longer notice period than regular-notice households. Outcomes are standardized within each survey wave using the mean and standard deviation of continued-assistance households. Midline (blue) and endline (red) estimates are reported separately. Estimates are obtained from regressions that include eligibility-space fixed effects. Lines indicate 95 percent confidence intervals. The underlying regression results are reported in Table A.13 in the Appendix.

Figure 8. Households' Appeals by Eligibility.



Notes: The figure reports appeal submission rates by the number of prioritization methods under which a household would have been eligible for continued assistance. Households are grouped according to whether they were eligible under zero, one, two, or three of the four methods. Households eligible under all four methods are omitted because they all continued receiving assistance and submitted no appeals. The sample includes all households whose assistance was discontinued. Bars show group means and lines indicate 95 percent confidence intervals.

Figure 9. Heterogeneity by Eligibility.



Notes: The figure plots the estimated impact of assistance discontinuation on the main outcomes separately for households that were eligible for continued assistance under one (blue circles), two (green squares), or three (orange diamonds) prioritization methods. Midline and endline estimates are presented in separate panels. Outcomes are standardized within each survey wave using the mean and standard deviation of continued-assistance households. Estimates are obtained from regressions that include eligibility-space fixed effects, a control for random assignment to the transition package, and its interactions with indicators for eligibility under two or three prioritization methods. Lines indicate 95 percent confidence intervals. The underlying regression results are reported in Table A.34 in the Appendix.

Table 1. Balance Checks across Prioritization Arms.

	PMT-Exp	PMT-FCS	PMT-CARI	Categorical	p-value
	(1)	(2)	(3)	(4)	(5)
Household size	3.81	3.83	3.77	3.81	0.738
	(1.79)	(1.81)	(1.80)	(1.82)	
School-age children	0.88	0.87	0.86	0.88	0.842
	(1.18)	(1.21)	(1.16)	(1.20)	
Female household head	0.34	0.35	0.31	0.32	0.046
	(0.47)	(0.48)	(0.46)	(0.47)	
Senior household head (65+)	0.26	0.27	0.27	0.27	0.924
	(0.44)	(0.44)	(0.44)	(0.45)	
Household head with primary education or lower	0.60	0.61	0.60	0.59	0.557
	(0.49)	(0.49)	(0.49)	(0.49)	
Any household member has chronic illness	0.74	0.74	0.73	0.75	0.866
	(0.44)	(0.44)	(0.44)	(0.44)	
Household head suffers from any disability	0.21	0.19	0.21	0.20	0.292
	(0.40)	(0.39)	(0.41)	(0.40)	
No household member employed	0.33	0.34	0.34	0.34	0.657
	(0.47)	(0.47)	(0.47)	(0.47)	
Dependency ratio	0.35	0.34	0.35	0.35	0.724
	(0.30)	(0.30)	(0.31)	(0.30)	
Household is located in Beirut	0.06	0.07	0.05	0.05	0.307
	(0.24)	(0.25)	(0.22)	(0.22)	
Household owns a smartphone	0.80	0.81	0.80	0.82	0.352
	(0.40)	(0.40)	(0.40)	(0.39)	
Household owns a car	0.28	0.29	0.30	0.31	0.181
	(0.45)	(0.46)	(0.46)	(0.46)	
Piped water	0.23	0.25	0.25	0.24	0.361
	(0.42)	(0.43)	(0.43)	(0.43)	
Gas for cooking	0.75	0.73	0.72	0.72	0.126
	(0.43)	(0.45)	(0.45)	(0.45)	
WiFi internet	0.47	0.48	0.46	0.48	0.645
	(0.50)	(0.50)	(0.50)	(0.50)	
Wealth index (PCA)	-0.20	-0.17	-0.16	-0.17	0.488
	(0.89)	(0.87)	(0.87)	(0.86)	
PMT-Expenditure score	15.19	15.19	15.19	15.19	0.879
	(0.38)	(0.37)	(0.39)	(0.38)	
PMT-FCS score	54.26	54.52	54.42	54.13	0.089
	(7.87)	(7.65)	(8.05)	(7.81)	
PMT-CARI score	0.40	0.40	0.40	0.40	0.654
	(0.13)	(0.14)	(0.14)	(0.13)	
Categorical score	39.54	39.52	39.21	39.19	0.965
	(43.74)	(43.54)	(44.55)	(43.84)	
Observations	1,600	1,600	1,600	1,600	6,400

Notes: The table reports baseline means for the original study sample, by prioritization arm. Standard deviations are reported in parentheses. The final column reports the p-value from a test of equality of means across all groups ($H_0 :=$ means are equal across groups). The table presents only a subset of the available variables. When considering the full set of baseline variables and targeting scores, there are 160 possible balance tests. Of these, 18 (11.3%) are significant at the 10 percent level ($p < 0.10$), 12 (7.5%) at the 5 percent level ($p < 0.05$), and 4 (2.5%) at the 1 percent level ($p < 0.01$).

Table 2. Attrition.

Panel A: Attrition by Targeting Arm

	PMT-Exp	PMT-FCS	PMT-CARI	Categorical	p-value
	(1)	(2)	(3)	(4)	(5)
Full sample					
Target sample	1,600	1,600	1,600	1,600	–
Attrition at Midline	172 (10.8%)	179 (11.2%)	161 (10.1%)	176 (11.0%)	0.721
Attrition at Endline	104 (6.5%)	96 (6.0%)	104 (6.5%)	94 (5.9%)	0.827
Sample excluding Transition Package					
Target sample	1,365	1,361	1,369	1,359	–
Attrition at Midline	4 (0.3%)	10 (0.7%)	14 (1.0%)	7 (0.5%)	0.055
Attrition at Endline	83 (6.1%)	84 (6.2%)	85 (6.2%)	80 (5.9%)	0.947

Panel B: Attrition by assistance status

	Continued	Regular Notice	Transition Package	p-value (1) vs (2)	p-value (1) vs (2) vs (3)
	(1)	(2)	(3)	(4)	(5)
Target sample	3,200	2,254	946	–	–
Attrition at Midline	23 (0.7%)	12 (0.5%)	653 (69.0%)	0.384	0.000
Attrition at Endline	197 (6.2%)	135 (6.0%)	66 (7.0%)	0.799	0.587

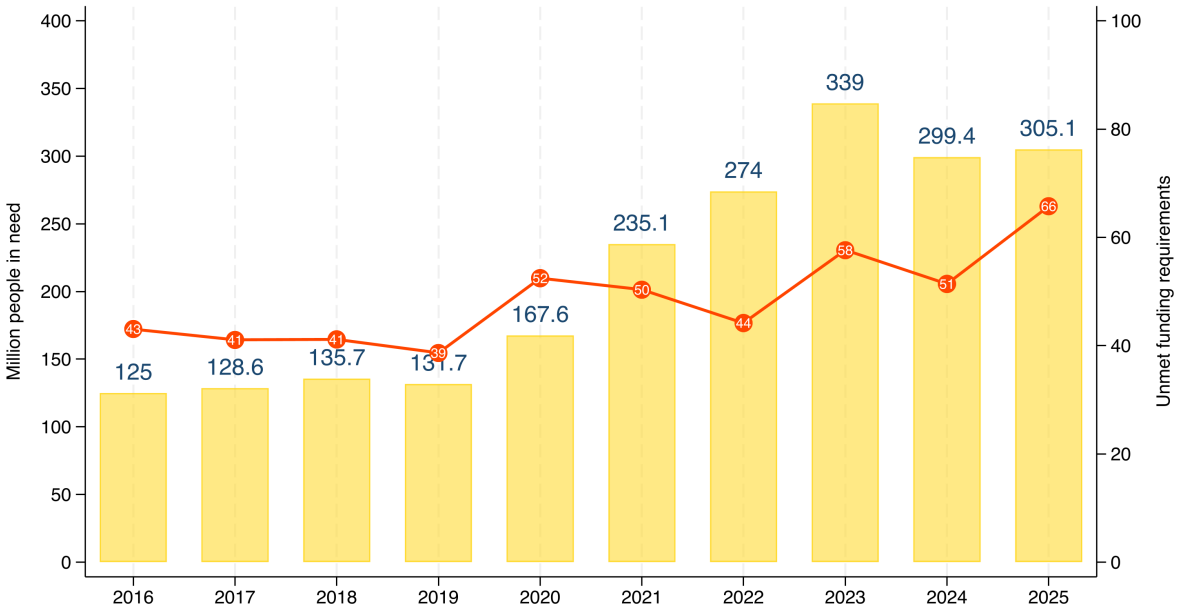
Notes: Cells report counts, with percentages in parentheses. All percentages are calculated relative to the original target sample. The final column reports the p-value from a test of equality of attrition rates across the displayed groups.

Table 3. Impact of Prioritization Methods on Main Outcomes.

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
PMT-FCS	0.08**	0.04	0.01	-0.03
	(0.03)	(0.03)	(0.04)	(0.04)
	[0.059]	[0.586]	[0.767]	[0.586]
PMT-CARI	0.03	0.03	0.04	-0.04
	(0.03)	(0.04)	(0.03)	(0.04)
	[0.466]	[0.466]	[0.466]	[0.466]
Categorical	0.06*	-0.00	0.00	-0.02
	(0.03)	(0.04)	(0.03)	(0.04)
	[0.397]	[0.972]	[0.972]	[0.972]
q-val(PMT-FCS=PMT-CARI)	0.405	0.789	0.789	0.789
q-val(PMT-FCS=Categorical)	0.830	0.830	0.842	0.842
q-val(PMT-CARI=Categorical)	0.535	0.535	0.535	0.565
Mean PMT-Exp.	0.00	0.00	0.00	0.00
Observations	6,358	6,358	6,358	6,356
Panel B: Endline				
PMT-FCS	0.04	0.01	-0.03	-0.04
	(0.03)	(0.04)	(0.04)	(0.03)
	[0.519]	[0.852]	[0.519]	[0.519]
PMT-CARI	-0.00	0.02	0.06	0.07**
	(0.03)	(0.04)	(0.03)	(0.03)
	[0.906]	[0.734]	[0.203]	[0.150]
Categorical	0.06*	-0.02	0.01	-0.01
	(0.03)	(0.04)	(0.03)	(0.04)
	[0.347]	[0.854]	[0.854]	[0.854]
q-val(PMT-FCS=PMT-CARI)	0.311	0.680	0.025	0.003
q-val(PMT-FCS=Categorical)	0.521	0.517	0.517	0.517
q-val(PMT-CARI=Categorical)	0.133	0.210	0.210	0.101
Mean PMT-Exp.	0.00	0.00	0.00	0.00
Observations	6,465	6,465	6,465	6,465

Notes: PMT-Expenditure is the omitted (reference) group. Food expenditure includes food purchased on credit as well as the estimated value of home production. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, reverse-coded worriedness, and mental health. The trust index combines seven measures of trust in different groups and institutions. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights. Outcomes are standardized within each survey wave using the mean and standard deviation of households assigned to the PMT-Expenditure arm. The q-value rows report FDR-adjusted q-values from tests of the null hypothesis that the coefficients on the indicated prioritization arms are equal. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

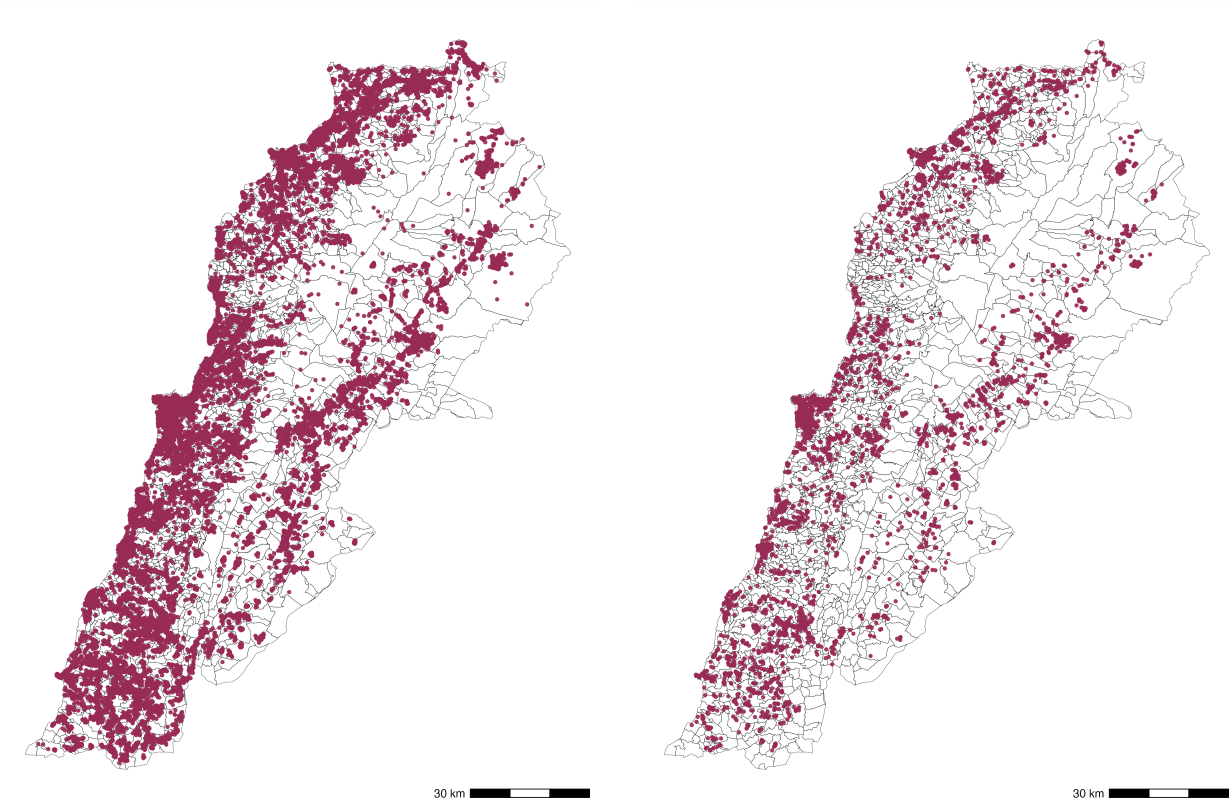
Figure A.1. Humanitarian Needs and Unmet Funding Requirements (2016–2025).



Notes: Bars report the estimated number of people in need, as reported in the annual Global Humanitarian Overview, measured in millions (left axis). The line reports the share of funding requirements that remained unmet (right axis), as recorded by the UN OCHA Financial Tracking Service. Line labels are rounded to the nearest percentage point.

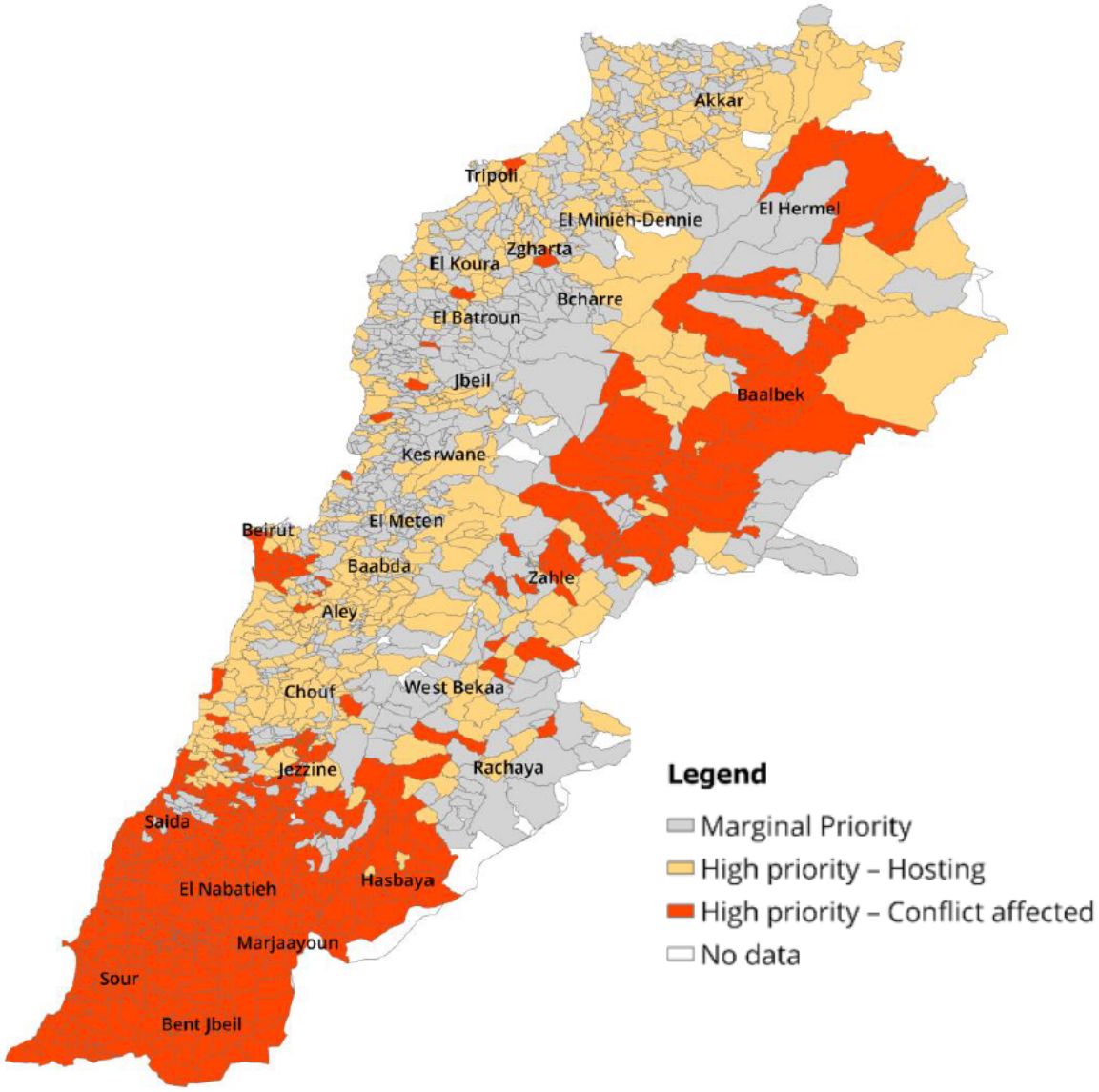
Figure A.2. Population and Sample Location.

Panel A: Beneficiary Population **Panel B: Study Sample**



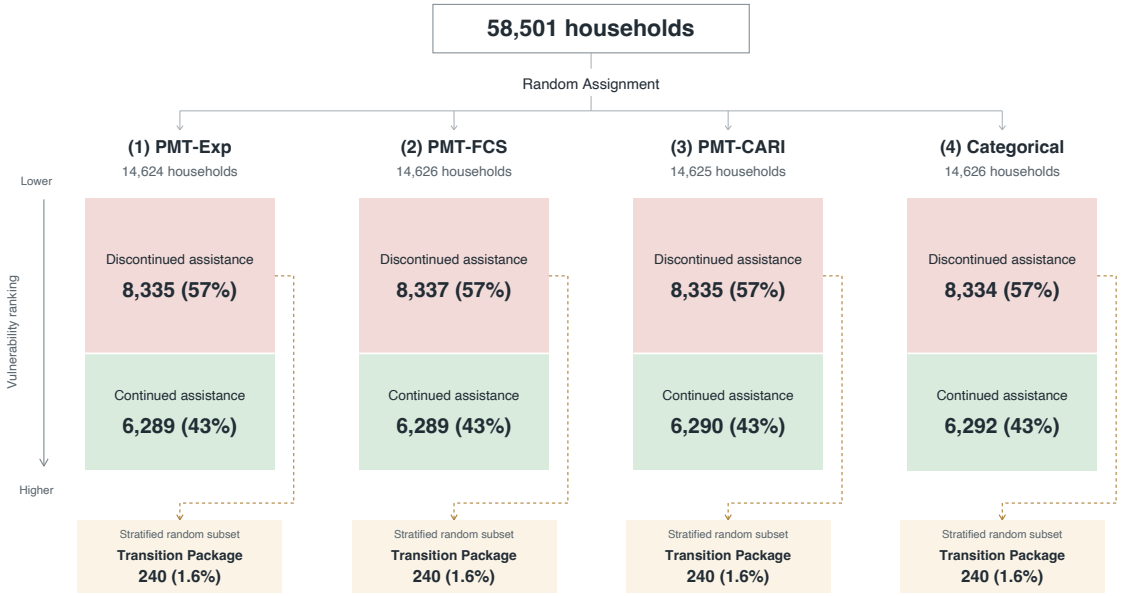
Notes: Panel A shows the geographic distribution of households in the full 2023 beneficiary population. Panel B shows the geographic distribution of the subset of households included in the study sample. Approximately 14 percent of households are omitted because valid GPS coordinates are unavailable, although their cadastre of residence is known in all cases.

Figure A.3. Map of Priority Areas due to Conflict.



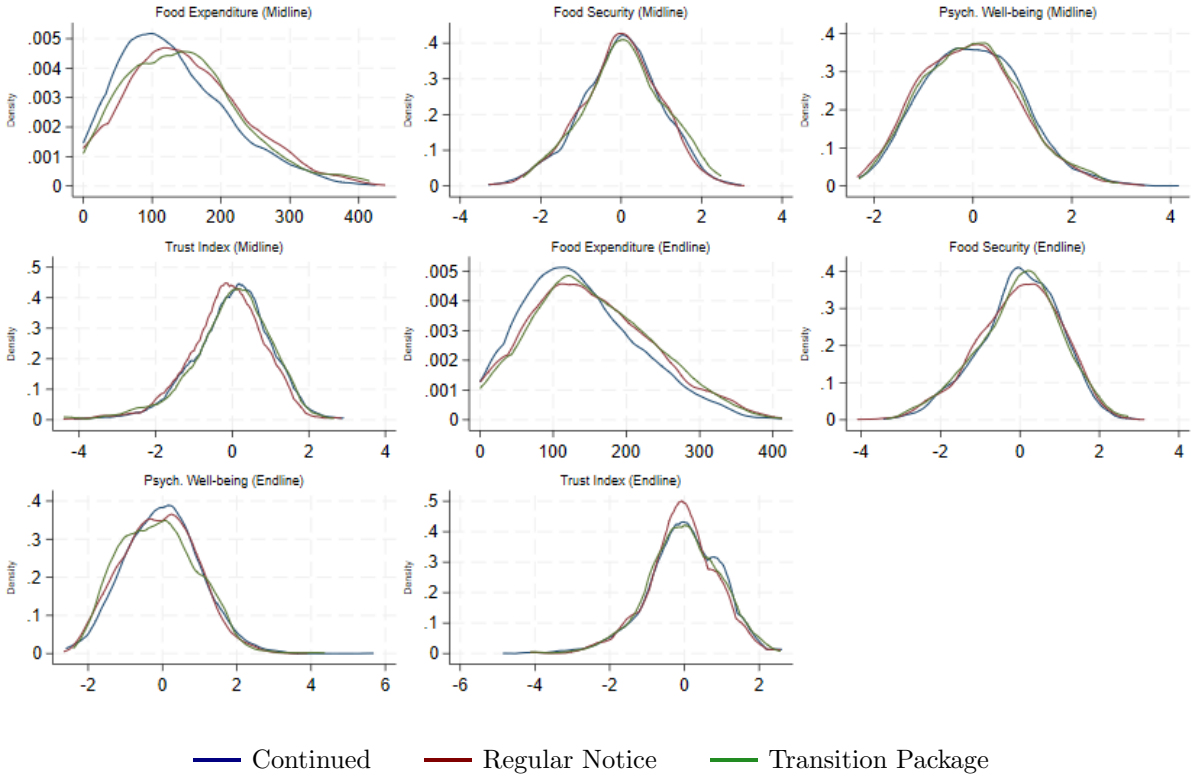
Notes: The figure was produced and kindly shared by the WFP Lebanon Country Office. Areas shown in red ("High priority - Conflict affected") are cadastres with a value of 3 or higher on the WFP Conflict Impact Index. The index takes into account the number of strikes and fatalities, whether civilians were targeted, whether residents had access to markets, whether a majority of the population had been displaced, and whether an evacuation order had been issued. Areas shown in orange ("High priority - Hosting") are cadastres that did not meet the conflict-affected criterion but hosted a large number of internally displaced households (either more than 500 households or more than 20% of the cadastre's population). Starting in October 2024, households located in either type of high-priority area received WFP emergency cash assistance, provided that they were already registered in the WFP system and were not receiving support through another WFP program.

Figure A.4. Random Assignment and Treatment Status.



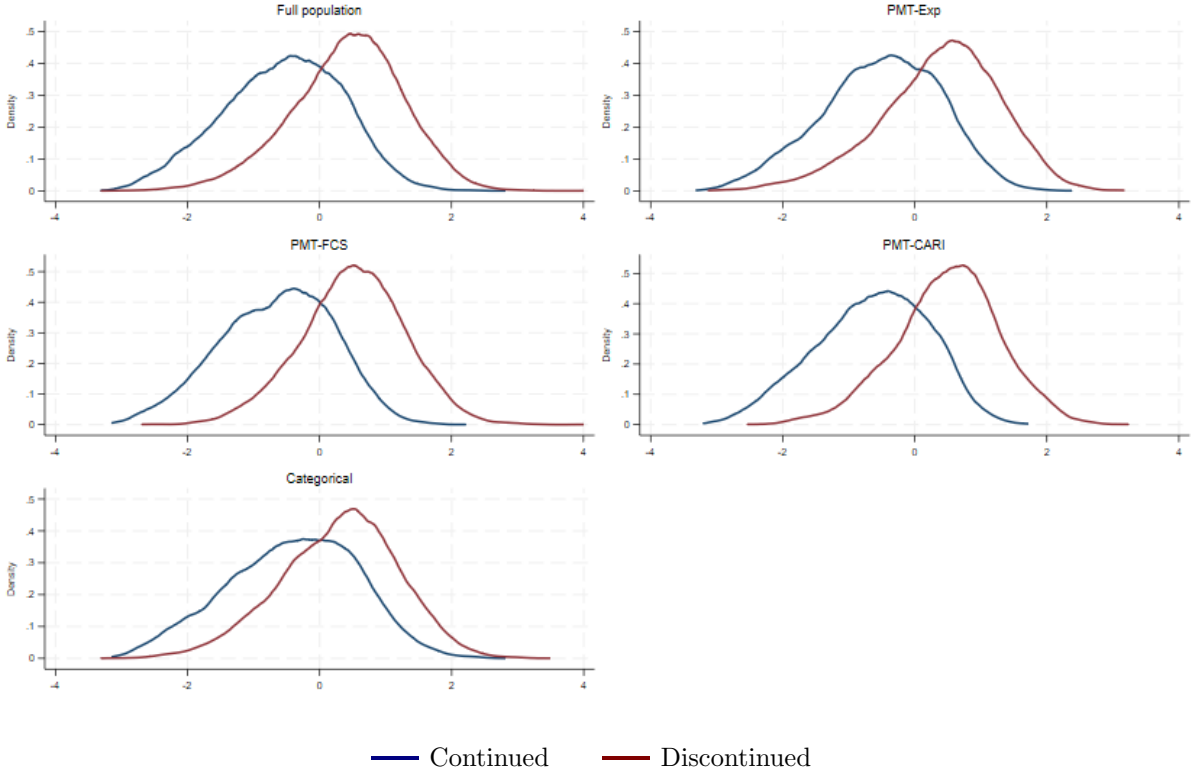
Notes: The figure summarizes the randomization process. The 58,501 beneficiary households were first randomly assigned to four prioritization groups of equal size. Within each group, households were ranked based on the specific prioritization method assigned to that group, and the 43 percent most vulnerable households were selected to keep receiving assistance while the remaining 57 percent were excluded from the program. Among households selected for discontinuation, a stratified random subset was assigned to the Transition Package, which extended assistance for six additional months before withdrawal.

Figure A.5. Distribution of Primary Outcomes across Assistance Status.



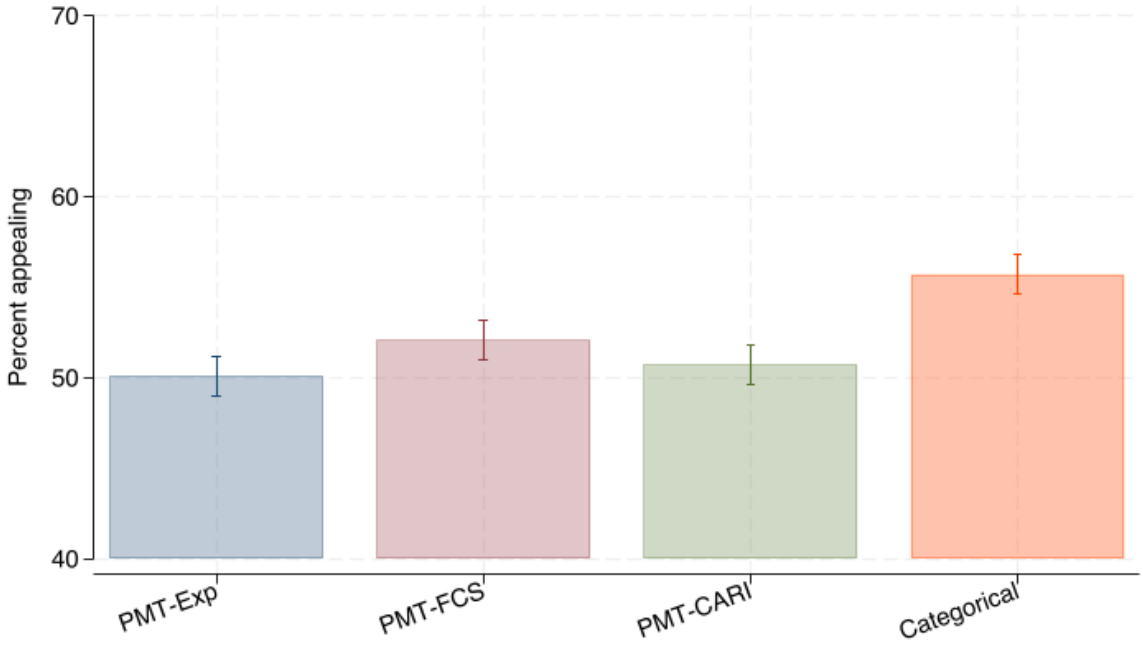
Notes: The figure plots kernel density estimates for the four main outcomes, separately by assistance status (continued assistance, regular-notice discontinued assistance, and transition package) and survey wave. Food expenditure is expressed in USD and includes food purchased on credit as well as the estimated value of home production. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, worriedness, and mental health. The trust index combines seven measures of trust toward other groups. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights.

Figure A.6. Baseline Wealth Distribution across Continued and Discontinued Households.



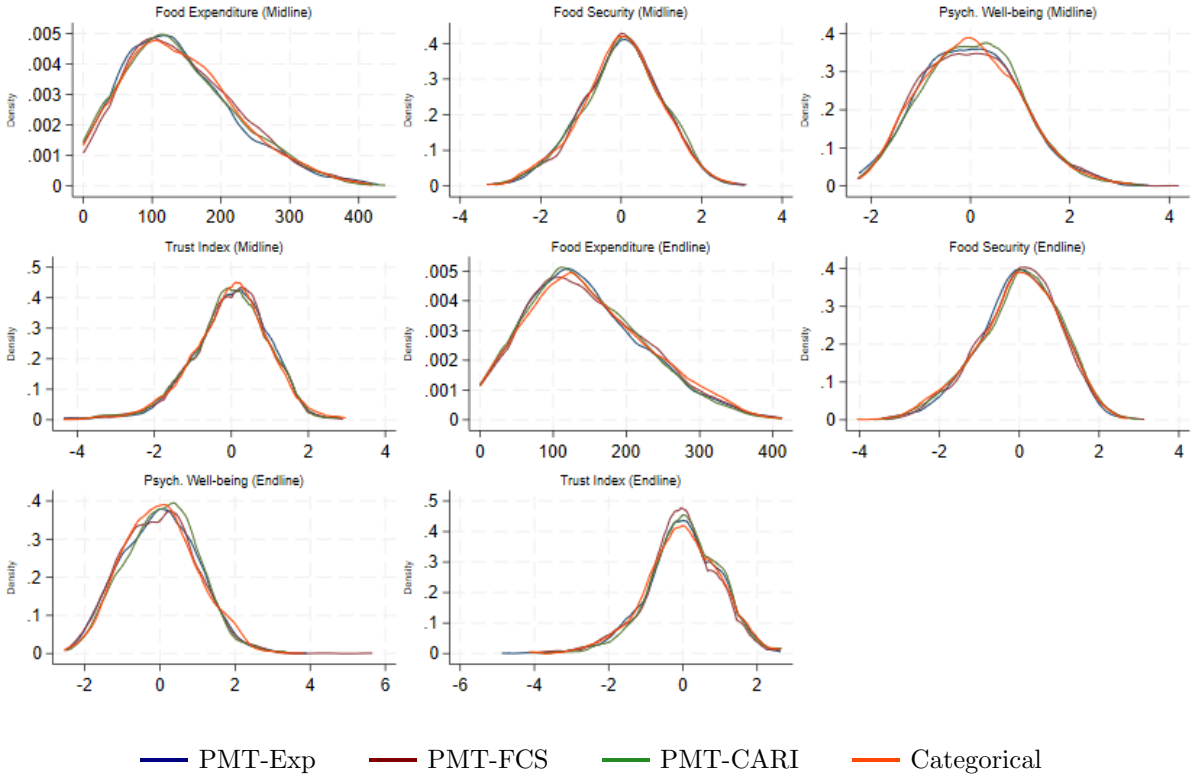
Notes: The figure plots kernel density estimates of the baseline wealth index for households assigned to continued versus discontinued assistance, in the full beneficiary population and separately by assigned prioritization method. Households assigned to the transition package are classified as discontinued. The wealth index is constructed from 36 indicators of asset ownership and connectivity using principal component analysis (PCA). It is standardized to have mean zero and standard deviation one in the full beneficiary population, with higher values indicating greater wealth.

Figure A.7. Households' Appeals by Prioritization Arm.



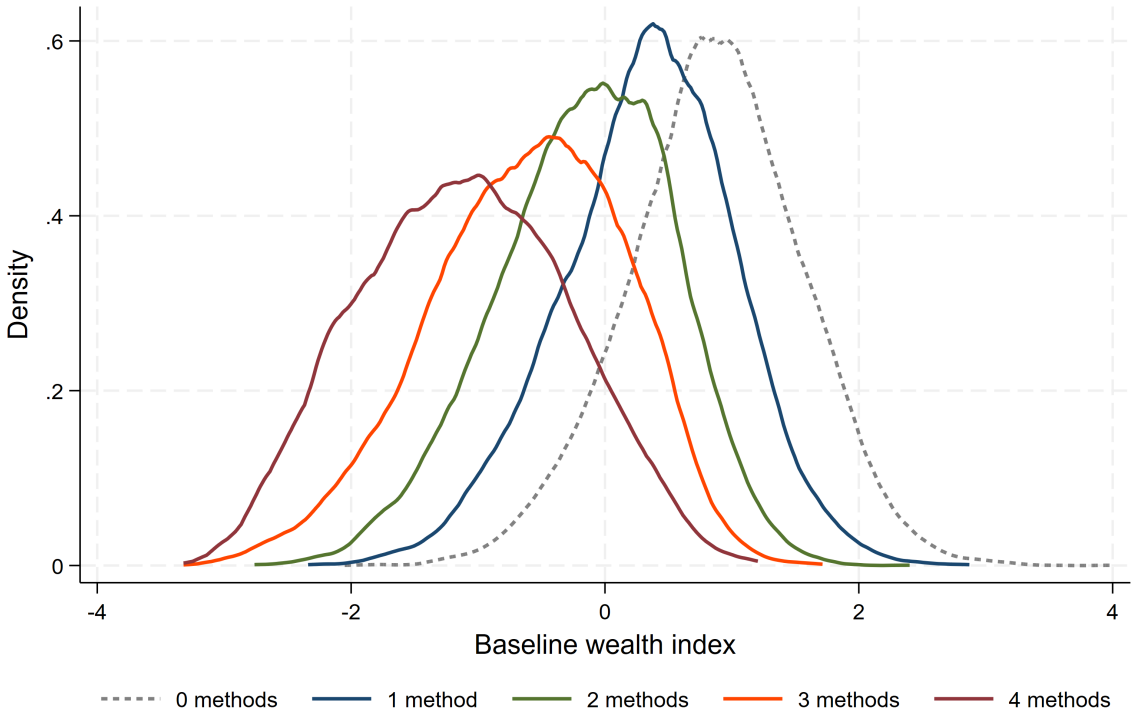
Notes: The figure reports appeal submission rates by assigned prioritization method. The sample includes all households whose assistance was discontinued. Bars show group means and lines indicate 95 percent confidence intervals.

Figure A.8. Distribution of Primary Outcomes Across the Four Prioritization Arms.



Notes: The figure plots kernel density estimates for the four main outcomes, separately by assigned prioritization method and survey wave. Food expenditure is expressed in USD and includes food purchased on credit as well as the estimated value of home production. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, worriedness, and mental health. The trust index combines seven measures of trust toward other groups. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights.

Figure A.9. Distribution of Baseline Wealth by Number of Eligible Methods.



Notes: The figure plots kernel density estimates of the baseline wealth index for households eligible under zero, one, two, three, or four targeting methods. The wealth index is constructed from 36 indicators of asset ownership and connectivity using principal component analysis (PCA). It is standardized to have mean zero and standard deviation one in the full beneficiary population, with higher values indicating greater wealth. The sample includes the full beneficiary population.

Table A.1. Balance Checks across Prioritization Arms (Full Population).

	PMT-Exp	PMT-FCS	PMT-CARI	Categorical	p-value
	(1)	(2)	(3)	(4)	(5)
Household size	3.86	3.85	3.86	3.85	0.679
	(1.83)	(1.81)	(1.82)	(1.83)	
School-age children	0.93	0.93	0.94	0.92	0.663
	(1.23)	(1.23)	(1.23)	(1.23)	
Female household head	0.30	0.30	0.30	0.30	0.887
	(0.46)	(0.46)	(0.46)	(0.46)	
Senior household head (65+)	0.25	0.25	0.25	0.25	0.952
	(0.43)	(0.43)	(0.43)	(0.43)	
Household head with primary education or lower	0.56	0.56	0.57	0.56	0.253
	(0.50)	(0.50)	(0.50)	(0.50)	
Any household member has chronic illness	0.74	0.74	0.74	0.74	0.849
	(0.44)	(0.44)	(0.44)	(0.44)	
Household head suffers from any disability	0.19	0.18	0.19	0.18	0.235
	(0.39)	(0.38)	(0.39)	(0.38)	
No household member employed	0.30	0.31	0.30	0.31	0.313
	(0.46)	(0.46)	(0.46)	(0.46)	
Dependency ratio	0.35	0.35	0.35	0.35	0.878
	(0.30)	(0.30)	(0.30)	(0.30)	
Household is located in Beirut	0.05	0.05	0.05	0.05	0.782
	(0.22)	(0.22)	(0.22)	(0.22)	
Household owns a smartphone	0.82	0.83	0.83	0.82	0.372
	(0.38)	(0.38)	(0.38)	(0.38)	
Household owns a car	0.37	0.37	0.37	0.37	0.827
	(0.48)	(0.48)	(0.48)	(0.48)	
Piped water	0.23	0.22	0.23	0.22	0.740
	(0.42)	(0.42)	(0.42)	(0.42)	
Gas for cooking	0.74	0.73	0.73	0.74	0.890
	(0.44)	(0.44)	(0.44)	(0.44)	
WiFi internet	0.52	0.52	0.52	0.51	0.244
	(0.50)	(0.50)	(0.50)	(0.50)	
Wealth index (PCA)	-0.00	0.01	0.00	-0.01	0.159
	(1.01)	(1.00)	(1.00)	(1.00)	
PMT-Expenditure score	15.25	15.25	15.25	15.25	0.569
	(0.49)	(0.49)	(0.49)	(0.49)	
PMT-FCS score	55.42	55.46	55.37	55.35	0.411
	(10.03)	(10.00)	(9.96)	(9.87)	
PMT-CARI score	0.37	0.37	0.37	0.37	0.331
	(0.19)	(0.19)	(0.19)	(0.19)	
Categorical score	33.80	33.40	33.77	34.04	0.189
	(44.56)	(44.41)	(44.75)	(44.52)	
Observations	14,624	14,626	14,625	14,626	58,501

Notes: The table reports baseline means for the full beneficiary population, by prioritization arm. Standard deviations are reported in parentheses. The final column reports the p-value from a test of equality of means across all groups (H_0 := means are equal across groups). The table presents only a subset of the available variables. When considering the full set of baseline variables and targeting scores, there are 160 possible balance tests. Of these, 11 (6.9%) are significant at the 10 percent level ($p < 0.10$), 6 (3.8%) at the 5 percent level ($p < 0.05$), and 1 (0.6%) at the 1 percent level ($p < 0.01$).

Table A.2. Baseline Comparison of Households Selected for Continued Assistance.

	PMT-Exp	PMT-FCS	PMT-CARI	Categorical	p-value
	(1)	(2)	(3)	(4)	(5)
Household size	4.74	4.02	4.52	3.19	0.000
	(1.81)	(2.03)	(1.97)	(1.98)	
School-age children	1.44	1.11	1.37	0.52	0.000
	(1.41)	(1.40)	(1.44)	(1.05)	
Female household head	0.25	0.32	0.28	0.53	0.000
	(0.43)	(0.47)	(0.45)	(0.50)	
Senior household head (65+)	0.17	0.25	0.19	0.47	0.000
	(0.38)	(0.43)	(0.40)	(0.50)	
Household head with primary education or lower	0.62	0.72	0.67	0.67	0.000
	(0.48)	(0.45)	(0.47)	(0.47)	
Any household member has chronic illness	0.70	0.75	0.74	0.85	0.000
	(0.46)	(0.43)	(0.44)	(0.35)	
Household head suffers from any disability	0.19	0.21	0.22	0.38	0.000
	(0.40)	(0.41)	(0.41)	(0.49)	
No household member employed	0.28	0.36	0.29	0.61	0.000
	(0.45)	(0.48)	(0.46)	(0.49)	
Dependency ratio	0.36	0.38	0.37	0.43	0.000
	(0.26)	(0.30)	(0.28)	(0.36)	
Household is located in Beirut	0.02	0.07	0.03	0.04	0.000
	(0.15)	(0.26)	(0.18)	(0.21)	
Household owns a smartphone	0.76	0.75	0.77	0.74	0.000
	(0.42)	(0.43)	(0.42)	(0.44)	
Household owns a car	0.17	0.16	0.18	0.22	0.000
	(0.37)	(0.36)	(0.39)	(0.41)	
Piped water	0.25	0.22	0.24	0.26	0.000
	(0.43)	(0.41)	(0.43)	(0.44)	
Gas for cooking	0.66	0.67	0.65	0.71	0.000
	(0.47)	(0.47)	(0.48)	(0.45)	
WiFi internet	0.40	0.28	0.35	0.37	0.000
	(0.49)	(0.45)	(0.48)	(0.48)	
Wealth index (PCA)	-0.51	-0.65	-0.65	-0.41	0.000
	(0.89)	(0.85)	(0.85)	(0.98)	
PMT-Expenditure score	14.80	14.99	14.88	15.22	0.000
	(0.29)	(0.43)	(0.36)	(0.47)	
PMT-FCS score	49.91	46.65	48.19	53.28	0.000
	(9.73)	(7.56)	(9.01)	(9.18)	
PMT-CARI score	0.52	0.51	0.55	0.41	0.000
	(0.15)	(0.15)	(0.12)	(0.19)	
Categorical score	36.48	45.96	42.90	75.87	0.000
	(41.26)	(45.38)	(44.00)	(32.57)	
Observations	6,289	6,289	6,290	6,292	25,160

Notes: The table reports baseline means for households selected to continue receiving assistance, by prioritization arm. Standard deviations are reported in parentheses. The final column reports the p-value from a test of equality of means across all groups ($H_0 :=$ means are equal across groups). The table presents only a subset of the available variables. When considering the full set of baseline variables and targeting scores, there are 160 possible balance tests. Of these, 144 (90.0%) are significant at the 10 percent level ($p < 0.10$), 139 (86.9%) at the 5 percent level ($p < 0.05$), and 135 (84.4%) at the 1 percent level ($p < 0.01$).

Table A.3. Population Distribution across Eligibility Spaces.

Eligibility Space	Total (%)	Method 1 PMT-Exp	Method 2 PMT-FCS	Method 3 PMT-CARI	Method 4 Categorical
E0	15295 (26.14%)	3824	3824	3824	3823
E1E2E3E4	7727 (13.21%)	1931	1932	1932	1932
E1	2724 (4.66%)	681	681	681	681
E1E2	706 (1.21%)	176	177	176	177
E1E2E3	7753 (13.25%)	1939	1938	1938	1938
E1E2E4	458 (0.78%)	114	115	114	115
E1E3	2825 (4.83%)	707	706	707	705
E1E3E4	1810 (3.09%)	453	452	452	453
E1E4	1153 (1.97%)	288	289	287	289
E2	1945 (3.32%)	485	486	487	487
E2E3	1390 (2.38%)	347	347	348	348
E2E3E4	2332 (3.99%)	583	583	583	583
E2E4	2845 (4.86%)	712	711	711	711
E3	703 (1.20%)	176	176	176	175
E3E4	616 (1.05%)	154	154	154	154
E4	8219 (14.05%)	2054	2055	2055	2055
Selected for assistance	25160 (43%)	6289	6289	6290	6292
Not selected	33341 (57%)	8335	8337	8335	8334
Total	**58501**	**14624**	**14626**	**14625**	**14626**

Notes: The table lists the 16 "eligibility spaces" generated by applying all four prioritization methods to the 2023 beneficiary population. Each space groups households that share the same eligibility pattern across the four methods, meaning that they would be selected for continued assistance under the same method or combination of methods. E0 denotes households that would not be selected by any method, while Ei indicates that households in that groups would be selected for continued assistance under method i. The Total (%) column reports the number and share of households in the 2023 beneficiary population that fall within each eligibility space. The remaining columns report the random assignment of households within each space across the four prioritization methods. Green numbers indicate that households were selected for continued assistance under the assigned method, while red numbers indicate that households were not selected for continued assistance under the assigned method.

Table A.4. Balance Checks across Assistance Status (Full Population).

			Discontinued			
	Continued	Discontinued	Regular Notice	Transition Package	p-value (1) vs. (2)	p-value (1) vs. (3) vs. (4)
	(1)	(2)	(3)	(4)	(5)	(6)
Household size	4.12	3.66	3.66	3.81	0.972	0.840
	(2.04)	(1.62)	(1.61)	(1.76)		
School-age children	1.11	0.79	0.79	0.83	0.934	0.456
	(1.38)	(1.09)	(1.09)	(1.12)		
Female household head	0.35	0.26	0.26	0.30	0.581	0.847
	(0.48)	(0.44)	(0.44)	(0.46)		
Senior household head (65+)	0.27	0.23	0.23	0.25	0.386	0.683
	(0.44)	(0.42)	(0.42)	(0.44)		
Household head with primary education or lower	0.67	0.48	0.47	0.59	0.650	0.510
	(0.47)	(0.50)	(0.50)	(0.49)		
Any household member has chronic illness	0.76	0.72	0.72	0.71	0.432	0.193
	(0.43)	(0.45)	(0.45)	(0.45)		
Household head suffers from any disability	0.25	0.13	0.13	0.19	0.053	0.118
	(0.43)	(0.34)	(0.34)	(0.40)		
No household member employed	0.39	0.24	0.24	0.30	0.262	0.529
	(0.49)	(0.43)	(0.43)	(0.46)		
Dependency ratio	0.38	0.32	0.32	0.34	0.944	0.242
	(0.30)	(0.30)	(0.30)	(0.31)		
Household is located in Beirut	0.04	0.06	0.06	0.06	0.499	0.642
	(0.20)	(0.23)	(0.23)	(0.25)		
Household owns a smartphone	0.75	0.88	0.88	0.82	0.890	0.895
	(0.43)	(0.33)	(0.32)	(0.38)		
Household owns a car	0.18	0.51	0.51	0.32	0.254	0.271
	(0.39)	(0.50)	(0.50)	(0.47)		
Piped water	0.24	0.21	0.21	0.23	0.760	0.946
	(0.43)	(0.41)	(0.41)	(0.42)		
Gas for cooking	0.67	0.78	0.78	0.74	0.866	0.954
	(0.47)	(0.41)	(0.41)	(0.44)		
WiFi internet	0.35	0.65	0.65	0.52	0.426	0.703
	(0.48)	(0.48)	(0.48)	(0.50)		
Wealth index (PCA)	-0.56	0.42	0.43	-0.06	0.831	0.679
	(0.90)	(0.85)	(0.85)	(0.77)		
PMT-Expenditure score	14.97	15.47	15.47	15.23	0.892	0.511
	(0.43)	(0.42)	(0.42)	(0.34)		
PMT-FCS score	49.51	59.84	59.98	55.43	0.307	0.466
	(9.24)	(7.99)	(7.99)	(6.74)		
PMT-CARI score	0.50	0.28	0.28	0.38	0.952	0.949
	(0.16)	(0.16)	(0.16)	(0.11)		
Categorical score	50.31	21.26	20.86	34.47	0.629	0.888
	(43.80)	(40.90)	(40.77)	(43.12)		
Observations	25,160	33,341	32,361	980	58,501	58,501

Notes: The table reports baseline means for the full beneficiary population, by assistance status. Standard deviations are reported in parentheses. Column (1) includes households assigned to continued assistance. Column (2) includes households assigned to discontinued assistance, combining regular-notice and transition-package households. Columns (3) and (4) report these two groups separately. Column (5) reports the p-value from a test of equality of means between continued and discontinued assistance households. Column (6) reports the p-value from a test of equality of means across continued assistance, regular-notice, and transition package households. Both tests control for eligibility-space fixed effects. The table presents only a subset of the available variables. Across the full set of baseline variables and targeting scores, there are 160 balance tests. Of these, for the comparison in column (5), 17 (10.6%) are significant at the 10 percent level ($p < 0.10$), 9 (5.6%) at the 5 percent level ($p < 0.05$), and 1 (0.6%) at the 1 percent level ($p < 0.01$). The corresponding numbers for the comparison in column (6) are 13 (8.1%), 8 (5.0%), and 2 (1.2%), respectively.

Table A.5. Comparison across Study Population and LHS Samples.

	Study Population	LHS	p-value
	(1)	(2)	(3)
Household size	3.86	3.77	0.006
	(1.82)	(1.80)	
School-age children	0.93	0.96	0.201
	(1.23)	(1.30)	
Female household head	0.30	0.17	0.000
	(0.46)	(0.38)	
Senior household head (65+)	0.25	0.28	0.000
	(0.43)	(0.45)	
Any household member has chronic illness	0.74	0.58	0.000
	(0.44)	(0.49)	
Household head suffers from any disability	0.18	0.08	0.000
	(0.39)	(0.28)	
Dependency ratio	0.35	0.37	0.000
	(0.30)	(0.31)	
Household is located in Beirut	0.05	0.03	0.000
	(0.22)	(0.18)	
Household owns a smartphone	0.83	0.94	0.000
	(0.38)	(0.24)	
Piped water	0.22	0.57	0.000
	(0.42)	(0.49)	
Gas for cooking	0.73	0.96	0.000
	(0.44)	(0.21)	
Household has WiFi internet	0.52	0.59	0.000
	(0.50)	(0.49)	
Household has mobile internet	0.49	0.59	0.000
	(0.50)	(0.49)	
Household owns a refrigerator	0.91	0.93	0.000
	(0.29)	(0.25)	
Household owns a washing machine	0.87	0.63	0.000
	(0.34)	(0.48)	
Household owns a flat screen TV	0.63	0.79	0.000
	(0.48)	(0.41)	
Household owns a computer/laptop	0.08	0.17	0.000
	(0.26)	(0.38)	
Flush toilet	0.86	0.96	0.000
	(0.35)	(0.20)	
Finished floor	0.91	0.93	0.001
	(0.28)	(0.26)	
Wealth index (PCA)	0.00	0.41	0.000
	(1.00)	(1.15)	
Observations	58,501	3,511	62,012

Notes: The table compares households in the 2023 beneficiary population with households in the 2023 Lebanon Household Survey (LHS) sample. Information on the beneficiary population comes from the population census. The final column reports the p-value from a test of equality of means across the two groups (H_0 := mean is equal across groups).

Table A.6. Balance Checks across Assistance Status (Study Sample).

			Discontinued			
	Continued	Discontinued	Regular Notice	Transition Package	p-value (1) vs. (2)	p-value (1) vs. (3) vs. (4)
	(1)	(2)	(3)	(4)	(5)	(6)
Household size	3.85	3.76	3.73	3.81	0.975	0.322
	(1.93)	(1.67)	(1.63)	(1.76)		
School-age children	0.91	0.83	0.83	0.83	0.321	0.599
	(1.26)	(1.11)	(1.11)	(1.12)		
Female household head	0.36	0.29	0.29	0.30	0.278	0.322
	(0.48)	(0.46)	(0.45)	(0.46)		
Senior household head (65+)	0.30	0.23	0.23	0.25	0.212	0.402
	(0.46)	(0.42)	(0.42)	(0.44)		
Household head with primary education or lower	0.64	0.55	0.54	0.59	0.787	0.595
	(0.48)	(0.50)	(0.50)	(0.49)		
Any household member has chronic illness	0.75	0.73	0.74	0.71	0.154	0.058
	(0.43)	(0.44)	(0.44)	(0.45)		
Household head suffers from any disability	0.24	0.16	0.15	0.19	0.749	0.306
	(0.43)	(0.37)	(0.36)	(0.40)		
No household member employed	0.40	0.28	0.26	0.30	0.595	0.845
	(0.49)	(0.45)	(0.44)	(0.46)		
Dependency ratio	0.37	0.33	0.33	0.34	0.833	0.427
	(0.31)	(0.30)	(0.30)	(0.31)		
Household is located in Beirut	0.06	0.06	0.05	0.06	0.096	0.153
	(0.23)	(0.23)	(0.23)	(0.25)		
Household owns a smartphone	0.77	0.84	0.84	0.82	0.546	0.828
	(0.42)	(0.37)	(0.36)	(0.38)		
Household owns a car	0.22	0.37	0.40	0.32	0.399	0.620
	(0.41)	(0.48)	(0.49)	(0.47)		
Piped water	0.26	0.22	0.22	0.23	0.632	0.883
	(0.44)	(0.42)	(0.42)	(0.42)		
Gas for cooking	0.70	0.76	0.77	0.74	0.669	0.854
	(0.46)	(0.43)	(0.42)	(0.44)		
WiFi internet	0.40	0.55	0.57	0.52	0.362	0.653
	(0.49)	(0.50)	(0.50)	(0.50)		
Wealth index (PCA)	-0.41	0.06	0.11	-0.06	0.542	0.802
	(0.85)	(0.83)	(0.85)	(0.77)		
PMT-Expenditure score	15.08	15.30	15.32	15.23	0.908	0.344
	(0.37)	(0.36)	(0.37)	(0.34)		
PMT-FCS score	52.18	56.46	56.91	55.43	0.591	0.809
	(7.85)	(7.20)	(7.35)	(6.74)		
PMT-CARI score	0.45	0.36	0.35	0.38	1.000	0.808
	(0.13)	(0.13)	(0.13)	(0.11)		
Categorical score	48.90	30.03	28.10	34.47	0.445	0.745
	(43.59)	(42.26)	(41.74)	(43.12)		
Observations	3,200	3,234	2,254	980	6,434	6,434

Notes: The table reports baseline means for the study sample, by assistance status. Standard deviations are reported in parentheses. Column (1) includes households assigned to continued assistance. Column (2) includes households assigned to discontinued assistance, combining regular-notice and transition-package households. Columns (3) and (4) report these two groups separately. Column (5) reports the p-value from a test of equality of means between continued and discontinued assistance households. Column (6) reports the p-value from a test of equality of means across continued assistance, regular-notice, and transition package households. Both tests control for eligibility-space fixed effects. The table presents only a subset of the available variables. Across the full set of baseline variables and targeting scores, there are 160 balance tests. Of these, for the comparison in column (5), 21 (13.1%) are significant at the 10 percent level ($p < 0.10$), 12 (7.5%) at the 5 percent level ($p < 0.05$), and 3 (1.9%) at the 1 percent level ($p < 0.01$). The corresponding numbers for the comparison in column (6) are 23 (14.4%), 12 (7.5%), and 3 (1.9%), respectively.

Table A.7. Balance Checks across Midline-Surveyed Long Notice.

	Continued	Regular Notice	Transition Package	p-value
	(1)	(2)	(3)	(4)
Household size	3.85	3.73	4.02	0.322
	(1.93)	(1.63)	(1.71)	
School-age children	0.91	0.83	0.89	0.742
	(1.26)	(1.11)	(1.18)	
Female household head	0.36	0.29	0.26	0.131
	(0.48)	(0.45)	(0.44)	
Senior household head (65+)	0.30	0.23	0.23	0.357
	(0.46)	(0.42)	(0.42)	
Household head with primary education or lower	0.64	0.54	0.65	0.041
	(0.48)	(0.50)	(0.48)	
Any household member has chronic illness	0.75	0.74	0.72	0.106
	(0.43)	(0.44)	(0.45)	
Household head suffers from any disability	0.24	0.15	0.20	0.588
	(0.43)	(0.36)	(0.40)	
No household member employed	0.40	0.26	0.28	0.895
	(0.49)	(0.44)	(0.45)	
Dependency ratio	0.37	0.33	0.32	0.728
	(0.31)	(0.30)	(0.29)	
Household is located in Beirut	0.06	0.05	0.04	0.096
	(0.23)	(0.23)	(0.20)	
Household owns a smartphone	0.77	0.84	0.80	0.673
	(0.42)	(0.36)	(0.40)	
Household owns a car	0.22	0.40	0.33	0.758
	(0.41)	(0.49)	(0.47)	
Piped water	0.26	0.22	0.25	0.817
	(0.44)	(0.42)	(0.43)	
Gas for cooking	0.70	0.77	0.74	0.806
	(0.46)	(0.42)	(0.44)	
WiFi internet	0.40	0.57	0.53	0.563
	(0.49)	(0.50)	(0.50)	
Wealth index (PCA)	-0.41	0.11	-0.12	0.771
	(0.85)	(0.85)	(0.76)	
PMT-Expenditure score	15.08	15.32	15.20	0.222
	(0.37)	(0.37)	(0.27)	
PMT-FCS score	52.18	56.91	55.27	0.475
	(7.85)	(7.35)	(5.71)	
PMT-CARI score	0.45	0.35	0.40	0.844
	(0.13)	(0.13)	(0.08)	
Categorical score	48.90	28.10	33.26	0.773
	(43.59)	(41.74)	(42.72)	
Observations	3,200	2,254	293	5,747

Notes: The table reports baseline means for the study sample, by assistance status. Standard deviations are reported in parentheses. Columns (1) and (2) mirror columns (1) and (3) of Table A.6. Column (3) restricts the focus to transition package households that were actually surveyed at midline. Assigned transition package households not surveyed at midline are excluded. The final column reports the p-value from a test of equality of means across the three groups (H_0 := mean is equal across groups), controlling for eligibility-strata fixed effects. The table presents only a subset of the available variables. Across the full set of baseline variables and targeting scores, there are 160 balance tests. Of these, 21 (13.1%) are significant at the 10 percent level ($p < 0.10$), 13 (8.1%) at the 5 percent level ($p < 0.05$), and 3 (1.9%) at the 1 percent level ($p < 0.01$).

Table A.8. Impact of Assistance Discontinuation on Food Expenditure.

		In WFP basket					Not in WFP basket			
	Food Exp. (USD)	Staples	Pulses	Oil/Fat	Sugar	Condiments	Vegetables	Fruits	Protein	Milk/Dairy
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Midline										
Discontinued	12.46***	6.04***	4.95***	3.60***	1.36***	0.52***	-1.03*	-0.77***	-1.12*	-0.42
	(2.21)	(0.77)	(0.27)	(0.20)	(0.17)	(0.17)	(0.59)	(0.29)	(0.59)	(0.36)
	—	[0.000]	[0.000]	[0.000]	[0.000]	[0.006]	[0.093]	[0.011]	[0.077]	[0.245]
Transition Package	-5.16	-3.90**	-4.61***	-2.79***	-1.50***	-0.52	2.40*	1.30**	3.78***	0.84
	(4.64)	(1.67)	(0.56)	(0.43)	(0.35)	(0.37)	(1.33)	(0.65)	(1.34)	(0.82)
	—	[0.035]	[0.000]	[0.000]	[0.000]	[0.177]	[0.091]	[0.068]	[0.011]	[0.305]
Discontinued + TP	7.30	2.14	0.34	0.81*	-0.14	-0.00	1.37	0.53	2.66**	0.42
q-val(Disc+TP=0)	—	0.589	0.779	0.277	0.783	0.996	0.676	0.741	0.277	0.779
Mean Continued	127.94	35.08	5.21	5.52	4.49	5.43	24.34	9.50	24.78	14.69
Observations	5,566	5,566	5,566	5,566	5,566	5,566	5,566	5,566	5,566	5,566
Panel B: Endline										
Discontinued	10.17***	5.22***	2.77***	2.10***	0.72***	0.27	-0.03	-0.26	0.18	0.20
	(2.34)	(0.83)	(0.23)	(0.19)	(0.18)	(0.21)	(0.51)	(0.29)	(0.66)	(0.38)
	—	[0.000]	[0.000]	[0.000]	[0.000]	[0.332]	[0.957]	[0.578]	[0.886]	[0.759]
Transition Package	2.64	1.02	0.02	0.05	-0.19	-0.03	0.90	0.57	0.01	-0.11
	(3.08)	(1.09)	(0.31)	(0.24)	(0.23)	(0.27)	(0.68)	(0.39)	(0.90)	(0.49)
	—	[0.932]	[0.988]	[0.988]	[0.932]	[0.988]	[0.823]	[0.823]	[0.988]	[0.988]
Discontinued + TP	12.80***	6.24***	2.79***	2.15***	0.53**	0.24	0.88	0.32	0.19	0.09
q-val(Disc+TP=0)	—	0.000	0.000	0.000	0.045	0.511	0.331	0.511	0.846	0.846
Mean Continued	134.55	37.14	5.19	6.93	5.46	6.63	21.91	10.08	26.87	15.53
Observations	5,654	5,654	5,654	5,654	5,654	5,654	5,654	5,654	5,654	5,654

Notes: The sample is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. The table reports estimated effects on total food expenditure (column (1)) and expenditure by food category (columns (2)-(10)). Expenditure is measured in USD and includes food purchased on credit and the estimated value of home production. Columns (2)-(6) cover food categories included in the WFP assistance basket, while columns (7)-(10) cover other food categories. The q-value row reports the FDR-adjusted q-value from a test of the null hypothesis that the sum of the coefficients on Discontinued and Transition Package equals zero. This corresponds to testing whether the effect for transition package households, relative to continued assistance households, is zero. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.9. Impact of Assistance Discontinuation on Credit and Non-Credit Purchases.

	Food Non-Credit	Food Credit	Non-Food Non-Credit	Non-Food Credit
	(1)	(2)	(3)	(4)
Panel A: Midline				
Discontinued	9.31***	3.95***	0.08	2.71
	(2.35)	(1.17)	(4.13)	(1.81)
	[0.000]	[0.002]	[0.985]	[0.179]
Transition Package	-6.22	1.16	15.21*	1.92
	(5.30)	(2.60)	(9.01)	(4.11)
	[0.481]	[0.654]	[0.366]	[0.654]
Discontinued + TP	3.09	5.11**	15.28*	4.63
q-val(Disc+TP=0)	0.557	0.176	0.176	0.346
Mean Continued	107.04	22.43	213.02	38.75
Observations	5,544	5,544	5,566	5,566
Panel B: Endline				
Discontinued	8.38***	3.04**	-3.81	4.83*
	(2.48)	(1.47)	(4.89)	(2.69)
	[0.003]	[0.076]	[0.436]	[0.096]
Transition Package	2.68	-0.54	3.00	-6.62**
	(3.26)	(1.91)	(6.63)	(3.35)
	[0.777]	[0.777]	[0.777]	[0.191]
Discontinued + TP	11.05***	2.50	-0.80	-1.79
q-val(Disc+TP=0)	0.002	0.357	0.901	0.778
Mean Continued	107.40	28.84	258.74	58.41
Observations	5,621	5,621	5,654	5,654

Notes: The sample is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. Columns (2) and (4) report the value of food and non-food purchases made on credit, respectively. Columns (1) and (3) report total food and non-food expenditure, respectively, net of the corresponding purchases on credit. Non-credit food expenditure includes the estimated value of home production. All outcomes are measured in USD. The q-value row reports the FDR-adjusted q-value from a test of the null hypothesis that the sum of the coefficients on Discontinued and Transition Package equals zero. This corresponds to testing whether the effect for transition package households, relative to continued assistance households, is zero. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.10. Impact of Assistance Discontinuation on the Secondary Outcomes.

	Food Security		Psychological Well-being								
	FCS	rCSI	Life Satisfaction	Worriedness	WHO-5	Non-food exp.	Social cohesion	Employed (%)	Hrs per member	HH has debt	Remittances
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Midline											
Discontinued	-1.60***	0.69*	-0.35***	0.29*	-1.02*	2.79	0.01	-1.07	-0.33	0.01	-0.01
	(0.48)	(0.37)	(0.09)	(0.17)	(0.57)	(5.00)	(0.03)	(0.72)	(0.33)	(0.01)	(0.01)
	[0.002]	[0.060]	[0.000]	[0.094]	[0.094]	[0.693]	[0.831]	[0.693]	[0.693]	[0.693]	[0.693]
Transition Package	1.12	-1.10	0.34*	-0.22	0.68	17.13	0.07	-1.83	-0.22	0.01	-0.04*
	(1.08)	(0.80)	(0.19)	(0.33)	(1.18)	(10.89)	(0.07)	(1.50)	(0.76)	(0.03)	(0.02)
	[0.300]	[0.300]	[0.206]	[0.564]	[0.564]	[0.348]	[0.502]	[0.443]	[0.770]	[0.770]	[0.348]
Discontinued + TP	-0.47	-0.41	-0.01	0.07	-0.34	19.92*	0.07	-2.91*	-0.55	0.02	-0.04**
q-val(Disc+TP=0)	0.664	0.664	0.969	0.969	0.969	0.133	0.419	0.133	0.463	0.463	0.133
Mean Continued	54.86	16.57	11.39	29.89	23.05	251.77	0.00	24.01	9.07	0.44	0.15
Observations	5,566	5,566	5,459	5,174	5,517	5,566	5,566	5,566	5,566	5,544	5,552
Panel B: Endline											
Discontinued	-0.10	1.21***	-0.51***	-0.01	-0.96*	1.02	0.03	-0.06	0.16	0.01	-0.00
	(0.52)	(0.36)	(0.09)	(0.16)	(0.53)	(6.04)	(0.03)	(0.75)	(0.33)	(0.02)	(0.01)
	[0.854]	[0.001]	[0.000]	[0.973]	[0.104]	[0.937]	[0.937]	[0.937]	[0.937]	[0.937]	[0.937]
Transition Package	-0.25	-0.21	0.06	0.21	-0.63	-3.62	-0.02	-1.11	-0.61	0.00	0.00
	(0.68)	(0.46)	(0.11)	(0.21)	(0.69)	(7.99)	(0.04)	(0.97)	(0.42)	(0.02)	(0.01)
	[0.710]	[0.710]	[0.565]	[0.545]	[0.545]	[0.861]	[0.861]	[0.764]	[0.764]	[0.861]	[0.861]
Discontinued + TP	-0.35	0.99**	-0.45***	0.20	-1.59**	-2.59	0.01	-1.17	-0.46	0.01	0.00
q-val(Disc+TP=0)	0.602	0.046	0.000	0.339	0.032	0.886	0.886	0.786	0.786	0.886	0.968
Mean Continued	53.18	14.66	11.38	30.67	20.78	317.15	0.00	23.43	8.50	0.52	0.06
Observations	5,654	5,654	5,565	5,480	5,643	5,654	5,654	5,653	5,653	5,644	5,652

Notes: The sample is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. The Food Consumption Score (FCS) ranges from 0 to 112, with higher values indicating higher food consumption. The reduced Coping Strategies Index (rCSI) ranges from 0 to 56, with higher values indicating greater reliance on coping strategies. Life satisfaction is the sum of five items (5-20), with higher values indicating greater satisfaction. Worriedness is the sum of ten items (10-40), with higher values indicating greater worry. The WHO-5 measures mental health on a 0-100 scale, with higher values indicating better well-being. Non-food expenditure is measured in USD. The social cohesion index combines lower community tension, volunteering, and any association membership using inverse-covariance weights applied to components standardized within each survey wave, with higher values indicating greater social cohesion. Employed share is the share of household members aged 15 or older who are employed. Hours worked is the total weekly hours worked by household members aged 15 or older divided by the number of household members in that group. Debt and remittances are indicators for having any outstanding debt and receiving any remittances during the survey reference period, respectively. The q-value row reports the FDR-adjusted q-value from a test of the null hypothesis that the sum of the coefficients on Discontinued and Transition Package equals zero. This corresponds to testing whether the effect for transition package households, relative to continued assistance households, is zero. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.11. Impact of Assistance Discontinuation on Consumption.

		In WFP basket					Not in WFP basket			
	Food Cons. (FCS)	Staples	Pulses	Oil/Fat	Sugar	Condiments	Vegetables	Fruits	Protein	Milk/Dairy
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Midline										
Discontinued	-1.60***	-0.06	-0.33***	-0.17***	-0.23***	-0.16**	-0.05	-0.09	-0.05	0.01
	(0.48)	(0.04)	(0.05)	(0.06)	(0.08)	(0.07)	(0.06)	(0.05)	(0.04)	(0.07)
	—	[0.216]	[0.000]	[0.012]	[0.012]	[0.036]	[0.430]	[0.183]	[0.319]	[0.901]
Transition Package	1.12	0.06	0.16	0.09	0.11	0.03	0.10	0.27**	0.04	-0.03
	(1.08)	(0.08)	(0.11)	(0.12)	(0.17)	(0.14)	(0.13)	(0.12)	(0.09)	(0.15)
	—	[0.748]	[0.594]	[0.748]	[0.748]	[0.850]	[0.748]	[0.193]	[0.823]	[0.850]
Discontinued + TP	-0.47	-0.00	-0.17	-0.08	-0.11	-0.14	0.05	0.18	-0.01	-0.02
q-val(Disc+TP=0)	—	0.982	0.548	0.904	0.904	0.904	0.982	0.548	0.982	0.982
Mean Continued	54.86	6.30	3.90	5.69	4.02	5.15	3.79	1.80	1.53	3.50
Observations	5,566	5,566	5,566	5,566	5,566	5,566	5,566	5,566	5,566	5,566
Panel B: Endline										
Discontinued	-0.10	0.07*	-0.21***	-0.01	-0.15*	0.03	-0.00	-0.05	0.03	0.10
	(0.52)	(0.04)	(0.05)	(0.06)	(0.09)	(0.06)	(0.06)	(0.06)	(0.04)	(0.08)
	—	[0.229]	[0.000]	[0.940]	[0.229]	[0.765]	[0.940]	[0.652]	[0.684]	[0.427]
Transition Package	-0.25	-0.09	0.09	0.05	0.00	0.01	0.00	0.02	-0.01	-0.08
	(0.68)	(0.06)	(0.06)	(0.08)	(0.11)	(0.08)	(0.08)	(0.08)	(0.05)	(0.10)
	—	[0.685]	[0.685]	[0.972]	[0.972]	[0.972]	[0.972]	[0.972]	[0.972]	[0.972]
Discontinued + TP	-0.35	-0.02	-0.12*	0.04	-0.15	0.04	-0.00	-0.04	0.02	0.01
q-val(Disc+TP=0)	—	0.976	0.562	0.976	0.783	0.976	0.993	0.976	0.976	0.990
Mean Continued	53.18	6.28	3.20	5.83	3.29	5.81	3.41	1.83	1.45	3.86
Observations	5,654	5,654	5,654	5,654	5,654	5,654	5,654	5,654	5,654	5,654

Notes: The sample is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. The table reports estimated effects on the Food Consumption Score (FCS) and food-group consumption frequencies. Column (1) reports effects in FCS points, with the score ranging from 0 to 112. Columns (2)-(10) report effects on the number of days, out of the previous seven, on which the household consumed each food category, ranging from 0 to 7. The FCS combines these frequencies using weights of 2 for staples, 3 for pulses, 0.5 each for oil/fat and sugar, 1 each for vegetables and fruits, and 4 each for animal protein and dairy. Condiments are reported separately and receive zero weight in the FCS. Columns (2)-(6) cover food categories included in the WFP assistance basket, while columns (7)-(10) cover other food categories. The q-value row reports the FDR-adjusted q-value from a test of the null hypothesis that the sum of the coefficients on Discontinued and Transition Package equals zero. This corresponds to testing whether the effect for transition package households, relative to continued assistance households, is zero. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.12. Impact of Assistance Discontinuation on Risky Coping Strategies.

	rCSI	Less Pref. Food	Borrowed Food	Reduced Meal Size	Reduced Meals	Reduced Adult Cons.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Midline						
Discontinued	0.69*	0.02	0.11*	0.15*	0.15*	0.05
	(0.37)	(0.07)	(0.06)	(0.09)	(0.09)	(0.07)
	—	[0.820]	[0.134]	[0.134]	[0.134]	[0.538]
Transition Package	-1.10	-0.17	-0.09	-0.28	-0.24	-0.08
	(0.80)	(0.15)	(0.12)	(0.18)	(0.18)	(0.14)
	—	[0.401]	[0.565]	[0.401]	[0.401]	[0.587]
Discontinued + TP	-0.41	-0.16	0.02	-0.13	-0.09	-0.02
q-val(Disc+TP=0)	—	0.880	0.880	0.880	0.880	0.880
Mean Continued	16.57	4.96	1.20	2.87	2.79	1.18
Observations	5,566	5,566	5,566	5,566	5,566	5,566
Panel B: Endline						
Discontinued	1.21***	0.07	0.21***	0.14	0.19**	0.13**
	(0.36)	(0.07)	(0.06)	(0.09)	(0.09)	(0.06)
	—	[0.336]	[0.002]	[0.147]	[0.072]	[0.072]
Transition Package	-0.21	0.09	-0.05	0.05	0.08	-0.11
	(0.46)	(0.09)	(0.08)	(0.12)	(0.12)	(0.08)
	—	[0.642]	[0.642]	[0.695]	[0.642]	[0.642]
Discontinued + TP	0.99**	0.16*	0.16**	0.19	0.27**	0.02
q-val(Disc+TP=0)	—	0.132	0.071	0.132	0.071	0.813
Mean Continued	14.66	4.93	0.98	2.79	2.48	0.83
Observations	5,654	5,654	5,654	5,654	5,654	5,654

Notes: The sample is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. The table reports estimated effects on the reduced Coping Strategies Index (rCSI) and its five components. Column (1) measures effects in rCSI points, with the index ranging from 0 to 56. Columns (2)-(6) measure effects on the number of days the household used each coping strategy during the previous seven days, ranging from 0 to 7. The rCSI combines these frequencies using weights of 1 for consuming less-preferred food, 2 for borrowing food, 1 for reducing meal size, 1 for reducing the number of meals, and 3 for reducing adults' consumption so children could eat. The q-value row reports the FDR-adjusted q-value from a test of the null hypothesis that the sum of the coefficients on Discontinued and Transition Package equals zero. This corresponds to testing whether the effect for transition package households, relative to continued assistance households, is zero. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.13. Impact of Assistance Discontinuation on Main Outcomes.

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
Discontinued	0.16***	-0.10***	-0.10***	-0.15***
	(0.03)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.000]	[0.000]	[0.000]
Transition Package	-0.06	0.10	0.08	0.12*
	(0.06)	(0.06)	(0.06)	(0.07)
	[0.337]	[0.201]	[0.241]	[0.201]
Discontinued + TP	0.10	0.00	-0.02	-0.02
q-val(Disc+TP=0)	0.402	0.963	0.963	0.963
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,566	5,566	5,566	5,564
Panel B: Endline				
Discontinued	0.14***	-0.08**	-0.12***	-0.06**
	(0.03)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.022]	[0.000]	[0.033]
Transition Package	0.03	0.00	-0.02	0.03
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.773]	[0.950]	[0.773]	[0.773]
Discontinued + TP	0.17***	-0.08*	-0.15***	-0.03
q-val(Disc+TP=0)	0.000	0.094	0.000	0.435
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,654	5,654	5,654	5,654

Notes: The sample is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. Food expenditure includes food purchased on credit as well as the estimated value of home production. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, reverse-coded worriedness, and mental health. The trust index combines seven measures of trust in different groups and institutions. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights. Outcomes are standardized within each survey wave using the mean and standard deviation of continued-assistance households The q-value row reports the FDR-adjusted q-value from a test of the null hypothesis that the sum of the coefficients on Discontinued and Transition Package equals zero. This corresponds to testing whether the effect for transition package households, relative to continued assistance households, is zero. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.14. Impact of Assistance Discontinuation on Trust.

	Trust Index	Family	Neighbors	Relig. Org.	UN/NGOs	Local Gov.	National Gov.	Migrants
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Midline								
Discontinued	-0.15***	-0.05	-0.00	-0.05	-0.25***	-0.06**	-0.02	-0.09***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
	—	[0.163]	[0.963]	[0.163]	[0.000]	[0.084]	[0.534]	[0.016]
Transition Package	0.12*	-0.01	0.09	-0.01	0.21***	-0.01	0.01	0.11*
	(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)
	—	[0.918]	[0.288]	[0.918]	[0.006]	[0.918]	[0.918]	[0.288]
Discontinued + TP	-0.02	-0.06	0.09	-0.06	-0.03	-0.07	-0.01	0.03
q-val(Disc+TP=0)	—	0.780	0.780	0.780	0.813	0.780	0.874	0.813
Mean Continued	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Observations	5,564	5,541	5,503	5,340	5,367	5,366	5,357	5,244
Panel B: Endline								
Discontinued	-0.06**	0.02	-0.05	-0.04	-0.13***	-0.04	-0.03	-0.04
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
	—	[0.434]	[0.352]	[0.352]	[0.000]	[0.352]	[0.363]	[0.352]
Transition Package	0.03	-0.00	0.02	0.05	0.05	0.03	0.02	-0.00
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
	—	[0.997]	[0.899]	[0.841]	[0.841]	[0.899]	[0.899]	[0.997]
Discontinued + TP	-0.03	0.02	-0.03	0.01	-0.08**	-0.01	-0.01	-0.04
q-val(Disc+TP=0)	—	0.811	0.811	0.811	0.307	0.811	0.811	0.811
Mean Continued	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Observations	5,654	5,643	5,610	5,470	5,562	5,486	5,396	5,371

Notes: The sample is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. The table reports estimated effects on the trust index (column 1) and its seven components (columns (2)–(8)): trust in family, neighbors, religious organizations, the UN and NGOs, local government, national government, and migrants. Individual responses are measured on a four-point scale and coded so that higher values indicate greater trust. The trust index combines the seven measures using inverse-covariance weights. Outcomes are standardized within each survey wave using the mean and standard deviation of continued-assistance households. The q-value row reports the FDR-adjusted q-value from a test of the null hypothesis that the sum of the coefficients on Discontinued and Transition Package equals zero. This corresponds to testing whether the effect for transition package households, relative to continued assistance households, is zero. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values:$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.15. Impact of Assistance Discontinuation: Heterogeneity by Initial Wealth.

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
Discontinued	0.17***	-0.10***	-0.10***	-0.15***
	(0.03)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.000]	[0.000]	[0.000]
Baseline Wealth Index	0.21***	0.11***	-0.03	-0.06**
	(0.02)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.000]	[0.310]	[0.031]
Discontinued × Wealth Index	0.09***	0.04	0.01	0.02
	(0.03)	(0.03)	(0.04)	(0.04)
	[0.030]	[0.555]	[0.850]	[0.748]
Transition Package	-0.05	0.10*	0.08	0.13*
	(0.06)	(0.06)	(0.06)	(0.07)
	[0.385]	[0.187]	[0.284]	[0.187]
Transition Package × Wealth Index	-0.07	0.02	0.11	-0.04
	(0.08)	(0.08)	(0.08)	(0.09)
	[0.749]	[0.835]	[0.737]	[0.835]
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,566	5,566	5,566	5,564
Panel B: Endline				
Discontinued	0.15***	-0.08**	-0.12***	-0.07**
	(0.03)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.020]	[0.000]	[0.025]
Baseline Wealth Index	0.15***	0.02	0.04	-0.10***
	(0.03)	(0.03)	(0.03)	(0.02)
	[0.000]	[0.344]	[0.162]	[0.000]
Discontinued × Wealth Index	0.05	0.04	0.01	0.00
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.650]	[0.650]	[0.934]	[0.934]
Transition Package	0.03	0.01	-0.03	0.04
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.720]	[0.822]	[0.720]	[0.720]
Transition Package × Wealth Index	0.01	-0.07	0.02	-0.05
	(0.05)	(0.06)	(0.05)	(0.05)
	[0.900]	[0.686]	[0.900]	[0.686]
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,654	5,654	5,654	5,654

Notes: The sample is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. Food expenditure includes food purchased on credit as well as the estimated value of home production. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, reverse-coded worriedness, and mental health. The trust index combines seven measures of trust in different groups and institutions. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights. Outcomes are standardized within each survey wave using the mean and standard deviation of continued-assistance households. The wealth index is constructed from 36 indicators of asset ownership and connectivity using principal component analysis (PCA). It is standardized to have mean zero and standard deviation one in the full beneficiary population, with higher values indicating greater wealth. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values:$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.16. Impact of Assistance Discontinuation: Heterogeneity by Household Size.

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
Discontinued	0.17***	-0.10***	-0.11***	-0.14***
	(0.03)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.000]	[0.000]	[0.000]
Baseline HH Size	0.18***	0.02	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
	[0.000]	[0.267]	[0.267]	[0.267]
Discontinued × HH Size	0.04***	0.02	-0.04**	0.02
	(0.02)	(0.02)	(0.02)	(0.02)
	[0.029]	[0.172]	[0.033]	[0.172]
Transition Package	-0.09	0.09	0.08	0.14**
	(0.06)	(0.06)	(0.06)	(0.07)
	[0.194]	[0.194]	[0.201]	[0.160]
Transition Package × HH Size	0.02	0.04	0.05	-0.07*
	(0.04)	(0.04)	(0.03)	(0.04)
	[0.644]	[0.324]	[0.324]	[0.281]
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,566	5,566	5,566	5,564
Panel B: Endline				
Discontinued	0.15***	-0.08**	-0.12***	-0.06**
	(0.03)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.026]	[0.000]	[0.033]
Baseline HH Size	0.17***	0.01	-0.03***	-0.03**
	(0.01)	(0.01)	(0.01)	(0.01)
	[0.000]	[0.481]	[0.014]	[0.016]
Discontinued × HH Size	0.05***	0.02	-0.01	0.01
	(0.02)	(0.02)	(0.02)	(0.02)
	[0.037]	[0.589]	[0.657]	[0.657]
Transition Package	0.01	0.00	-0.02	0.04
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.989]	[0.989]	[0.989]	[0.989]
Transition Package × HH Size	-0.01	-0.02	-0.01	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)
	[0.799]	[0.799]	[0.799]	[0.799]
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,654	5,654	5,654	5,654

Notes: The sample is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. Food expenditure includes food purchased on credit as well as the estimated value of home production. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, reverse-coded worriedness, and mental health. The trust index combines seven measures of trust in different groups and institutions. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights. Outcomes are standardized within each survey wave using the mean and standard deviation of continued-assistance households. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.17. Impact of Assistance Discontinuation: Heterogeneity by Gender of Household Head.

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
Discontinued	0.20***	-0.08**	-0.13***	-0.14***
	(0.04)	(0.04)	(0.04)	(0.03)
	[0.000]	[0.031]	[0.000]	[0.000]
Female HH Head	-0.24***	-0.08*	-0.03	-0.05
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.000]	[0.155]	[0.554]	[0.303]
Discontinued × Female HH Head	-0.09	-0.07	0.08	-0.02
	(0.06)	(0.06)	(0.06)	(0.06)
	[0.315]	[0.315]	[0.315]	[0.728]
Transition Package	-0.10	0.09	0.11	0.13*
	(0.07)	(0.07)	(0.07)	(0.07)
	[0.208]	[0.208]	[0.208]	[0.208]
Transition Package × Female HH Head	0.11	0.02	-0.10	-0.05
	(0.15)	(0.14)	(0.13)	(0.17)
	[0.891]	[0.891]	[0.891]	[0.891]
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,566	5,566	5,566	5,564
Panel B: Endline				
Discontinued	0.18***	-0.06*	-0.09**	-0.06
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.000]	[0.120]	[0.023]	[0.120]
Female HH Head	-0.23***	-0.12***	-0.03	0.02
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.000]	[0.015]	[0.620]	[0.620]
Discontinued × Female HH Head	-0.11*	-0.04	-0.08	-0.03
	(0.06)	(0.07)	(0.06)	(0.06)
	[0.278]	[0.647]	[0.391]	[0.647]
Transition Package	0.00	-0.04	-0.04	0.02
	(0.05)	(0.05)	(0.05)	(0.05)
	[0.953]	[0.820]	[0.820]	[0.820]
Transition Package × Female HH Head	0.08	0.13	0.05	0.03
	(0.09)	(0.09)	(0.09)	(0.09)
	[0.703]	[0.692]	[0.703]	[0.703]
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,654	5,654	5,654	5,654

Notes: The sample is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. Food expenditure includes food purchased on credit as well as the estimated value of home production. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, reverse-coded worriedness, and mental health. The trust index combines seven measures of trust in different groups and institutions. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights. Outcomes are standardized within each survey wave using the mean and standard deviation of continued-assistance households. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.18. Impact of Assistance Discontinuation on Main Outcomes (Calibrated Population Weights).

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
Discontinued	0.16***	-0.11***	-0.10**	-0.13***
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.000]	[0.007]	[0.010]	[0.002]
Transition Package	-0.02	0.01	0.06	0.18*
	(0.10)	(0.09)	(0.08)	(0.11)
	[0.934]	[0.934]	[0.934]	[0.376]
Discontinued + TP	0.14	-0.10	-0.05	0.05
q-val(Disc+TP=0)	0.537	0.537	0.652	0.652
Mean Continued	0.06	-0.03	0.00	-0.01
Observations	5,564	5,564	5,564	5,562
Panel B: Endline				
Discontinued	0.13***	-0.11**	-0.15***	-0.05
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.005]	[0.021]	[0.003]	[0.278]
Transition Package	0.06	0.04	-0.07	-0.02
	(0.06)	(0.06)	(0.06)	(0.05)
	[0.652]	[0.675]	[0.652]	[0.675]
Discontinued + TP	0.19***	-0.07	-0.22***	-0.07
q-val(Disc+TP=0)	0.002	0.235	0.000	0.235
Mean Continued	0.05	-0.04	0.03	-0.07
Observations	5,654	5,654	5,654	5,654

Notes: The table replicates Table A.13 using wave-specific calibrated population weights. Standard errors account for the sampling weights and survey stratification and are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.19. Impact of Assistance Discontinuation on Secondary Outcomes (Calibrated Population Weights).

	Food Security		Psychological Well-being								
	FCS	rCSI	Life Satisfaction	Worriedness	WHO-5	Non-food exp.	Social cohesion	Employed (%)	Hrs per member	HH has debt	Remittances
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Midline											
Discontinued	-1.75***	0.69	-0.31***	0.46**	-0.97	10.02	-0.04	-0.50	-0.15	0.02	0.00
	(0.66)	(0.48)	(0.12)	(0.23)	(0.79)	(6.70)	(0.04)	(0.95)	(0.42)	(0.02)	(0.01)
	[0.015]	[0.148]	[0.027]	[0.070]	[0.217]	[0.552]	[0.552]	[0.861]	[0.861]	[0.552]	[0.861]
Transition Package	0.09	-0.07	0.05	-0.09	1.75	25.29	0.02	-1.94	1.04	0.02	-0.01
	(1.70)	(0.98)	(0.27)	(0.41)	(1.75)	(17.76)	(0.09)	(2.12)	(1.40)	(0.04)	(0.03)
	[0.956]	[0.956]	[0.841]	[0.841]	[0.841]	[0.842]	[0.842]	[0.842]	[0.842]	[0.842]	[0.842]
Discontinued + TP	-1.66	0.62	-0.26	0.37	0.78	35.31**	-0.02	-2.43	0.89	0.04	-0.01
q-val(Disc+TP=0)	0.531	0.531	0.562	0.562	0.661	0.270	0.785	0.706	0.785	0.706	0.785
Mean Continued	54.70	16.93	11.51	30.14	23.26	251.58	-0.02	23.75	8.85	0.46	0.13
Observations	5,564	5,564	5,457	5,172	5,515	5,564	5,564	5,564	5,564	5,542	5,550
Panel B: Endline											
Discontinued	-0.41	1.46***	-0.52***	0.27	-0.75	3.14	0.07	0.80	0.59	0.01	-0.00
	(0.72)	(0.49)	(0.11)	(0.22)	(0.77)	(8.30)	(0.05)	(1.03)	(0.46)	(0.02)	(0.01)
	[0.574]	[0.006]	[0.000]	[0.330]	[0.330]	[0.846]	[0.595]	[0.846]	[0.595]	[0.846]	[0.977]
Transition Package	0.12	-0.54	-0.07	0.37	-1.09	-5.58	-0.06	-1.91	-1.43***	-0.04	-0.01
	(0.93)	(0.61)	(0.15)	(0.29)	(0.97)	(10.64)	(0.06)	(1.32)	(0.52)	(0.03)	(0.01)
	[0.901]	[0.750]	[0.641]	[0.387]	[0.387]	[0.600]	[0.475]	[0.351]	[0.036]	[0.351]	[0.582]
Discontinued + TP	-0.29	0.93	-0.59***	0.64**	-1.84*	-2.43	0.02	-1.10	-0.84	-0.02	-0.01
q-val(Disc+TP=0)	0.754	0.248	0.000	0.043	0.059	0.819	0.819	0.692	0.619	0.692	0.692
Mean Continued	52.83	15.04	11.43	30.65	21.34	313.53	-0.06	23.91	8.62	0.54	0.06
Observations	5,654	5,654	5,565	5,480	5,643	5,654	5,654	5,653	5,653	5,644	5,652

Notes: The table replicates Table A.10 using wave-specific calibrated population weights. Standard errors account for the sampling weights and survey stratification and are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.20. Impact of Assistance Discontinuation on Main Outcomes (Centered-Strata Interaction Correction).

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
Discontinued	0.16***	-0.10***	-0.11***	-0.14***
	(0.03)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.000]	[0.000]	[0.000]
Transition Package	-0.06	0.11*	0.10	0.12*
	(0.06)	(0.06)	(0.06)	(0.07)
	[0.381]	[0.156]	[0.156]	[0.156]
Discontinued + TP	0.11*	0.00	-0.02	-0.02
q-val(Disc+TP=0)	0.387	0.953	0.953	0.953
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,566	5,566	5,566	5,564
Panel B: Endline				
Discontinued	0.15***	-0.08**	-0.13***	-0.06**
	(0.03)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.021]	[0.000]	[0.048]
Transition Package	0.03	0.00	-0.02	0.03
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.763]	[0.943]	[0.763]	[0.763]
Discontinued + TP	0.17***	-0.08*	-0.15***	-0.03
q-val(Disc+TP=0)	0.000	0.092	0.000	0.507
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,654	5,654	5,654	5,654

Notes: The table replicates Table A.13, allowing the effect of assistance discontinuation to vary across eligibility spaces by interacting the Discontinued indicator with eligibility-space indicators centered at their sample means. With this parametrization, the reported coefficient on Discontinued represents the average discontinuation effect across eligibility spaces, weighted by each space's share of the estimation sample. Eligibility spaces with no variation in discontinuation status are excluded from the estimation. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.21. Impact of Assistance Discontinuation on Secondary Outcomes (Centered-Strata Interaction Correction).

	Food Security		Psychological Well-being								
	FCS	rCSI	Life Satisfaction	Worriedness	WHO-5	Non-food exp.	Social cohesion	Employed (%)	Hrs per member	HH has debt	Remittances
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Midline											
Discontinued	-1.62***	0.72*	-0.37***	0.32*	-1.12**	3.22	0.01	-1.12	-0.38	0.01	-0.01
	(0.48)	(0.37)	(0.09)	(0.17)	(0.57)	(5.02)	(0.03)	(0.72)	(0.33)	(0.01)	(0.01)
	[0.002]	[0.053]	[0.000]	[0.058]	[0.058]	[0.738]	[0.813]	[0.699]	[0.699]	[0.699]	[0.738]
Transition Package	1.07	-1.20	0.37*	-0.19	0.98	19.42*	0.06	-1.48	-0.07	0.01	-0.03*
	(1.11)	(0.82)	(0.19)	(0.33)	(1.20)	(11.03)	(0.07)	(1.53)	(0.78)	(0.03)	(0.02)
	[0.333]	[0.286]	[0.163]	[0.566]	[0.566]	[0.238]	[0.622]	[0.622]	[0.931]	[0.862]	[0.238]
Discontinued + TP	-0.55	-0.48	-0.00	0.13	-0.15	22.65**	0.07	-2.60*	-0.45	0.02	-0.04**
q-val(Disc+TP=0)	0.620	0.620	0.986	0.986	0.986	0.133	0.505	0.179	0.562	0.505	0.133
Mean Continued	54.86	16.57	11.39	29.89	23.05	251.77	0.00	24.01	9.07	0.44	0.15
Observations	5,566	5,566	5,459	5,174	5,517	5,566	5,566	5,566	5,566	5,544	5,552
Panel B: Endline											
Discontinued	-0.08	1.24***	-0.52***	0.03	-0.94*	2.11	0.04	-0.13	0.15	0.01	-0.00
	(0.52)	(0.36)	(0.09)	(0.17)	(0.53)	(6.11)	(0.03)	(0.76)	(0.33)	(0.02)	(0.01)
	[0.883]	[0.001]	[0.000]	[0.871]	[0.118]	[0.862]	[0.862]	[0.862]	[0.862]	[0.862]	[0.862]
Transition Package	-0.25	-0.22	0.06	0.21	-0.63	-3.70	-0.02	-1.09	-0.60	0.00	0.00
	(0.68)	(0.46)	(0.11)	(0.21)	(0.69)	(7.99)	(0.04)	(0.98)	(0.42)	(0.02)	(0.01)
	[0.711]	[0.711]	[0.572]	[0.547]	[0.547]	[0.883]	[0.883]	[0.786]	[0.786]	[0.883]	[0.883]
Discontinued + TP	-0.33	1.02**	-0.46***	0.24	-1.57**	-1.59	0.02	-1.23	-0.45	0.02	0.00
q-val(Disc+TP=0)	0.624	0.041	0.000	0.263	0.036	0.941	0.894	0.804	0.804	0.879	0.941
Mean Continued	53.18	14.66	11.38	30.67	20.78	317.15	0.00	23.43	8.50	0.52	0.06
Observations	5,654	5,654	5,565	5,480	5,643	5,654	5,654	5,653	5,653	5,644	5,652

Notes: The table replicates Table A.10, allowing the effect of assistance discontinuation to vary across eligibility spaces by interacting the Discontinued indicator with eligibility-space indicators centered at their sample means. With this parametrization, the reported coefficient on Discontinued represents the average discontinuation effect across eligibility spaces, weighted by each space's share of the estimation sample. Eligibility spaces with no variation in discontinuation status are excluded from the estimation. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.22. Impact of Local Exposure to Continued Assistance.

	Appeal	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)	(5)
Panel A: Administrative Data, All Discontinued Household					
No. Continued HHs <200m	0.222***				
	(0.082)				
Mean dep. var.	54.44				
Observations	14,946				
Panel B: Midline Survey Data					
Discontinued		0.166***	-0.097***	-0.114***	-0.142***
		(0.039)	(0.037)	(0.038)	(0.039)
		[0.000]	[0.009]	[0.004]	[0.000]
Discontinued x No. Continued HHs <200m		-0.001	-0.002	0.002	-0.001
		(0.002)	(0.002)	(0.002)	(0.002)
		[0.777]	[0.777]	[0.777]	[0.777]
No. Continued HHs <200m		-0.000	0.016***	0.024***	0.009**
		(0.003)	(0.003)	(0.003)	(0.003)
		[0.877]	[0.000]	[0.000]	[0.015]
Mean Continued		-0.02	0.02	-0.01	0.02
Observations		4,819	4,819	4,819	4,817
Panel C: Endline Survey Data					
Discontinued		0.123***	-0.055	-0.117***	-0.120***
		(0.040)	(0.040)	(0.039)	(0.039)
		[0.004]	[0.171]	[0.004]	[0.004]
Discontinued x No. Continued HHs <200m		0.001	-0.002	-0.001	0.004*
		(0.002)	(0.002)	(0.002)	(0.002)
		[0.814]	[0.645]	[0.645]	[0.314]
No. Continued HHs <200m		0.011***	0.021***	0.011***	-0.007*
		(0.003)	(0.003)	(0.003)	(0.004)
		[0.000]	[0.000]	[0.000]	[0.057]
Mean Continued		-0.01	0.01	-0.01	0.01
Observations		4,896	4,896	4,896	4,896

Notes: Panel A uses administrative appeals data and includes all households whose assistance was discontinued. The outcome indicates whether a household submitted an appeal against the discontinuation decision. Panels B and C use the study sample and report midline and endline estimates, respectively, for the four main outcomes. *No. Continued HHs <200m* measures the number of continued-assistance households within 200 meters of each household. All regressions control for the total number of program beneficiary households within the same 200-meter radius and include eligibility-space fixed effects. Panels B and C additionally control for random assignment to the transition package. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.23. Impact of Local Exposure to Continued Assistance among Households with the Same Eligibility Profile.

	Appeal	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)	(5)
Panel A: Administrative Data, All Discontinued Household					
No. Continued same-eligibility HHs <200m	0.279				
	(0.322)				
Mean dep. var.	54.44				
Observations	14,946				
Panel B: Midline Survey Data					
Discontinued		0.155***	-0.093***	-0.085**	-0.158***
		(0.034)	(0.033)	(0.034)	(0.035)
		[0.000]	[0.007]	[0.012]	[0.000]
Discontinued x No. Continued same-eligibility HHs <200m		0.005	-0.017	-0.008	0.006
		(0.011)	(0.014)	(0.014)	(0.014)
		[0.677]	[0.677]	[0.677]	[0.677]
No. Continued same-eligibility HHs <200m		-0.022*	0.035**	0.018	-0.001
		(0.012)	(0.015)	(0.014)	(0.015)
		[0.145]	[0.093]	[0.268]	[0.939]
Mean Continued		-0.02	0.02	-0.01	0.02
Observations		4,819	4,819	4,819	4,817
Panel C: Endline Survey Data					
Discontinued		0.129***	-0.048	-0.114***	-0.096***
		(0.035)	(0.037)	(0.035)	(0.034)
		[0.000]	[0.195]	[0.002]	[0.007]
Discontinued x No. Continued same-eligibility HHs <200m		0.001	-0.029*	-0.018	0.020
		(0.015)	(0.015)	(0.013)	(0.017)
		[0.959]	[0.213]	[0.304]	[0.304]
No. Continued same-eligibility HHs <200m		0.021*	0.031**	0.023*	-0.007
		(0.013)	(0.015)	(0.014)	(0.017)
		[0.130]	[0.130]	[0.130]	[0.668]
Mean Continued		-0.01	0.01	-0.01	0.01
Observations		4,896	4,896	4,896	4,896

Notes: Panel A uses administrative appeals data and includes all households whose assistance was discontinued. The outcome indicates whether a household submitted an appeal against the discontinuation decision. Panels B and C use the study sample and report midline and endline estimates, respectively, for the four main outcomes. *No. Continued same-eligibility HHs <200m* measures the number of continued-assistance households within 200 meters that belong to the same eligibility space as the household. All regressions control for the total number of program beneficiary households in the same eligibility space within the 200-meter radius and include eligibility-space fixed effects. Panels B and C additionally control for random assignment to the transition package. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.24. Impact of Assistance Discontinuation on Main Outcomes, by Priority Status.

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
β_N: Discontinued × Non-Priority	0.132**	-0.075	-0.089	-0.112*
	(0.066)	(0.060)	(0.065)	(0.063)
	[0.153]	[0.211]	[0.211]	[0.153]
β_P: Discontinued × Priority	0.173***	-0.112***	-0.111***	-0.150***
	(0.033)	(0.033)	(0.033)	(0.034)
	[0.000]	[0.000]	[0.000]	[0.000]
p-val($\beta_N = \beta_P$)	0.583	0.593	0.760	0.587
q-val($\beta_N = \beta_P$)	0.760	0.760	0.760	0.760
Observations	5,566	5,566	5,566	5,564
Panel B: Endline				
β_N: Discontinued × Non-Priority	0.221***	-0.082	-0.237***	-0.156**
	(0.070)	(0.063)	(0.072)	(0.070)
	[0.003]	[0.195]	[0.003]	[0.036]
β_P: Discontinued × Priority	0.117***	-0.077**	-0.095***	-0.036
	(0.035)	(0.038)	(0.034)	(0.033)
	[0.003]	[0.055]	[0.011]	[0.279]
p-val($\beta_N = \beta_P$)	0.182	0.946	0.077	0.123
q-val($\beta_N = \beta_P$)	0.243	0.946	0.243	0.243
Observations	5,654	5,654	5,654	5,654
p-value(ΔMidline=ΔEndline)	0.093	0.613	0.069	0.089
q-value(ΔMidline=ΔEndline)	0.125	0.613	0.125	0.125

Notes: The table reports estimated effects of assistance discontinuation on the main outcomes separately for households in non-priority and priority areas, by survey round. Between the two survey rounds, discontinued households residing in priority areas received WFP emergency cash assistance (Figure A.3 shows the geographic distribution of priority status). All regressions include eligibility-space fixed effects, a control for random assignment to the transition package, and their full interactions with an indicator for priority location. The p-value (q-value) row in each panel reports the p-value (FDR-adjusted q-value) from a test of the null hypothesis that the discontinuation effect is equal across non-priority and priority areas. The bottom p-value (q-value) reports the p-value (FDR-adjusted q-value) from a test of the null hypothesis that the difference in discontinuation effects between non-priority and priority areas is the same at midline and endline. This test is based on a regression that includes both survey rounds, with standard errors clustered at the household level. Robust standard errors for the round-specific estimates are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.25. Impact of Assistance Discontinuation on Labor Supply at Midline.

	Adults			Women			Men		
	Employed (%)	Hrs per member	Job quality	Employed (%)	Hrs per member	Job quality	Employed (%)	Hrs per member	Job quality
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Full sample									
Discontinued	-1.07	-0.33	0.011	-0.56	-0.07	0.021	-1.82	-1.09	0.009
	(0.72)	(0.33)	(0.024)	(0.72)	(0.24)	(0.049)	(1.34)	(0.67)	(0.026)
	[0.518]	[0.694]	[0.761]	[0.761]	[0.761]	[0.761]	[0.518]	[0.518]	[0.761]
Transition Package	-1.83	-0.22	0.074	-1.39	-0.42	0.135	-2.49	0.02	0.086
	(1.50)	(0.76)	(0.051)	(1.49)	(0.51)	(0.100)	(2.67)	(1.36)	(0.054)
	[0.498]	[0.866]	[0.498]	[0.527]	[0.529]	[0.498]	[0.527]	[0.991]	[0.498]
Discontinued + TP	-2.91*	-0.55	0.085*	-1.94	-0.49	0.156	-4.32	-1.08	0.095*
q-val(Disc+TP=0)	0.216	0.463	0.216	0.291	0.435	0.216	0.216	0.463	0.216
Mean Continued	24.01	9.07	0.969	10.00	2.97	1.183	43.80	17.46	0.916
Observations	5,566	5,566	3,298	5,424	5,424	920	4,661	4,661	2,800
Panel B: Restricted sample									
Discontinued	-0.63	-0.35	0.010	0.05	-0.02	-0.010	-1.82	-1.03	0.014
	(0.77)	(0.36)	(0.025)	(0.79)	(0.26)	(0.056)	(1.39)	(0.70)	(0.027)
	[0.937]	[0.937]	[0.954]	[0.954]	[0.954]	[0.954]	[0.854]	[0.854]	[0.954]
Transition Package	-3.37**	-0.97	0.093*	-3.11**	-0.99**	0.178	-1.94	-0.06	0.106*
	(1.51)	(0.71)	(0.053)	(1.43)	(0.48)	(0.113)	(2.80)	(1.41)	(0.055)
	[0.115]	[0.220]	[0.144]	[0.115]	[0.115]	[0.171]	[0.550]	[0.966]	[0.118]
Discontinued + TP	-4.00***	-1.33*	0.102*	-3.07**	-1.01**	0.169	-3.76	-1.09	0.120**
q-val(Disc+TP=0)	0.073	0.092	0.092	0.078	0.078	0.169	0.203	0.440	0.078
Mean Continued	26.12	10.11	0.960	9.22	2.81	1.242	44.33	17.76	0.903
Observations	4,363	4,363	2,916	4,331	4,331	703	4,271	4,271	2,610

Notes: The table focuses on midline results. Panel A includes households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. Panel B restricts this sample to households that, at baseline, contained at least one man and one woman aged 15 or older. Outcomes are always defined at the household level. Columns (1)–(3) consider all household members aged 15 or older, columns (4)–(6) restrict to women, and columns (7)–(9) to men. Employed (%) is the share of household members in the relevant group who are employed. Hours per member is total weekly hours worked by members of the relevant group divided by the number of members in that group. The job-quality index ranges from 0 to 2 for each household member aged 15 or older that is employed, assigning one point for salaried employment and one point for working 1–48 hours during the previous week (summed across all jobs). The q-value row reports the FDR-adjusted q-value from a test of the null hypothesis that the sum of the coefficients on Discontinued and Transition Package equals zero. This corresponds to testing whether the effect for transition package households, relative to continued assistance households, is zero. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.26. Baseline Wealth across Continued and Discontinued Households.

	Wealth Index (PCA)	
	(1)	(2)
Discontinued	0.974***	0.896***
	(0.007)	(0.015)
PMT-FCS		-0.139***
		(0.016)
PMT-CARI		-0.134***
		(0.016)
Categorical		0.101***
		(0.017)
Discontinued × PMT-FCS		0.261***
		(0.020)
Discontinued × PMT-CARI		0.243***
		(0.020)
Discontinued × Categorical		-0.193***
		(0.022)
Mean Continued	-0.56	-0.56
Observations	58,501	58,501
R-squared	0.233	0.241

Notes: The sample includes the full 2023 beneficiary population. The outcome is the baseline wealth index, constructed from 36 indicators of asset ownership and connectivity using principal component analysis (PCA). It is standardized to have mean zero and standard deviation one in the full beneficiary population, with higher values indicating greater wealth. Households assigned to the transition package are classified as discontinued. In column (2), PMT-Expenditure is the omitted reference group. Robust standard errors are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.27. Baseline Correlates of Appeal Submission among Discontinued Households.

	Household Appealed		
	(1)	(2)	(3)
Wealth Index (PCA)	-3.20***		
	(0.28)		
Rank distance to cutoff (1,000)		-0.34***	
		(0.03)	
Household head with primary education or lower			0.24
			(0.56)
Household owns a smartphone			2.38**
			(0.99)
Household owns a phone			2.21**
			(0.94)
Observations	32,361	32,361	32,361
R-squared	0.004	0.004	0.000

Notes: The table uses administrative appeals as well as baseline census data and includes all households whose assistance was discontinued. The outcome is an indicator equal to one if the household submitted an appeal against the discontinuation decision. The wealth index is constructed from the baseline census using 36 indicators of asset ownership and connectivity and principal component analysis (PCA). It is standardized to have mean zero and standard deviation one, with higher values indicating greater wealth. Rank distance to the cutoff is calculated under each household?s assigned prioritization method using the full beneficiary population. It is measured in thousands of rank positions, with higher values indicating that a household lies farther below the inclusion cutoff. Robust standard errors are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.28. Impact of Targeting Methods on Secondary Outcomes.

	Food Security		Psychological Well-being								
	FCS	rCSI	Life Satisfaction	Worriedness	WHO-5	Non-food exp.	Social cohesion	Employed (%)	Hrs per member	HH has debt	Remittances
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Midline											
PMT-FCS	0.17	-0.53	-0.02	-0.19	0.37	9.60	0.04	-0.15	0.06	0.01	0.01
	(0.58)	(0.43)	(0.11)	(0.20)	(0.67)	(6.00)	(0.04)	(0.87)	(0.38)	(0.02)	(0.01)
	[0.769]	[0.447]	[0.841]	[0.841]	[0.841]	[0.460]	[0.460]	[0.868]	[0.868]	[0.868]	[0.460]
PMT-CARI	1.08*	0.33	0.12	-0.38*	0.12	4.67	0.01	-0.20	0.06	0.02	0.00
	(0.58)	(0.44)	(0.11)	(0.20)	(0.68)	(5.96)	(0.04)	(0.87)	(0.39)	(0.02)	(0.01)
	[0.123]	[0.447]	[0.396]	[0.173]	[0.860]	[0.869]	[0.869]	[0.869]	[0.869]	[0.869]	[0.869]
Categorical	-0.04	0.04	0.05	-0.08	-0.04	9.22	0.02	-0.09	0.30	0.03	-0.01
	(0.58)	(0.44)	(0.11)	(0.20)	(0.66)	(6.05)	(0.04)	(0.87)	(0.39)	(0.02)	(0.01)
	[0.945]	[0.945]	[0.954]	[0.954]	[0.954]	[0.450]	[0.636]	[0.916]	[0.636]	[0.450]	[0.636]
q-val(PMT-FCS=PMT-CARI)	0.115	0.100	0.508	0.508	0.711	0.814	0.814	0.999	0.999	0.892	0.814
q-val(PMT-FCS=Categorical)	0.719	0.379	0.594	0.594	0.594	0.950	0.846	0.950	0.846	0.846	0.224
q-val(PMT-CARI=Categorical)	0.109	0.512	0.771	0.399	0.814	0.840	0.840	0.901	0.840	0.840	0.840
Mean PMT-Exp.	54.47	16.62	11.26	30.32	23.00	255.57	0.00	24.91	9.34	0.44	0.13
Observations	6,358	6,358	6,240	5,900	6,300	6,358	6,358	6,358	6,358	6,325	6,342
Panel B: Endline											
PMT-FCS	0.24	0.05	0.01	0.23	-0.38	-5.68	-0.00	-1.64**	-1.01***	0.03*	0.01
	(0.58)	(0.39)	(0.10)	(0.19)	(0.59)	(6.76)	(0.03)	(0.83)	(0.35)	(0.02)	(0.01)
	[0.895]	[0.895]	[0.902]	[0.649]	[0.784]	[0.526]	[0.933]	[0.102]	[0.024]	[0.102]	[0.526]
PMT-CARI	0.22	-0.20	0.23**	0.07	0.70	-10.75	0.03	-0.44	-0.50	0.01	-0.00
	(0.57)	(0.39)	(0.10)	(0.18)	(0.60)	(6.77)	(0.03)	(0.84)	(0.36)	(0.02)	(0.01)
	[0.695]	[0.695]	[0.056]	[0.713]	[0.364]	[0.500]	[0.599]	[0.599]	[0.500]	[0.599]	[0.599]
Categorical	0.28	0.59	0.11	0.11	0.05	-4.77	-0.00	-0.12	0.10	0.03*	0.01
	(0.58)	(0.39)	(0.10)	(0.19)	(0.61)	(6.70)	(0.03)	(0.85)	(0.37)	(0.02)	(0.01)
	[0.629]	[0.268]	[0.776]	[0.847]	[0.931]	[0.952]	[0.984]	[0.984]	[0.984]	[0.539]	[0.952]
q-val(PMT-FCS=PMT-CARI)	0.986	0.986	0.078	0.382	0.104	0.454	0.419	0.278	0.278	0.278	0.278
q-val(PMT-FCS=Categorical)	0.942	0.347	0.511	0.511	0.511	0.949	0.949	0.204	0.011	0.949	0.949
q-val(PMT-CARI=Categorical)	0.927	0.095	0.428	0.832	0.428	0.459	0.459	0.707	0.435	0.459	0.435
Mean PMT-Exp.	53.30	14.96	11.09	30.66	20.54	327.94	0.00	25.02	9.34	0.50	0.05
Observations	6,465	6,465	6,363	6,263	6,453	6,465	6,465	6,464	6,464	6,450	6,462

Notes: PMT-Expenditure is the omitted (reference) group. The Food Consumption Score (FCS) ranges from 0 to 112, with higher values indicating higher food consumption. The reduced Coping Strategies Index (rCSI) ranges from 0 to 56, with higher values indicating greater reliance on coping strategies. Life satisfaction is the sum of five items (5-20), with higher values indicating greater satisfaction. Worriedness is the sum of ten items (10-40), with higher values indicating greater worry. The WHO-5 measures mental health on a 0-100 scale, with higher values indicating better well-being. Non-food expenditure is measured in USD. The social cohesion index combines lower community tension, volunteering, and any association membership using inverse-covariance weights applied to components standardized within each survey wave, with higher values indicating greater social cohesion. Employed share is the share of household members aged 15 or older who are employed. Hours worked is the total weekly hours worked by household members aged 15 or older divided by the number of household members in that group. Debt and remittances are indicators for having any outstanding debt and receiving any remittances during the survey reference period, respectively. The q-value rows report FDR-adjusted q-values from tests of the null hypothesis that the coefficients on the indicated prioritization arms are equal. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.29. Impact of Targeting Methods on Primary Outcomes (Calibrated Population Weights).

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
PMT-FCS	0.09*	-0.01	0.01	0.02
	(0.05)	(0.05)	(0.05)	(0.05)
	[0.394]	[0.797]	[0.797]	[0.797]
PMT-CARI	0.06	0.02	-0.02	-0.04
	(0.05)	(0.05)	(0.05)	(0.05)
	[0.709]	[0.709]	[0.709]	[0.709]
Categorical	0.08	-0.03	-0.04	0.05
	(0.05)	(0.05)	(0.05)	(0.05)
	[0.534]	[0.534]	[0.534]	[0.534]
q-val(PMT-FCS=PMT-CARI)	0.680	0.680	0.680	0.680
q-val(PMT-FCS=Categorical)	0.836	0.836	0.836	0.836
q-val(PMT-CARI=Categorical)	0.837	0.649	0.837	0.380
Mean PMT-Exp.	0.04	0.06	0.04	-0.04
Observations	6,356	6,356	6,356	6,354
Panel B: Endline				
PMT-FCS	0.04	0.04	-0.06	-0.04
	(0.05)	(0.05)	(0.05)	(0.05)
	[0.520]	[0.520]	[0.520]	[0.520]
PMT-CARI	-0.03	0.03	-0.00	0.05
	(0.05)	(0.05)	(0.05)	(0.05)
	[0.804]	[0.804]	[0.981]	[0.804]
Categorical	0.01	-0.07	-0.05	-0.03
	(0.05)	(0.05)	(0.05)	(0.05)
	[0.874]	[0.740]	[0.740]	[0.835]
q-val(PMT-FCS=PMT-CARI)	0.339	0.863	0.339	0.325
q-val(PMT-FCS=Categorical)	0.786	0.174	0.786	0.786
q-val(PMT-CARI=Categorical)	0.472	0.278	0.472	0.340
Mean PMT-Exp.	0.05	0.01	0.06	-0.02
Observations	6,465	6,465	6,465	6,465

Notes: The table replicates Table 3 using wave-specific calibrated population weights. Standard errors account for the sampling weights and survey stratification and are reported in parentheses. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.30. Impact of Targeting Methods on the Secondary Outcomes (Calibrated Population Weights).

	Food Security		Psychological Well-being								
	FCS	rCSI	Life Satisfaction	Worriedness	WHO-5	Non-food exp.	Social cohesion	Employed (%)	Hrs per member	HH has debt	Remittances
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Midline											
PMT-FCS	-0.26	0.08	-0.06	-0.37	0.01	3.30	0.01	-1.97	-0.18	0.02	0.02
	(0.91)	(0.61)	(0.16)	(0.31)	(1.02)	(8.90)	(0.06)	(1.33)	(0.55)	(0.03)	(0.02)
	[0.898]	[0.898]	[0.991]	[0.722]	[0.991]	[0.893]	[0.901]	[0.550]	[0.893]	[0.716]	[0.550]
PMT-CARI	1.02	0.41	-0.10	-0.42	-1.37	10.81	0.00	-0.88	0.36	0.02	-0.01
	(0.89)	(0.60)	(0.17)	(0.32)	(1.05)	(9.19)	(0.06)	(1.33)	(0.56)	(0.03)	(0.02)
	[0.491]	[0.491]	[0.562]	[0.291]	[0.291]	[0.675]	[0.967]	[0.675]	[0.675]	[0.675]	[0.675]
Categorical	-0.07	0.53	-0.11	-0.16	-1.23	5.95	-0.05	-1.78	0.33	0.04	-0.01
	(0.89)	(0.62)	(0.16)	(0.32)	(1.01)	(9.02)	(0.06)	(1.34)	(0.57)	(0.03)	(0.02)
	[0.938]	[0.788]	[0.608]	[0.608]	[0.608]	[0.563]	[0.563]	[0.547]	[0.563]	[0.547]	[0.563]
q-val(PMT-FCS=PMT-CARI)	0.317	0.588	0.864	0.864	0.514	0.614	0.931	0.614	0.614	0.931	0.383
q-val(PMT-FCS=Categorical)	0.829	0.829	0.783	0.769	0.588	0.880	0.703	0.880	0.718	0.880	0.225
q-val(PMT-CARI=Categorical)	0.435	0.853	0.954	0.954	0.954	0.896	0.896	0.896	0.961	0.896	0.961
Mean PMT-Exp.	54.85	15.77	11.45	30.44	24.19	259.80	0.03	26.47	9.33	0.43	0.13
Observations	6,356	6,356	6,238	5,898	6,298	6,356	6,356	6,356	6,356	6,323	6,340
Panel B: Endline											
PMT-FCS	0.60	-0.20	-0.16	0.32	-0.14	-14.54	-0.01	-1.87	-1.07*	0.04	-0.01
	(0.90)	(0.61)	(0.15)	(0.29)	(0.95)	(10.16)	(0.05)	(1.32)	(0.59)	(0.03)	(0.01)
	[0.737]	[0.737]	[0.403]	[0.403]	[0.885]	[0.238]	[0.851]	[0.238]	[0.238]	[0.238]	[0.790]
PMT-CARI	0.07	-0.41	0.11	0.26	-0.15	-6.04	0.07	-0.46	-0.76	0.05*	-0.01
	(0.90)	(0.59)	(0.15)	(0.30)	(0.97)	(10.16)	(0.06)	(1.32)	(0.58)	(0.03)	(0.01)
	[0.935]	[0.935]	[0.720]	[0.720]	[0.878]	[0.663]	[0.431]	[0.729]	[0.431]	[0.395]	[0.431]
Categorical	-0.44	0.89	-0.01	0.32	-0.60	-4.57	-0.02	-1.49	-0.50	0.04	0.00
	(0.88)	(0.63)	(0.15)	(0.28)	(0.96)	(10.29)	(0.05)	(1.34)	(0.60)	(0.03)	(0.01)
	[0.619]	[0.315]	[0.933]	[0.799]	[0.799]	[0.808]	[0.808]	[0.805]	[0.808]	[0.805]	[0.808]
q-val(PMT-FCS=PMT-CARI)	0.731	0.731	0.195	0.992	0.992	0.673	0.673	0.673	0.673	0.673	0.673
q-val(PMT-FCS=Categorical)	0.240	0.173	0.870	0.980	0.946	0.947	0.974	0.974	0.947	0.974	0.947
q-val(PMT-CARI=Categorical)	0.564	0.073	0.830	0.830	0.830	0.885	0.516	0.835	0.835	0.835	0.516
Mean PMT-Exp.	53.60	15.00	11.21	30.44	21.42	329.31	0.01	26.48	9.99	0.49	0.06
Observations	6,465	6,465	6,363	6,263	6,453	6,465	6,465	6,464	6,464	6,450	6,462

Notes: The table replicates Table A.28 using wave-specific calibrated population weights. Standard errors account for the sampling weights and survey stratification and are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.31. Impact of Targeting Methods: Heterogeneity by Initial Wealth.

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
PMT-FCS	0.07**	0.03	0.01	-0.03
	(0.03)	(0.03)	(0.04)	(0.04)
	[0.106]	[0.622]	[0.776]	[0.622]
PMT-CARI	0.02	0.02	0.04	-0.04
	(0.03)	(0.04)	(0.03)	(0.04)
	[0.621]	[0.621]	[0.570]	[0.570]
Categorical	0.05	-0.01	0.00	-0.02
	(0.03)	(0.04)	(0.03)	(0.04)
	[0.587]	[0.983]	[0.983]	[0.983]
Wealth Index (PCA)	0.26***	0.16***	0.01	-0.05*
	(0.03)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.000]	[0.751]	[0.123]
PMT-FCS × Wealth Index	-0.08**	-0.07*	-0.05	-0.00
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.122]	[0.122]	[0.352]	[0.948]
PMT-CARI × Wealth Index	0.01	0.01	-0.02	-0.03
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.877]	[0.877]	[0.877]	[0.877]
Categorical × Wealth Index	-0.04	-0.06	-0.03	0.03
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.541]	[0.448]	[0.541]	[0.541]
Mean PMT-Exp.	0.00	0.00	0.00	0.00
Observations	6,358	6,358	6,358	6,356
Panel B: Endline				
PMT-FCS	0.03	0.00	-0.03	-0.04
	(0.03)	(0.04)	(0.04)	(0.03)
	[0.492]	[0.892]	[0.492]	[0.492]
PMT-CARI	-0.01	0.02	0.05	0.08**
	(0.03)	(0.04)	(0.03)	(0.03)
	[0.726]	[0.726]	[0.235]	[0.104]
Categorical	0.06*	-0.03	0.01	-0.00
	(0.03)	(0.04)	(0.03)	(0.03)
	[0.396]	[0.886]	[0.886]	[0.886]
Wealth Index (PCA)	0.20***	0.08**	0.07**	-0.10***
	(0.03)	(0.03)	(0.03)	(0.03)
	[0.000]	[0.013]	[0.029]	[0.000]
PMT-FCS × Wealth Index	-0.09**	-0.08**	-0.01	0.02
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.077]	[0.077]	[0.759]	[0.759]
PMT-CARI × Wealth Index	0.01	-0.04	-0.00	-0.05
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.985]	[0.528]	[0.985]	[0.528]
Categorical × Wealth Index	-0.01	-0.02	-0.01	0.01
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.841]	[0.841]	[0.841]	[0.841]
Mean PMT-Exp.	0.00	0.00	0.00	0.00
Observations	6,465	6,465	6,465	6,465

Notes: PMT-Expenditure is the omitted (reference) group. Food expenditure includes food purchased on credit as well as the estimated value of home production. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, reverse-coded worriedness, and mental health. The trust index combines seven measures of trust in different groups and institutions. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights. Outcomes are standardized within each survey wave using the mean and standard deviation of households assigned to the PMT-Expenditure arm. The wealth index is constructed from 36 indicators of asset ownership and connectivity using principal component analysis (PCA). It is standardized to have mean zero and standard deviation one in the full beneficiary population, with higher values indicating greater wealth. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.32. Impact of Targeting Methods: Heterogeneity by Household Size.

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
PMT-FCS	0.08**	0.04	0.01	-0.03
	(0.03)	(0.03)	(0.04)	(0.04)
	[0.047]	[0.573]	[0.734]	[0.580]
PMT-CARI	0.04	0.03	0.04	-0.04
	(0.03)	(0.04)	(0.03)	(0.04)
	[0.387]	[0.411]	[0.387]	[0.387]
Categorical	0.06*	-0.00	0.00	-0.02
	(0.03)	(0.04)	(0.03)	(0.04)
	[0.259]	[0.974]	[0.974]	[0.974]
Baseline HH Size	0.19***	0.06***	-0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)
	[0.000]	[0.000]	[0.533]	[0.774]
PMT-FCS × HH Size	-0.02	-0.07***	-0.04**	0.01
	(0.02)	(0.02)	(0.02)	(0.02)
	[0.292]	[0.005]	[0.071]	[0.566]
PMT-CARI × HH Size	-0.01	-0.03	0.00	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)
	[0.916]	[0.804]	[0.948]	[0.814]
Categorical × HH Size	-0.01	-0.04**	-0.02	0.01
	(0.02)	(0.02)	(0.02)	(0.02)
	[0.588]	[0.080]	[0.588]	[0.588]
Mean PMT-Exp.	0.00	0.00	0.00	0.00
Observations	6,358	6,358	6,358	6,356
Panel B: Endline				
PMT-FCS	0.04	0.01	-0.03	-0.04
	(0.03)	(0.04)	(0.04)	(0.03)
	[0.517]	[0.841]	[0.517]	[0.517]
PMT-CARI	0.00	0.02	0.06	0.07**
	(0.03)	(0.04)	(0.03)	(0.03)
	[0.929]	[0.719]	[0.214]	[0.167]
Categorical	0.06*	-0.02	0.01	-0.01
	(0.03)	(0.04)	(0.03)	(0.04)
	[0.244]	[0.832]	[0.832]	[0.832]
Baseline HH Size	0.17***	0.04***	-0.03**	-0.06***
	(0.01)	(0.01)	(0.01)	(0.01)
	[0.000]	[0.007]	[0.036]	[0.000]
PMT-FCS × HH Size	0.01	-0.03	0.00	0.06***
	(0.02)	(0.02)	(0.02)	(0.02)
	[0.889]	[0.241]	[0.889]	[0.014]
PMT-CARI × HH Size	-0.01	-0.05**	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.02)
	[0.755]	[0.049]	[0.433]	[0.433]
Categorical × HH Size	0.02	-0.05***	-0.03*	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
	[0.382]	[0.039]	[0.145]	[0.179]
Mean PMT-Exp.	0.00	0.00	0.00	0.00
Observations	6,465	6,465	6,465	6,465

Notes: PMT-Expenditure is the omitted (reference) group. Food expenditure includes food purchased on credit as well as the estimated value of home production. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, reverse-coded worriedness, and mental health. The trust index combines seven measures of trust in different groups and institutions. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights. Outcomes are standardized within each survey wave using the mean and standard deviation of households assigned to the PMT-Expenditure arm. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values:$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.33. Impact of Targeting Methods: Heterogeneity by Gender of Household Head.

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
PMT-FCS	0.07*	-0.01	0.00	-0.01
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.398]	[0.922]	[0.922]	[0.922]
PMT-CARI	0.02	-0.02	0.06	-0.04
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.618]	[0.618]	[0.572]	[0.572]
Categorical	0.08*	-0.04	-0.00	0.01
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.245]	[0.822]	[0.911]	[0.911]
Female HH Head	-0.25***	-0.19***	0.00	-0.03
	(0.05)	(0.05)	(0.05)	(0.06)
	[0.000]	[0.000]	[0.994]	[0.800]
PMT-FCS × Female Head	0.04	0.14**	0.02	-0.06
	(0.07)	(0.07)	(0.07)	(0.08)
	[0.754]	[0.194]	[0.805]	[0.754]
PMT-CARI × Female Head	0.00	0.14*	-0.07	0.01
	(0.07)	(0.07)	(0.07)	(0.08)
	[0.977]	[0.248]	[0.736]	[0.977]
Categorical × Female Head	-0.09	0.09	0.02	-0.09
	(0.07)	(0.07)	(0.07)	(0.08)
	[0.362]	[0.362]	[0.803]	[0.362]
Mean PMT-Exp., Male-headed	0.16	0.06	0.00	0.00
Observations	6,358	6,358	6,358	6,356
Panel B: Endline				
PMT-FCS	0.06	-0.03	-0.08*	-0.03
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.312]	[0.516]	[0.312]	[0.516]
PMT-CARI	-0.02	-0.01	0.02	0.05
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.776]	[0.776]	[0.776]	[0.776]
Categorical	0.11***	-0.05	-0.05	-0.05
	(0.04)	(0.04)	(0.04)	(0.04)
	[0.039]	[0.268]	[0.268]	[0.268]
Female HH Head	-0.20***	-0.17***	-0.16***	-0.04
	(0.05)	(0.05)	(0.05)	(0.05)
	[0.000]	[0.003]	[0.005]	[0.493]
PMT-FCS × Female Head	-0.06	0.10	0.14*	-0.03
	(0.07)	(0.07)	(0.07)	(0.07)
	[0.483]	[0.327]	[0.249]	[0.688]
PMT-CARI × Female Head	0.02	0.09	0.12	0.06
	(0.07)	(0.08)	(0.07)	(0.07)
	[0.765]	[0.461]	[0.449]	[0.500]
Categorical × Female Head	-0.17**	0.07	0.19**	0.13*
	(0.07)	(0.08)	(0.07)	(0.08)
	[0.029]	[0.367]	[0.029]	[0.128]
Mean PMT-Exp., Male-headed	0.13	0.06	0.06	0.00
Observations	6,465	6,465	6,465	6,465

Notes: PMT-Expenditure is the omitted (reference) group. Food expenditure includes food purchased on credit as well as the estimated value of home production. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, reverse-coded worriedness, and mental health. The trust index combines seven measures of trust in different groups and institutions. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights. Outcomes are standardized within each survey wave using the mean and standard deviation of households assigned to the PMT-Expenditure arm. All regressions include eligibility-space fixed effects. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values:$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.34. Impact of Assistance Discontinuation: Heterogeneity by Eligibility.

	Food Exp.	Food Security	Psych. Well-being	Trust
	(1)	(2)	(3)	(4)
Panel A: Midline				
Discontinued (1 method)	0.27***	-0.05	-0.19***	-0.18***
	(0.06)	(0.06)	(0.06)	(0.06)
	[0.000]	[0.375]	[0.002]	[0.002]
+ 2 methods	-0.08	-0.04	0.15**	0.01
	(0.07)	(0.07)	(0.07)	(0.07)
	[0.562]	[0.818]	[0.181]	[0.905]
+ 3 methods	-0.26***	-0.13	0.07	0.12
	(0.08)	(0.08)	(0.08)	(0.09)
	[0.005]	[0.205]	[0.422]	[0.207]
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,566	5,566	5,566	5,564
Panel B: Endline				
Discontinued (1 method)	0.20***	-0.02	-0.17***	-0.02
	(0.06)	(0.06)	(0.06)	(0.06)
	[0.003]	[0.741]	[0.014]	[0.741]
+ 2 methods	-0.08	-0.07	0.07	-0.07
	(0.07)	(0.08)	(0.08)	(0.07)
	[0.366]	[0.366]	[0.366]	[0.366]
+ 3 methods	-0.10	-0.16*	0.02	-0.07
	(0.09)	(0.09)	(0.09)	(0.09)
	[0.507]	[0.348]	[0.832]	[0.595]
Mean Continued	0.00	0.00	0.00	0.00
Observations	5,654	5,654	5,654	5,654

Notes: The sample is restricted to households in the 14 eligibility spaces where the prioritization methods disagree on whether households should continue receiving assistance. The table reports estimated effects of assistance discontinuation separately by the number of prioritization methods under which a household would have been eligible for continued assistance. The coefficient on Discontinued reports the effect for households eligible under one method, while the interaction terms indicate how the discontinuation effect differs for households eligible under two or three methods. The food security index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI). The psychological well-being index combines measures of life satisfaction, reverse-coded worriedness, and mental health. The trust index combines seven measures of trust in different groups and institutions. The food security, psychological well-being, and trust indices are constructed using inverse-covariance weights. Outcomes are standardized within each survey wave using the mean and standard deviation of continued-assistance households. All regressions include eligibility-space fixed effects, an indicator for assignment to the transition package, and its interactions with indicators for eligibility under two or three prioritization methods. Robust standard errors are reported in parentheses. FDR-adjusted q-values are reported in square brackets. Adjustments are calculated separately for each survey wave and treatment comparison. Significance stars are based on unadjusted p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.35. Arab Barometer Wave VII.

	Accepts foreign aid	Accepts UN role: social/economic	Accepts UN role: governance/security
	(1)	(2)	(3)
Panel A: pooled sample			
UN favorable	1.86***	3.17***	3.17***
	(0.36)	(0.47)	(0.52)
CSO trust	1.04***	1.97***	1.80***
	(0.33)	(0.45)	(0.49)
Mean	97.6%	95.8%	94.9%
Observations	9,887	9,866	9,723
Panel B: Lebanon sample			
UN favorable	3.18***	4.52***	3.96***
	(0.96)	(1.05)	(1.20)
CSO trust	-0.14	0.95	2.42**
	(0.81)	(0.70)	(0.95)
Mean	96.6%	96.0%	95.1%
Observations	1,183	1,175	1,153

Notes: The table reports associations between attitudes toward the United Nations (UN), trust in civil society organizations (CSOs), and willingness to accept foreign aid or UN involvement, using data from the Arab Barometer Wave VII. Panel A pools respondents across all participating countries, while Panel B includes Lebanon only. *UN favorable* equals one if the respondent reports a very or somewhat favorable opinion of the United Nations, and zero if they report a somewhat or very unfavorable opinion. *CSO trust* equals one if the respondent reports a great deal or quite a lot of trust in civil society organizations, and zero if they report not a lot of trust or no trust at all. *Accepts foreign aid* equals one if, when asked which area foreign aid to their country should prioritize, the respondent selects any area, and zero if they state that they do not want foreign aid. *Accepts UN role: social/economic* equals one if, when asked what the UN's first priority in the region should be, the respondent selects any of the listed social or economic issues (the specific list might vary across countries). It equals zero if the respondent states that they do not want the UN to work on any of these issues. *Accepts UN role: governance/security* equals one if, when asked what the UN's first priority in the region should be, the respondent selects one of the listed government or security issues (the specific list might vary across countries). It equals zero if the respondent states that they do not want the UN to work on any of these issues. Don't-know responses and refusals are excluded. All regressions include controls for age, gender, education, employment, settlement type, household size, and income hardship. Panel A includes country fixed effects; Panel B includes governorate fixed effects. Regressions use survey weights, as reported in the Arab Barometer. Robust standard errors are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix B: Definition of the Four Targeting Methods

This appendix provides additional details on the construction and performance of the prioritization methods. Section B.1 describes the four targeting models and defines the vulnerability outcome underlying each of them. Section B.2 details the estimation procedure used to construct the three proxy means tests. Section B.3 then evaluates the performance of the resulting prioritization rules.

B.1 The Four Prioritization Models

The four prioritization rules draw on targeting approaches commonly used by aid and humanitarian organizations worldwide. They reflect distinct approaches (data- vs. human- driven) and notions of need (low economic resources vs. food deprivation vs. multidimensional vulnerability). This section describes each rule in detail.

B.1.1 PMT-Expenditure

The first model follows a proxy means test (PMT) approach and targets per capita consumption expenditure, which is a commonly used proxy for poverty. Like all of our PMT-based models, it is estimated using the Lebanon Household Survey (LHS), which collects very detailed information on household consumption and expenditure across a range of food and non-food items. Table B.1 summarizes the construction of the target variable, including its components and valuation rules.

The food module records consumption of 128 individual products over a seven-day recall period. For each product, households report the quantity consumed, in units appropriate for that product, as well as the source of that consumption (market purchases, home production, or transfers). Quantities are valued using the household's own reported unit value rather than prices from an external source. As a result, food obtained outside the market enters consumption expenditure on the same basis as purchased food. For example, suppose a household reports consuming three kilograms of tomatoes during the previous week, one kilogram of which is home-grown, and reports that its most recent purchase was two kilograms for 300,000 Lebanese pounds. The implied unit value of 150,000 pounds per kilogram is applied to all three kilograms consumed. This means that the home-grown tomatoes enter consumption expenditure at the same value as the purchased tomatoes. Food received as transfers (gifts, in-kind wages, or assistance) is recorded separately and is valued in the same way.

When households report receiving food assistance, the survey records both its amount and the institution providing it. We use this information to subtract the value of food received from humanitarian sources to construct a final measure of consumption expenditure net of humanitarian assistance. Across the LHS sample, 13.8% of the Lebanese households report receiving humanitarian assistance in the month before the interview (10.4% cash, 13.8% including in-kind).

Non-food consumption is recorded using three different recall periods, depending on the expected frequency of purchase: 30 days for frequently purchased items, such as soap; three months for less frequent items, such as school uniforms; and 12 months for infrequently purchased items, such as refrigerators.

All expenditure amounts are converted to US dollars and monthly frequency. The final outcome

is the natural logarithm of monthly per capita consumption expenditure, net of humanitarian assistance.

Table B.1. PMT-Expenditure: Construction of the Target Variable

Outcome	Natural logarithm of monthly per capita consumption expenditure, net of humanitarian assistance, built along the lines recommended for welfare analysis by Deaton and Zaidi (2002) and Mancini and Vecchi (2022)	
Construction	*Component*	*Definition*
	Food, purchased	Seven-day recall over 128 individual products. Quantity is reported in product-specific units.
	Food, own production	Same module. It is valued at the household's own unit value.
	Food, in kind	Same module. It also records the specific source and allows identify food received from a humanitarian source.
	Non-food, 30 days	Frequently purchased items.
	Non-food, 3 months	Less frequently purchased items.
	Non-food, 12 months	Lumpy and infrequent purchases.
	Health, education, communication expenses	Collected in their own specific modules.
	Dwelling utilities	Electricity, water and other utilities of the dwelling. Collected in their own specific modules.
	Aggregation	All components are placed on a common 30 days basis, summed, and divided by household size
Weights	None. Components are summed at their monetary value. Transactions reported in Lebanese pounds are converted to US dollars.	

B.1.2 PMT-FCS

The second model also follows a proxy means test (PMT) approach and targets the Food Consumption Score (FCS). FCS is a widely used measure of household food security that captures dietary diversity and consumption frequency. Households report the number of days, over the preceding seven days, on which they consumed foods from each of eight food groups. These frequencies are multiplied by group-specific weights reflecting relative nutritional value and then summed to obtain a score ranging from 0 to 112. The FCS is constructed following the same procedure described in the main text, and reported in Table B.2.

Table B.2. PMT-FCS: Construction of the Target Variable

Outcome	Food Consumption Score, constructed following WFP standard guidelines (WFP, 2021)	
Construction	*Component*	*Definition*
	Staples	Cereals, bread, roots and tubers. Weight: 2
	Pulses	Beans, peas, nuts and seeds. Weight: 3
	Meat, fish and eggs	Weight: 4
	Milk and dairy	Weight: 4
	Vegetables	Weight: 1
	Fruit	Weight: 1
	Oils and fats	Weight: 0.5
	Sugar	Sugar and confectionery. Weight: 0.5
Weights	Each group enters as the number of days in the preceding seven on which the household consumed anything from it, multiplied by the group's weight. The weights are intended to reflect the nutritional density of each group. Since each group can contribute at most seven days, the attainable range of the score is 0 to $7 \times 16 = 112$	

B.1.3 PMT-CARI

The third model also follows a proxy means test (PMT) approach and targets a binary indicator based on the Consolidated Approach for Reporting Indicators of Food Security (CARI). CARI is a multidimensional measure of vulnerability that combines information on economic capacity, food consumption, and coping capacity (WFP, 2021). The first two components are measured using the same outcomes targeted by the PMT models described above (per capita expenditure and the Food Consumption Score), while coping capacity is captured by the reduced Coping Strategies Index (rCSI) and the Livelihood Coping Strategies (LCS).

These components are combined into a four-point CARI index, indicating increasing vulnerability. Households with a rounded CARI score above 2 are classified as vulnerable. This binary indicator of vulnerability is the outcome targeted by this PMT model. Table B.3 reports the components and aggregation rules used to construct the index.

Table B.3. PMT-CARI: Construction of the CARI Score

Outcome	Indicator equal to one if the rounded CARI score is 3 or 4, corresponding to moderate or severe food insecurity, and zero otherwise. The score follows WFP's CARI methodology (WFP, 2021).	
Construction	*Component*	*Definition*
	Per capita expenditure	Same as reported in Table B.1
	FCS	Same as reported in Table B.2
	Reduced coping strategies index (rCSI)	As described in the main text, rCSI is the weighted count of short-term food strategies used in the past seven days: lower-quality food (1), borrowed food (2), fewer meals (1), smaller portions (1), adults skipping meals (3)
	Livelihood Coping Strategies (LCS)	Measured using a set of questions on livelihood coping strategies adopted in response to insufficient resources. Strategies are classified as stress (2), crisis (3), or emergency (4). Each household is assigned the score of the most severe strategy it reports, or 1 if none is reported. Examples include selling household goods (stress), productive assets (crisis), or a house or land (emergency).
Weights	Each component is classified on a four-point severity scale, following WFP methodology (WFP, 2021). The scores are then averaged and rounded to obtain the final CARI score.	

B.1.4 Expert (Categorical) score

The final method follows an expert-based categorical approach. In November 2023, WFP administered a survey to 28 experts with operational or research experience in food security and humanitarian assistance in Lebanon.[1]

Experts were presented with the same set of household characteristics available to the PMT models and asked to select and rank, in descending order of importance, up to ten characteristics they considered most relevant for prioritizing food assistance among Lebanese households. Each characteristic was formulated as either an 'include if" or an 'exclude if" criterion. Twenty-two of the 28 experts ranked ten criteria, while the remaining six ranked between one and nine, for an average of 9.3 criteria per expert.

We aggregate these rankings using a Borda-count procedure. A criterion receives ten points when ranked first, nine when ranked second, and so forth, down to one point when ranked tenth; criteria not selected by an expert receive zero points. Points are then summed across the 28 experts, producing a weight for each criterion that reflects both the frequency with which it was selected and its position in experts' rankings.

For each household, the resulting vulnerability score is calculated as the sum of the weights associated with the inclusion criteria it satisfies, minus the corresponding sum for the exclusion criteria it satisfies. Higher scores therefore indicate greater vulnerability according to the aggregated expert assessment. Table B.4 reports the criteria receiving the largest weights.

[1]The survey was administered online, with a median completion time of approximately nine minutes. Twenty-two respondents completed the survey in English and six in Arabic. Seventeen respondents worked for non-governmental organizations, six for United Nations agencies, three for government bodies, and one in research. One respondent declined to report an affiliation. Twenty-three respondents were based in Lebanon at the time of the survey. Twenty respondents were women and eight were men. Seven were under 30 years of age, twelve were aged 30–39, five 40–49, three 50–59, and one 60–69.

Table B.4. Expert (Categorical) Score: Most Frequent Criteria

Characteristic	Weight
Inclusion criteria	
Household has no income	27.0
Widowed/divorced/single (monoparental) head	23.4
Household head is a minor	17.0
Head has a physical disability	15.8
No working-age member (15–64)	15.2
Household head is elderly	15.0
Household head is female	15.0
Any household member has a disability	15.0
Household head has any disability	10.6
Any member has a chronic disease	7.4
Exclusion criteria	
Owns a car/truck/bus	−27.0
Has a dryer	−16.0
Has a washing machine	−14.0
Main income from a large business (>5 workers)	−14.0
Has a microwave	−11.0
Has an air conditioner	−8.0
Any child in private school	−7.0
Has a dishwasher	−7.0
Has a heating system	−6.0
Main income from unskilled regular labor	−5.0

B.2 PMT Estimation and Model Selection

Implementation of the categorical approach is mechanical, once the expert weights have been constructed. The PMT-based approaches instead require defining a common estimation sample, fitting alternative prediction models, and selecting the specification with the strongest out-of-sample performance.

B.2.1 Training sample

As described above, the three PMTs are estimated using data from the Lebanese Household Survey (LHS), a nationally representative survey of 3,597 households conducted between January and May 2023.

We begin with the 3,511 households with non-missing information on the core outcomes used across the three PMTs. This forms the common base sample for all models. To better align the estimation sample with the population targeted by the assistance program and to limit the influence of high-expenditure outliers, we exclude households in the top quartile of the per capita food expenditure distribution. This removes 877 households and leaves a final estimation sample of 2,634 households.[2]

We randomly assign 75% of this sample (1,976 households) to the training sample and reserve the

[2]Within this restricted sample, 43% of households have per capita expenditure below the Minimum Expenditure Basket. This share is very similar to that observed in WFP's post-distribution monitoring survey, conducted as part of regular operations during the first half of 2023 among a representative sample of program recipients.

remaining 25% (658 households) as a held-out evaluation sample. The same split is used for all three outcomes and all candidate estimators. Model fitting and tuning-parameter selection are conducted exclusively within the training sample, while predictive performance is assessed on the evaluation sample.

B.2.2 Predictor construction and model selection

We harmonize the LHS and the WFP beneficiary census, retaining only variables observed in both datasets. We standardize continuous variables, convert categorical variables into indicators, and add squared and interaction terms. This procedure yields 192 candidate predictors spanning demographic, socioeconomic, housing, service-access, and asset characteristics, together with governorate fixed effects.

For each outcome, we estimate five alternative prediction models. We begin with forward stepwise regression, which sequentially adds predictors according to improvements in model fit. While widely used, stepwise selection can overfit and may be sensitive to the sequence in which predictors enter the model (Harrell, 2015; Smith, 2018). We therefore also consider three regularized alternatives: Ridge, the Least Absolute Shrinkage and Selection Operator (LASSO), and Elastic Net. Finally, we also fit a generalized random forest (GRF), which provides a flexible nonlinear benchmark (Athey et al., 2019).

For log per capita expenditure and Food Consumption Score, the regression-based estimators use linear specifications. For the binary CARI outcome, the regularized estimators use logistic specifications, while the stepwise model is estimated as a linear probability model. Governorate fixed effects are left unpenalized in the regularized specifications, so that penalization applies only to household characteristics. Penalty parameters are selected by cross-validation within the training sample to minimize cross-validated prediction error.

We evaluate all five estimators on the common held-out evaluation sample. Because households in this sample are excluded from model estimation, this exercise approximates the practical targeting problem of predicting welfare for households whose outcomes are not observed. Table B.5 reports mean squared error (MSE), together with R^2 for the continuous outcomes and McFadden's D^2 for CARI. Lower MSE indicates more accurate predictions. Higher R^2 indicates that the model explains a larger share of the variation in the outcome, while higher D^2 indicates a greater improvement in fit relative to an intercept-only model.

The regularized estimators perform similarly and provide the strongest predictive fit for two of the three outcomes. For log per capita expenditure, LASSO and Elastic Net attain an out-of-sample R^2 of 0.43, compared with 0.42 for Ridge and stepwise OLS and 0.37 for the GRF. For CARI, the regularized logit models achieve $D^2 = 0.17$, compared with 0.15 for the GRF and 0.12 for the stepwise linear probability model. For the Food Consumption Score, stepwise OLS performs best, with $R^2 = 0.15$, followed closely by LASSO and Elastic Net at $R^2 = 0.14$; the GRF yields $R^2 = 0.12$. Thus, all five estimators explain less than one-sixth of the out-of-sample variation in the FCS.

The GRF does not improve predictive performance relative to the regression-based approaches, consistent with the evidence in Aiken et al. (2025). One possible explanation is that the squared terms and interactions among the candidate predictors capture some nonlinearities relevant for prediction. The different approaches differ more substantially in parsimony. LASSO retains 126 predictors for expenditure and 71 for each of the two food-security outcomes, compared with 40,

25, and 29 predictors, respectively, under stepwise selection. Ridge retains non-zero coefficients on the full set of predictors, while the GRF is estimated using the full predictor set.

Our ultimate objective, however, is not to predict welfare per se but to rank households for program prioritization. We therefore also assess how differences in predictive performance translate into targeting accuracy. In the beneficiary census, each fitted model is used to rank households by predicted need, and the 43% ranked as most vulnerable (with the lowest predicted expenditure or FCS, or the highest predicted vulnerability under CARI) are selected for continued assistance. We replicate this exercise in the evaluation sample and compare the selected households with those classified as eligible using their observed outcomes.

Table B.6 reports precision, defined as the share of households selected by a model that also belong to the eligible group according to observed welfare. Precision ranges from 0.71 to 0.73 for expenditure, from 0.57 to 0.59 for the FCS, and from 0.64 to 0.65 for CARI, compared with an eligibility share of approximately 0.43 for each outcome. Differences across estimators are therefore modest. Notably, despite its weaker overall predictive fit, the GRF is among the strongest performers in terms of precision for two of the three outcomes.[3]

Taken together, these results motivate our choice of LASSO for the construction of the final PMT scores. LASSO delivers the strongest predictive performance for expenditure and CARI, jointly with other regularized estimators, and performs only slightly below the best-performing model for the FCS. Differences in targeting precision across estimators are small, with LASSO performing at or near the top across outcomes. At the same time, LASSO yields substantially more parsimonious specifications than Ridge or the GRF, retaining 126 predictors for expenditure and 71 for the other two outcomes. We therefore use the cross-validated LASSO predictions to rank households in the beneficiary census under each of the three PMT-based prioritization rules. This choice is also consistent with Altındağ et al. (2021), who select a cross-validated LASSO, benchmarked against Ridge, Elastic Net, and a random forest, to predict per capita expenditure for targeting Syrian refugees in Lebanon.

Table B.5. Out-of-sample Model Comparison.

	Expenditure			FCS			CARI > 2		
Method	MSE	R^2	#cov	MSE	R^2	#cov	MSE	D^2	#cov
Stepwise OLS	0.332	0.415	40	592.7	0.148	25	0.198	0.124	29
Ridge	0.328	0.421	192	604.9	0.131	192	0.193	0.168	192
Elastic Net	0.325	0.426	128	601.9	0.135	73	0.191	0.175	78
LASSO	0.325	0.427	126	601.6	0.136	71	0.191	0.175	71
GRF	0.360	0.366	192	614.4	0.117	192	0.198	0.147	192

Notes: All models use the same random 75% training sample and held-out 25% evaluation sample. Outcomes are per capita expenditure, FCS, and binary CARI, as defined above. MSE is mean squared error (the Brier score for CARI). Goodness of fit is measured by R^2 for continuous outcomes and McFadden's D^2 for CARI. # cov. is the number of predictors retained.

[3]With a fixed coverage rate, targeting depends primarily on households' relative positions around the selection cutoff. Improvements in overall predictive accuracy therefore need not translate into improvements in targeting precision.

Table B.6. Precision, by Estimator

Method	Expenditure	FCS	CARI > 2
Stepwise OLS	0.706	0.587	0.640
Ridge	0.724	0.576	0.643
Elastic Net	0.731	0.572	0.650
LASSO	0.731	0.572	0.647
GRF	0.713	0.587	0.654
Share eligible in test sample	0.430	0.436	0.424

Notes: Each estimator selects the 43% ranked as most in need in the held-out sample. Precision is the share selected who are eligible: the bottom 43% of observed per capita expenditure or FCS, or CARI above 2. The final row reports the eligible share, equal to the expected precision of random selection.

To assess which characteristics contribute most to each PMT score, we rank the retained predictors by the absolute value of their estimated coefficient multiplied by the predictor's standard deviation. We then normalize these values to sum to 100% within each model. This measure places coefficients expressed in different units on a common scale and provides an indication of each predictor's relative contribution to the resulting score.

Table B.7 reports the ten predictors receiving the greatest weight in each PMT. Their composition differs substantially across models. In PMT-Exp, durable assets and demographic characteristics together account for 57% of total predictor importance, with household size entering both directly and through indicators for its quartiles. By contrast, housing conditions and access to services account for roughly half of total importance in both PMT-FCS and PMT-CARI. Demographic characteristics contribute least to PMT-FCS, at 11%, while governorate fixed effects account for 12% of total importance (their largest contribution across the three models). The expert-based categorical approach described above places yet another emphasis: its highest-weighted inclusion criteria primarily capture demographic vulnerability, while its strongest exclusion criteria are largely related to asset ownership (Table B.4).

There is also limited overlap in the individual characteristics receiving the greatest weight across PMT models. Of the 30 entries across the three top-ten lists, 24 are distinct, and only two appear in all three models: the number of luxury assets owned and one governorate fixed effect. Thus, despite being estimated on the same households and drawing from the same set of candidate predictors, the three PMTs place the greatest weight on different household characteristics depending on the definition of vulnerability they are designed to predict.

Table B.7. Highest Weight Predictors, by PMT Model

Predictor	Type	Share of importance (%)
PMT-Exp		
Luxury assets owned (count)	Assets	10.0
Household size	Demographic	3.7
Household size, quartile	Demographic	3.1
Owns an air conditioner	Assets	2.8
Electricity from a neighbour's grid	Socioeconomic	2.6
Owns a microwave	Assets	2.3
Any child in private school	Socioeconomic	2.1
Owns an independent electric oven	Assets	2.0
Governorate: Bekaa	Geography	1.9
Any school-age child not enrolled	Demographic	1.9
Ten most important, combined		*32.4*
PMT-FCS		
Luxury assets owned (count)	Assets	5.4
Has WiFi internet	Assets	4.2
Governorate: Mount Lebanon	Geography	4.1
Flush toilet	Socioeconomic	3.8
Share of school-age children enrolled	Demographic	3.2
Owns solar panels	Socioeconomic	3.0
Concrete floor	Socioeconomic	3.0
Electricity from a private grid	Socioeconomic	2.9
Governorate: Bekaa	Geography	2.9
Head has tertiary education	Socioeconomic	2.9
Ten most important, combined		*35.3*
PMT-CARI		
Luxury assets owned (count)	Assets	8.7
Household size quartile, squared	Demographic	5.4
Uses a private generator	Socioeconomic	3.8
Governorate: Bekaa	Geography	3.5
Any school-age child not enrolled	Demographic	2.8
Owns a water heating system	Assets	2.7
Apartment in a residential complex	Socioeconomic	2.6
Household size, squared	Demographic	2.4
Electricity from a private grid	Socioeconomic	2.4
Does not own the dwelling	Socioeconomic	2.3
Ten most important, combined		*36.4*

Notes: Predictor importance is defined as the absolute value of the penalized LASSO coefficient multiplied by the predictor's standard deviation, normalized to sum to 100% within each model. The reported models are the same specifications used in Table B.5. LASSO retains 126 predictors for PMT-Exp and 71 for each of PMT-FCS and PMT-CARI.

B.3 Validating the Targeting Scores

We assess how well the four prioritization rules identify vulnerable households using the observed welfare outcomes available in the LHS. Because the rules are designed to capture different dimensions of need, we evaluate them in two ways. First, for the three PMT-based rules, which are explicitly trained to predict a specific welfare outcome, we assess each model against its intended outcome. The expert-based categorical rule does not target a single observed welfare measure and is therefore not included in this first exercise. Second, we compare all four rules against common

welfare benchmarks, allowing us to assess how differently they identify vulnerability when evaluated against the same outcomes.

For the first exercise, we apply each PMT to the held-out 25% LHS evaluation sample and replicate the prioritization procedure within this sample. Specifically, we select the 43% of households ranked as most in need by each model, matching the program's coverage rate, and compare this classification with eligibility based on the corresponding observed welfare outcome from LHS data. We report standard measures of targeting performance: precision, recall, inclusion and exclusion errors, and the area under the receiver operating characteristic curve (AUC). Precision measures the share of selected households that are eligible according to observed welfare, while recall measures the share of eligible households that are selected. Inclusion and exclusion errors are their respective complements. AUC measures how well the model ranks eligible households ahead of ineligible households across the full ranking, without imposing a particular selection cutoff.

Table B.8 reports these measures and includes random targeting as a benchmark. PMT-Exp achieves the highest precision (74%) and AUC (0.83), followed by PMT-CARI (64% and 0.78), and PMT-FCS (58% and 0.71). All three rules outperform random targeting, which yields precision of 46% and an AUC of 0.50. These comparisons indicate how well each method identifies households according to its intended targeting objective. Differences across methods, however, also reflect differences in how predictable the underlying welfare outcomes are.

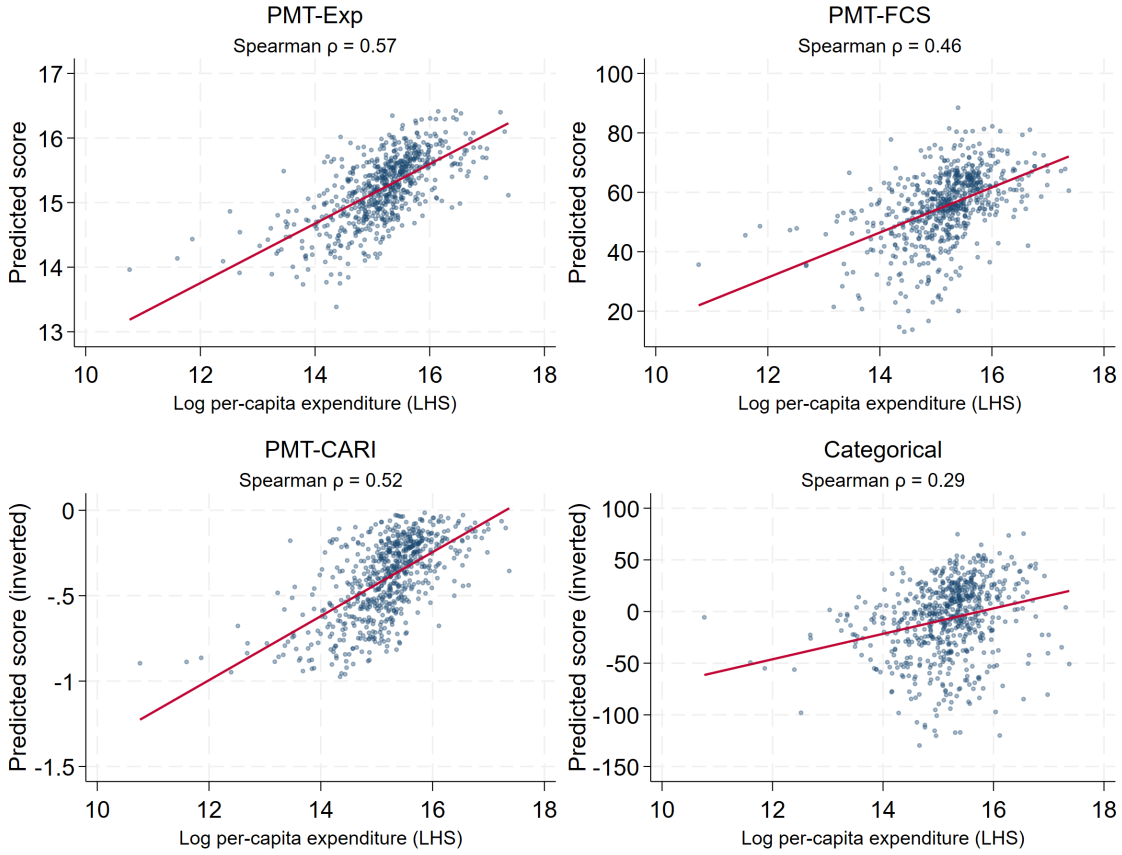
Table B.8. Targeting accuracy: each method against the outcome it predicts

Method	Spearman ρ	AUC	Precision	Recall	Incl. err.	Excl. err.
PMT-Exp	0.572	0.831	0.739	0.723	0.261	0.277
	[0.516, 0.621]	[0.799, 0.860]	[0.687, 0.785]	[0.669, 0.769]		
PMT-FCS	0.353	0.706	0.581	0.621	0.419	0.379
	[0.277, 0.418]	[0.662, 0.743]	[0.524, 0.637]	[0.556, 0.675]		
PMT-CARI	0.471	0.775	0.636	0.696	0.364	0.304
	[0.406, 0.531]	[0.738, 0.809]	[0.583, 0.687]	[0.641, 0.747]		
Random	0.000	0.500	0.461	0.431	0.539	0.569
	[-0.079, 0.075]	[0.454, 0.543]	[0.400, 0.520]	[0.387, 0.471]		

Notes: Accuracy on the 655-household held-out fold of the Lebanese Household Survey, each method evaluated against the outcome it targets (PMT-Exp: expenditure; PMT-FCS: the FCS; PMT-CARI: food insecurity). Precision, recall and the error rates are computed at the 43% eligibility cut-off; Spearman's ρ and the AUC use the full distribution. Inclusion error = 1−precision; exclusion error = 1−recall. Coverage and the deprivation rate are both set at 43%, but tie-breaking at the percentile cut and item non-response leave them slightly apart, so the two errors do not coincide exactly. Random is the no-information benchmark. Brackets: 95% bootstrap confidence intervals (1000 replications).

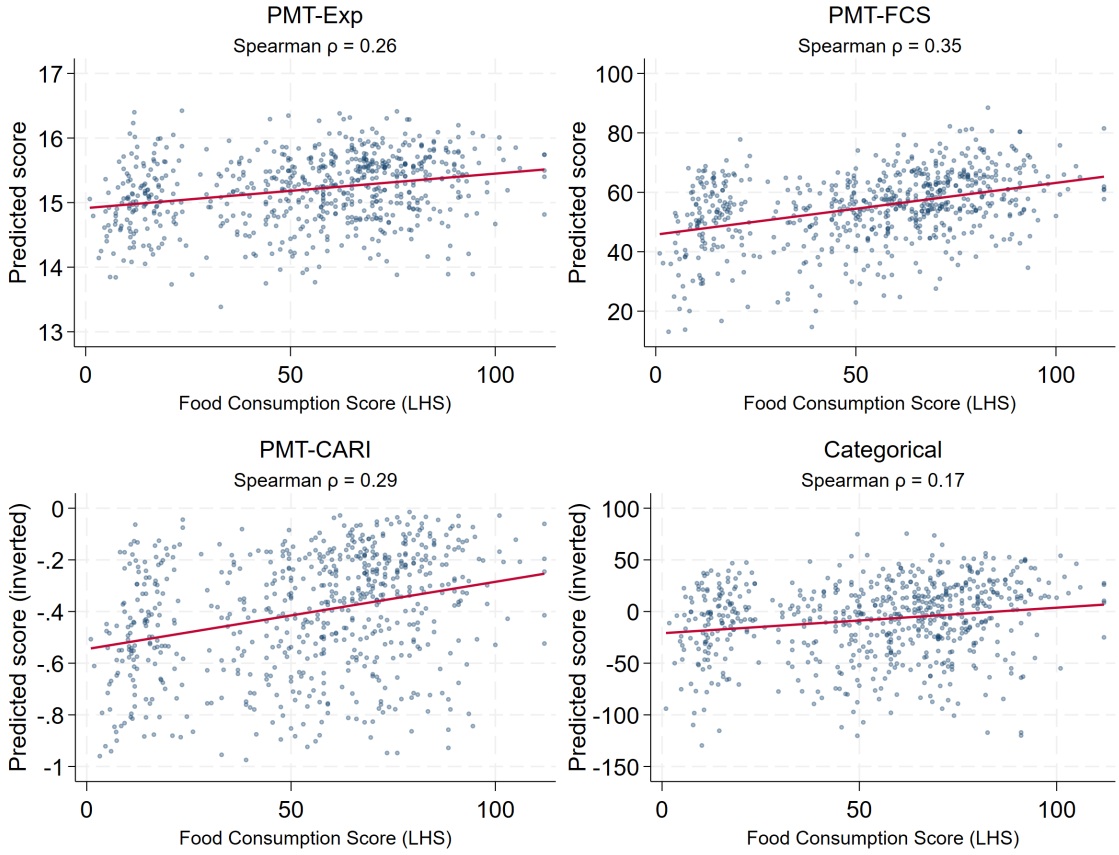
For the second exercise, we evaluate all four prioritization rules against three common welfare benchmarks. We consider per capita expenditure and the Food Consumption Score (FCS), which coincide with the outcomes targeted by PMT-Exp and PMT-FCS, respectively, but are also widely used measures of poverty and food security. As an additional benchmark, we use a household wealth index, which captures a distinct dimension of economic well-being and is not itself the target of any of the prioritization rules. Evaluating all four rules against the same outcomes allows us to compare how closely they align with alternative dimensions of household welfare. Figure B.1 compares each rule with observed per capita expenditure, Figure B.2 with observed FCS, and Figure B.3 with the household wealth index. All comparisons always use the held-out evaluation sample.

Figure B.1. Predicted Prioritization Score and Observed Expenditure, by Method



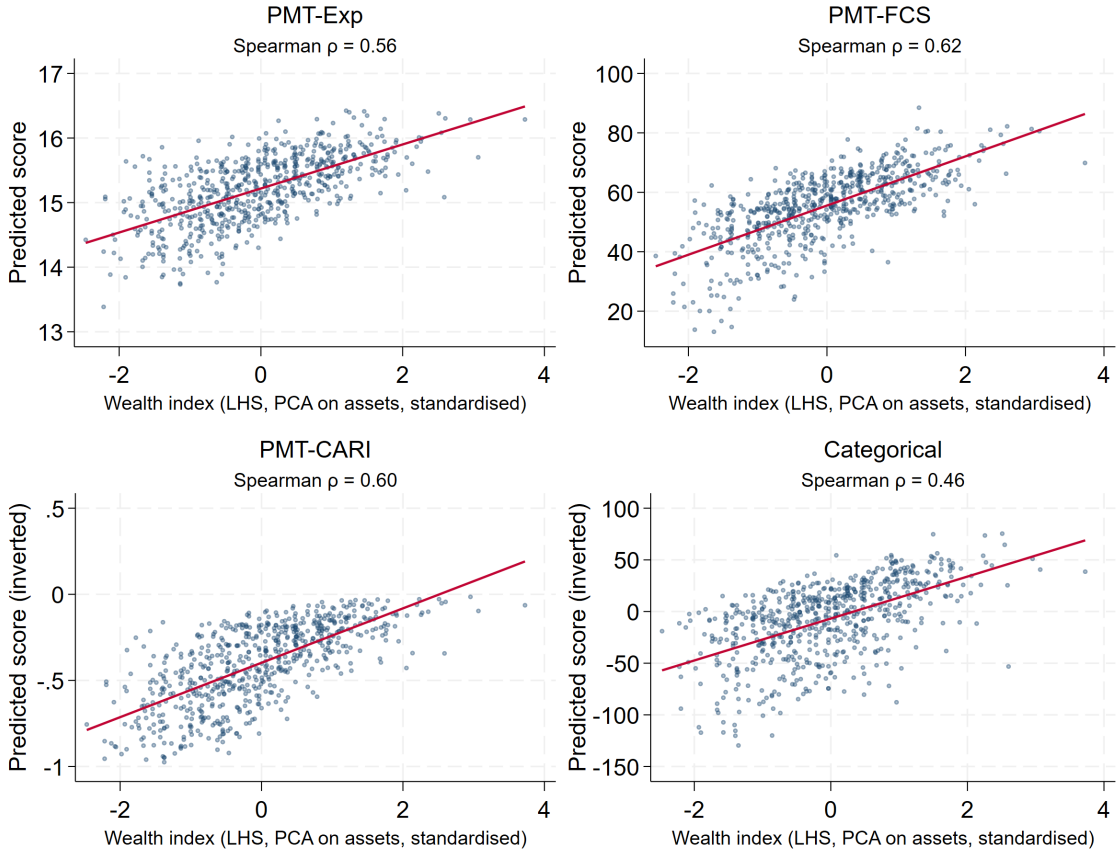
Notes: Each panel plots a method's predicted prioritization score against observed log per capita expenditure in the held-out 25% evaluation sample of the Lebanese Household Survey (LHS) data. Spearman's ρ is reported in each panel. For PMT-CARI and the expert-based categorical method, higher scores indicate greater vulnerability. We therefore reverse the sign of these scores in the figure so that, for all methods, higher values correspond to higher welfare.

Figure B.2. Predicted Prioritization Score and Observed Food Consumption Score, by Method



Notes: Each panel plots a method's predicted prioritization score against observed Food Consumption Score in the held-out 25% evaluation sample of the Lebanese Household Survey (LHS) data. Spearman's ρ is reported in each panel. For PMT-CARI and the expert-based categorical method, higher scores indicate greater vulnerability. We therefore reverse the sign of these scores in the figure so that, for all methods, higher values correspond to higher welfare.

Figure B.3. Predicted Prioritization Score and Observed Wealth Index, by Method



Notes: Each panel plots a method's predicted prioritization score against observed wealth index in the held-out 25% evaluation sample of the Lebanese Household Survey (LHS) data. The wealth index is constructed from 31 indicators of asset ownership and connectivity using principal component analysis (PCA). Spearman's ρ is reported in each panel. For PMT-CARI and the expert-based categorical method, higher scores indicate greater vulnerability. We therefore reverse the sign of these scores in the figure so that, for all methods, higher values correspond to higher welfare.

The three figures show that all four prioritization scores are positively correlated with each welfare benchmark, once the scores are oriented so that higher values indicate greater welfare. The associations are stronger for expenditure and wealth than for the FCS. As expected, PMT-Exp is most strongly associated with observed expenditure, while PMT-FCS is most strongly associated with observed FCS, the outcomes each model was designed to predict. The categorical score, despite being constructed through a very different approach, is also positively correlated with all three welfare measures, although generally less strongly than the PMTs. Overall, the results suggest that the different rules capture related dimensions of household welfare, while the dispersion around the fitted relationships also illustrates the limits of predicting welfare from observable household characteristics alone.

Finally, Table B.9 compares our results with estimates from the targeting literature.[4] We focus primarily on PMT-Exp, since most of these studies evaluate targeting against consumption expenditure. Its AUC of 0.83 and inclusion error of 26% fall within the range reported for other PMTs. Its inclusion error is below the 31% reported for the administrative data model in Altındağ et al. (2021) for Syrian refugees in Lebanon and the 32% average reported by Brown et al. (2018) across nine countries at 40% coverage. For broader context, AUCs across our three PMTs range from 0.71 to 0.83 when evaluated against their respective outcomes, compared with 0.61 to 0.73 for models using mobile phone data to identify consumption poverty (Aiken et al., 2022, 2025). Differences in coverage, evaluation populations, and welfare benchmarks limit direct comparisons, but these results suggest that the performance of our models is broadly consistent with, and in many cases compares favorably to, that documented in the targeting literature.

[4]We assembled the comparison sample in two stages. First, we manually screened all references cited in Coady et al. (2004), Grosh et al. (2022), and Devereux et al. (2017), covering approximately 405 entries, to identify studies evaluating household targeting rules against an observed welfare measure and reporting at least one quantitative measure of targeting performance. To extend the search to include more recent studies, we then used OpenAlex to retrieve works citing or cited by Alatas et al. (2012); Brown et al. (2018); Altındağ et al. (2021); Aiken et al. (2022, 2025). This produced 1,021 unique records. After screening titles and abstracts, 64 studies remained, of which 21 evaluated statistical targeting methods. Full text review using the same inclusion criteria retained 11 studies. We used a large language model (Claude Opus) to assist with extracting the relevant estimates from each paper. The reported values were then checked manually against the original tables and text, and the corresponding table and page references were recorded in the replication files. We also harmonised the reported measures across studies, noting whether targeting errors were calculated for the full population or only within programme areas, and whether welfare was measured using consumption or an index. This checking process identified only one extraction error, which was corrected. We excluded three studies whose reported error measures that were not comparable with ours because errors were defined against specific programme eligibility, were defined against a predicted index, or were only reported as aggregated misclassified households, without distinguishing inclusion and exclusion errors. We also excluded community based targeting rules, since our comparison focuses on statistical prediction methods.

Table B.9. Targeting performance in comparative perspective

Study	Country	Method	Cov.	ρ	AUC	Incl. err
Our study (Lebanon)						
		PMT-Exp	43%	0.57	0.83	0.26
		PMT-CARI	43%	0.47	0.78	0.36
		PMT-FCS	43%	0.35	0.71	0.42
		Expert (categ.)	43%	0.21	0.62	0.45
Literature: proxy means tests						
Brown et al. (2018)	Burkina Faso	PMT	40%	–	–	0.29
	Ethiopia	PMT	40%	–	–	0.39
	Ghana	PMT	40%	–	–	0.26
	Malawi	PMT	40%	–	–	0.34
	Mali	PMT	40%	–	–	0.32
	Niger	PMT	40%	–	–	0.35
	Nigeria	PMT	40%	–	–	0.24
	Tanzania	PMT	40%	–	–	0.29
	Uganda	PMT	40%	–	–	0.33
	Mean	PMT	40%	–	–	*0.32*
Altındağ et al. (2021)	Lebanon (ref.)	PMT	50%	–	–	0.31
Aiken et al. (2025)	Bangladesh	PMT	21%	0.65	0.82	0.48[a]
Aiken et al. (2022)	Togo	PMT	29%	0.72	0.85	0.37
Schnitzer and Stoeffler (2024)	Cameroon	PMT	67%	–	–	0.24[a]
	Mali	PMT	66%	–	–	0.27[a]
	Burkina Faso (I)	PMT	60%	–	–	0.27[a]
	Niger (I)	PMT	45%	–	–	0.48[a]
	Niger (II)	PMT	45%	–	–	0.47[a]
	Senegal (II)	PMT	41%	–	–	0.36[a]
	Senegal (I)	PMT	36%	–	–	0.41[a]
	Burkina Faso (II)	PMT	21%	–	–	0.54[a]
	Median	PMT	45%	–	–	*0.39*[a]
Stoeffler et al. (2016)	Cameroon	PMT	67%	–	–	0.21[a]
Premand and Schnitzer (2021)	Niger	PMT	40%	–	–	0.39[a]
Klasen and Lange (2015)	Bolivia (2008)	PMT parsimonious (M1)[d]	50%	–	0.74	0.30
	Bolivia (2011)	PMT parsimonious (M1)	50%	–	0.74	0.31
	Bolivia (2008)	PMT full (M4)	50%	–	0.84	0.23
	Bolivia (2011)	PMT full (M4)	50%	–	0.85	0.22
Hillebrecht et al. (2020)	Burkina Faso	PMT	20%	–	–	0.53[c]
Wobcke and Mariyah (2023)	Indonesia (Brebes)	PMT[e]	40%	–	–	0.33
	Indonesia (S. Tengah)	PMT	40%	–	–	0.31
	Indonesia (Surabaya)	PMT	40%	–	–	0.24
Alatas et al. (2012)	Indonesia	PMT	30%	0.41	–	0.18[b]
Literature: machine learning (phone)						
Aiken et al. (2022)	Togo	ML	29%	0.45	0.73	0.50
Aiken et al. (2025)	Bangladesh	ML	21%	0.23	0.61	0.68[a]
	Togo	ML	29%	0.44	0.72	0.50

Notes: Coverage indicates the share of households selected. ρ denotes Spearman's rank correlation between the targeting score and the welfare measure. Both ρ and the AUC use the full ranking and therefore do not depend on the selection cutoff, whereas inclusion error depends on both the coverage rate and the population over which it is evaluated. Inclusion error is defined as $1 -$ precision, that is, the share of selected households that are not classified as vulnerable. For multi-country studies, we report each country or programme separately together with the cross-study mean in Brown et al. (2018) or median in Schnitzer and Stoeffler (2024). [a] Computed within programme areas only. [b] Reported as the share of non-poor households selected. [c] The original study reports false-positive rate. Here we recomputed it as $1 -$ precision at 20% coverage. [d] Four specifications of increasing complexity, M1 to M4, evaluated at 50% coverage. [e] Here we report LASSO estimates at 40% coverage, as results across six different estimators considered in the study differ by at most four percentage points.

Additional References

Altındağ, Onur, Stephen D. O'Connell, Aytuğ Şaşmaz, Zeynep Balcıoğlu, Pinella Cadoni, Matthew Jerneck, and Aaron K. Foong, "Targeting humanitarian aid using administrative data: Model design and validation," *Journal of Development Economics*, 2021, *148*, 102564.

Athey, Susan, Julie Tibshirani, and Stefan Wager, "Generalized Random Forests," *The Annals of Statistics*, 2019, *47* (2), 1148–1178.

Coady, David, Margaret Grosh, and John Hoddinott, "Targeting of Transfers in Developing Countries," in "Targeting of Transfers in Developing Countries," World Bank, 2004.

Deaton, Angus and Salman Zaidi, *Guidelines for Constructing Consumption Aggregates for Welfare Analysis*, Vol. 135, World Bank Publications, 2002.

Devereux, Stephen, Edoardo Masset, Rachel Sabates-Wheeler, Michael Samson, Althea-Maria Rivas, and Dolf te Lintelo, "The Targeting Effectiveness of Social Transfers," *Journal of Development Effectiveness*, 2017, *9* (2), 162–211.

Grosh, Margaret, Phillippe Leite, Matthew Wai-Poi, and Emil Tesliuc, *Revisiting Targeting in Social Assistance: A New Look at Old Dilemmas*, World Bank, 2022.

Harrell, Frank E., *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*, 2 ed., Cham: Springer, 2015.

Hillebrecht, Michael, Stefan Klonner, Noraogo A. Pacere, and Aurélia Souares, "Community-Based versus Statistical Targeting of Anti-Poverty Programs: Evidence from Burkina Faso," *Journal of African Economies*, 2020, *29* (3), 271–305.

Klasen, Stephan and Simon Lange, "Targeting Performance and Poverty Effects of Proxy Means-Tested Transfers: Trade-offs and Challenges," IAI Discussion Papers 231, Ibero-America Institute for Economic Research, Göttingen 2015.

Mancini, Giulia and Giovanni Vecchi, "On the Construction of a Consumption Aggregate for Inequality and Poverty Analysis," Technical Report, World Bank, Washington, DC 2022.

Schnitzer, Pascale and Quentin Stoeffler, "Targeting Social Safety Nets: Evidence from Nine Programs in the Sahel," *The Journal of Development Studies*, 2024, *60* (4), 574–595.

Smith, Gary, "Step Away from Stepwise," *Journal of Big Data*, 2018, *5* (32), 1–12.

Stoeffler, Quentin, Bradford Mills, and Carlo del Ninno, "Reaching the poor: Cash transfer program targeting in Cameroon," *World Development*, 2016, *83*, 244–263.

Wobcke, Wayne and Siti Mariyah, "Machine Learning and Data Augmentation in the Proxy Means Test for Poverty Targeting," *Statistical Journal of the IAOS*, 2023, *39*, 961–977.

Appendix C: Variables Definition

This appendix describes the construction of the main variables used in the paper. Table C1 presents the four primary outcome families in the order in which they appear in the analysis. Table C2 then defines their underlying components, the secondary outcomes, and the baseline measures used in descriptive and heterogeneity analyses. "Don't know" and "Prefer not to disclose" responses are always treated as missing.

Table C1: Construction of the Main Outcome Variables

Outcome family	Definition and construction	Scale and direction
Food expenditure	For each of nine food groups, the survey records the value consumed during the previous 30 days from purchases made in cash or on credit, as well as the estimated value of food obtained through own production, gathering, or in-kind payment. The nine food groups are cereals and tubers; vegetables; fruits; meat, fish, and eggs; pulses and nuts; milk and dairy; oil and fat; sugar and sweets; and condiments and spices. Values are converted to US dollars and top-coded at the wave-specific 95th percentile. In the main outcome analysis, total food expenditure is standardized within each survey wave.	Depending on the table, values are reported in USD or within-wave standard deviation units.[1]
Food security index	The index combines the Food Consumption Score (FCS) and the reverse-coded reduced Coping Strategies Index (rCSI) using inverse-covariance weights (Anderson, 2008). The FCS is coded so that more frequent and diverse food consumption increases the index, while the rCSI is reverse-coded so that greater reliance on harmful coping strategies decreases it.	Within-wave standard deviation units.[1] Higher values indicate greater food security.
Psychological well-being index	The index combines measures of life satisfaction, reverse-coded worriedness, and mental health using inverse-covariance weights (Anderson, 2008). Each component is coded so that higher values indicate better psychological well-being.	Within-wave standard deviation units.[1] Higher values indicate better psychological well-being.
Trust index	The index combines seven measures of trust in different groups and institutions using inverse-covariance weights (Anderson, 2008). Each component is coded so that higher values indicate greater trust.	Within-wave standard deviation units.[1] Higher values indicate greater trust.

[1]For ease of interpretation, we always standardize outcomes relative to the relevant reference group. In the assistance discontinuation analysis, outcomes are standardized within each survey wave using the mean and standard deviation among continued assistance households. In the prioritization analysis, we instead use the mean and standard deviation among households assigned to the PMT-Exp arm, which is the omitted reference group.

Table C2: Construction of Component, Secondary, and Baseline Variables

Variable	Definition and construction	Scale
	Food expenditure	
Food expenditure by food group	*Question wording:* For each of nine food groups (cereals and tubers; vegetables; fruits; meat, fish, and eggs; pulses and nuts; milk and dairy; oil and fat; sugar and sweets; and condiments and spices), respondents are asked about consumption during the previous 30 days and separately report the "Value purchased using cash or credit" and the "Value produced, gathered, and/or received in exchange of labor." *Construction:* The main food-expenditure measure sums both reported values across the nine food groups. All values are converted to US dollars.	USD.
	Food security	
Food Consumption Score (FCS)	*Question wording:* "In the past 7 days, how many days did members of the household eat [cereals and tubers; vegetables; fruits; meat, fish, and eggs; pulses and nuts; milk and dairy; oil and fat; sugar and sweets; and condiments and spices]?" *Construction:* FCS equals: cereals, roots, and tubers (×2) + pulses (×3) + vegetables (×1) + fruits (×1) + meat, fish, and eggs (×4) + milk and dairy (×4) + sugar (×0.5) + oil and fat (×0.5) + condiments and spices (×0).	0–112, with higher values indicating higher food consumption.
Reduced Coping Strategies Index (rCSI)	*Question wording:* "In the past 7 days, how many days did anyone from the household [rely on less preferred or less expensive food; borrow food or rely on help from relatives or friends; reduce meal portions; restrict adults' consumption so children can eat; reduce the number of meals] ?" *Construction:* rCSI equals: rely on less preferred or less expensive food (×1) + borrow food or rely on help from relatives or friends (×2) + reduce meal portions (×1) + restrict adults' consumption so children can eat (×3) + reduce the number of meals (×1).	0–56, with higher values indicating greater reliance on coping strategies.
	Psychological well-being	
Life satisfaction	*Question wording:* "How satisfied are you with [the household's nutrition/food security; overall health; financial situation; interactions with the community and neighbors; or your life in general]?" *Construction:* Responses range from 1 (very satisfied) to 4 (very unsatisfied). Responses are reverse-coded and summed so that higher values indicate greater satisfaction.	5–20, with higher values indicating greater satisfaction.
Worriedness	*Question wording:* "How worried are you about [food; health care; education; indebtedness; employment; family; extremism; conflict or sectarian tensions; natural disasters; and migration or displacement]?" *Construction:* Responses range from 1 (not worried at all) to 4 (very worried). Responses are summed across the ten items, so that higher values indicate greater worry.	10–40, with higher values indicating greater worry.

continued on next page

Table C2 continued from previous page

Variable	Definition and construction	Scale
WHO-5 well-being score	*Question wording:* "For the following questions, please indicate for each which is closest to how you have been feeling over the past week." The five statements are: "I have felt cheerful and in good spirits"; "I have felt calm and relaxed"; "I have felt active and vigorous"; "I woke up feeling fresh and rested"; and "My daily life has been filled with things that interest me." *Construction:* Responses range from 0 (never) to 5 (all of the time). Responses are summed across the five items and multiplied by four.	0–100, with higher values indicating better well-being.
	Trust	
Trust in groups and institutions	*Question wording:* "How much do you currently trust the following people and organizations?" The seven items are: family; neighbors; religious organizations; international organizations (UN)/NGOs; local government; national government; and migrants from outside Lebanon. *Construction:* Responses range from 1 (trust completely) to 4 (do not trust at all). Each item is reverse-coded so that higher values indicate greater trust.	1–4 per item, with higher values indicating greater trust.
	Secondary outcomes	
Non-food expenditure	*Question wording:* "In the last 30 days, what is the estimated amount the household spent in cash and credit on [rent; electricity and lighting, including generator electricity; drinking water and reservoir water; heating, excluding electric heating; soap and cleaning products; communications, including telephone, mobile data, and WiFi; transportation, including fuel for cars and motorbikes; alcohol and tobacco; cooking gas; or household utilities and assets, such as cooking utensils and furniture]?". Respondents are also asked: "For the following items, please tell me the estimated amount spent in both cash and credit by the household since the beginning of the year / in the past 6 months: [medical expenses and health care; clothing and shoes; education; celebrations and social events; agricultural inputs; construction and house repairs; or other non-food expenditure not previously reported]". *Construction:* Expenditures on the seven less frequent categories are converted to monthly equivalents and added to expenditure on the ten categories measured over the previous 30 days. All amounts are converted to US dollars.	USD.
Social cohesion index	*Question wording:* "Since January 2024/July 2024, has there been tension between members of your community/neighborhood?"; "How many groups or associations have you or any household member been a member of since January 2024/July 2024?"; "Has any member of your household volunteered in community/neighborhood activities in the past 6 months?" *Construction:* The index combines the three measures using inverse-covariance weights (Anderson, 2008). The community tension question is reverse-coded so that lower tension increases the index. Components are standardized within each survey wave before aggregation.	Within-wave standard deviation units.[1] Higher values indicate greater social cohesion.

continued on next page

Table C2 continued from previous page

Variable	Definition and construction	Scale
Share of household members employed	*Question wording:* For each household member: "During the past 7 days, did [household member] do any work for wage, salary, commission, tip, or any other pay, for one or more hours?" and "During the past 7 days, did [household member] run/operate or do any kind of business, farming, or other activity to generate income, for one or more hours?" *Construction:* The employed share is the proportion of household members aged 15 or above who report either type of work.	0–1, with higher values indicating a larger employed share.
Weekly hours worked	*Question wording:* For each household member and reported job: "During the past 7 days, how many hours did [household member] work on this activity?" *Construction:* Weekly hours in paid employment and own-account work are summed across jobs and household members and divided by the total number of household members aged 15 or above (including those who are not employed).	Weekly hours per household member aged 15 or above.
Household has debt	*Question wording:* "Does your household currently have any debt with any institution or person, regardless if that institution/person is inside Lebanon or abroad?" *Construction:* The outcome is an indicator equal to one if the household answers yes.	0/1 indicator.
Remittances	*Question wording:* "Since January 2024/July 2024, has the household received any money from any person (e.g. relative, friend) living outside the household that you do NOT have to pay back?" *Construction:* The outcome is an indicator equal to one if the household answers yes.	0/1 indicator.
	Additional variables used in the analyses	
Purchases on credit	*Question wording:* After each expenditure module: "Out of this amount, please estimate the total amount spent on credit." *Construction:* Purchases on credit are summed across the relevant food or non-food expenditure categories. Non-credit expenditure is calculated as total expenditure minus the corresponding amount purchased on credit. All values are converted to US dollars.	USD.
Job quality index	*Question wording:* For each reported paid job: "What type of job is it?" For each paid or own-account activity: "During the past 7 days, how many hours did [household member] work on this activity?" *Construction:* Each employed household member aged 15 or above receives one point for holding at least one salaried job in the private or public sector, and one point for working between 1 and 48 hours per week, summed across all paid and own-account jobs. The two indicators are added, and the resulting scores are averaged across employed household members. The index is undefined when no member of the relevant group is employed.	0–2, with higher values indicating better job quality among employed members.

continued on next page

Table C2 continued from previous page

Variable	Definition and construction	Scale
Baseline wealth index	*Construction:* The index is constructed from 36 indicators of asset ownership and connectivity from the baseline census using principal component analysis (PCA). In the analysis it is standardized to have mean zero and standard deviation one in the full beneficiary population, with higher values indicating greater wealth.	Standard deviation units. Higher values indicate greater wealth.
Appeal submission	*Construction:* Using WFP administrative records, the outcome is an indicator equal to one if the household submitted an appeal against the decision to discontinue assistance, and zero otherwise.	0/1 indicator.
Priority-area status	*Construction:* The variable is an indicator equal to one if the household's cadastre of residence was classified by WFP as a high-priority area because it was affected by conflict and/or hosted a large number of internally displaced households, and zero otherwise. This geographic classification determined the areas targeted for emergency cash assistance between the midline and endline surveys. Figure A.3 shows the geographic distribution of priority areas.	0/1 indicator.
Rank distance to the eligibility cutoff	*Construction:* Households in the full beneficiary population are ranked according to their score under their randomly assigned prioritization method, with more vulnerable households ranked first. For each method, the cutoff rank equals the number of households selected for continued assistance under that method. Rank distance is defined as the household's rank minus the cutoff rank, with negative values set to zero, and is divided by 1,000.	Thousands of rank positions. Higher values indicate that the household lies farther below the eligibility cutoff.

Notes: Quoted text reproduces the English questionnaire, with square brackets denoting the item or person inserted by the survey instrument. Questions referring to "January 2024/July 2024" use January 2024 as the reference period at midline and July 2024 at endline.

Appendix D. Prioritization Methods and Assistance Discontinuation in Disagreement Eligibility Spaces

This appendix complements the full-sample comparison of prioritization methods reported in the main text, by focusing on the 14 eligibility spaces in which the methods disagree. In these spaces, households are eligible for continued assistance under at least one, but not all, of the four methods. Random assignment to a prioritization method therefore generates within-space variation in assistance continuation, allowing us to estimate the effect of assistance discontinuation separately within each eligibility space and survey wave. We use these estimates to assess whether the four prioritization methods tend to select households in eligibility spaces where discontinuation generates larger welfare losses.

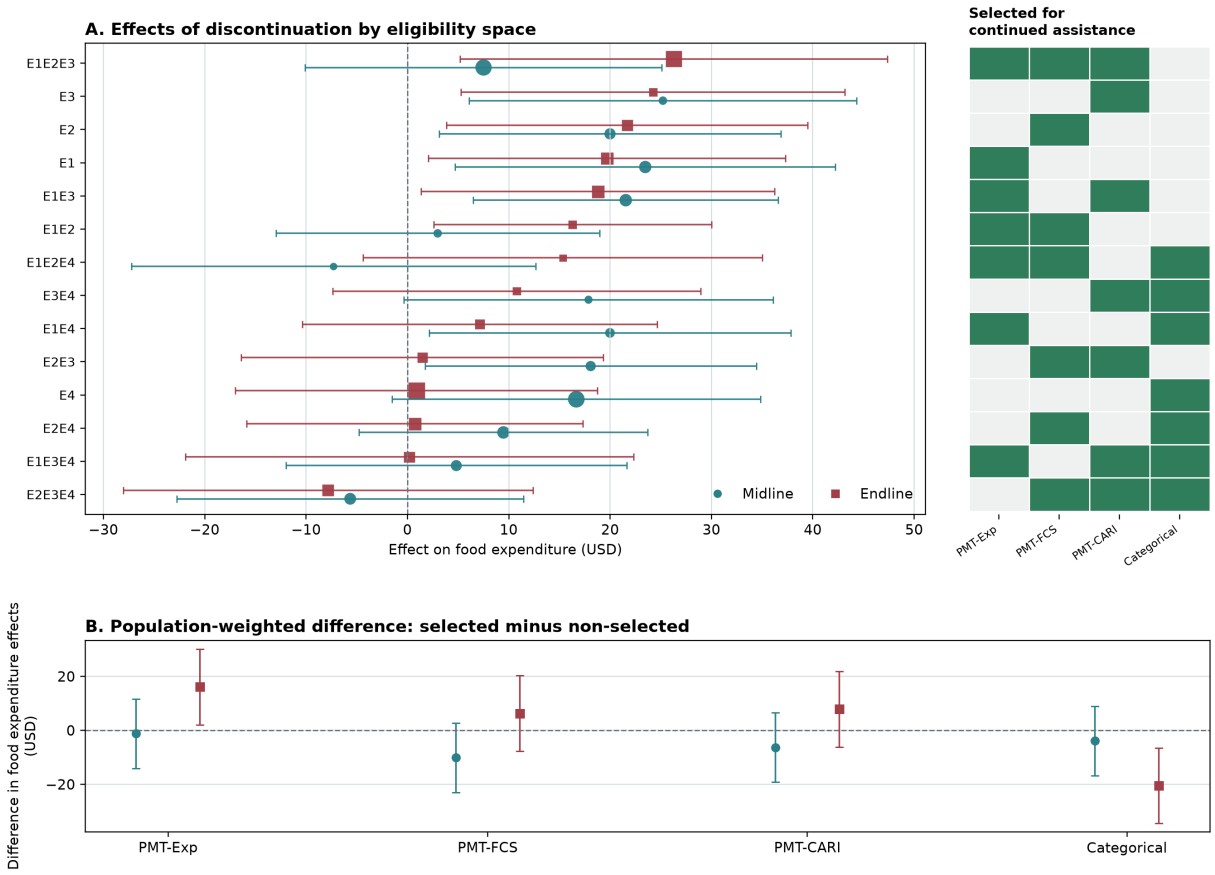
Figure D.1 reports results for food expenditure in USD. Figures D.2–D.4 report results for the three main welfare indices: food security, psychological well-being, and trust. Figure D.5 summarizes these outcomes using a composite welfare index equal to the unweighted average of the three standardized indices. For the three welfare indices and the composite measure, we report the negative of the estimated discontinuation effect, so that positive values indicate larger welfare losses from discontinuation. For food expenditure, we report the discontinuation effect in its original sign.

In each figure, Panel A reports the midline and endline estimates for each eligibility space, together with 95 percent confidence intervals. Eligibility spaces are ordered by the endline estimate. The point estimate marker size is proportional to the number of households in the space in the full 2023 beneficiary population (Table A.3). The adjacent matrix indicates in green the prioritization methods under which households in each eligibility space would be selected for continued assistance.

Panel B summarizes the eligibility-space estimates by prioritization method. For each method and survey wave, we calculate the population-weighted average estimated loss among the eligibility spaces selected by that method and subtract the corresponding average among spaces not selected by that method. Weights are based on each eligibility space's share of the full beneficiary population. A positive value therefore indicates that the method selects (for continued assistance) eligibility spaces in which estimated losses from discontinuation would be larger, while a negative value indicates larger estimated losses among the spaces it does not select (for continued assistance). For food expenditure, the same calculation compares expenditure responses rather than welfare losses.

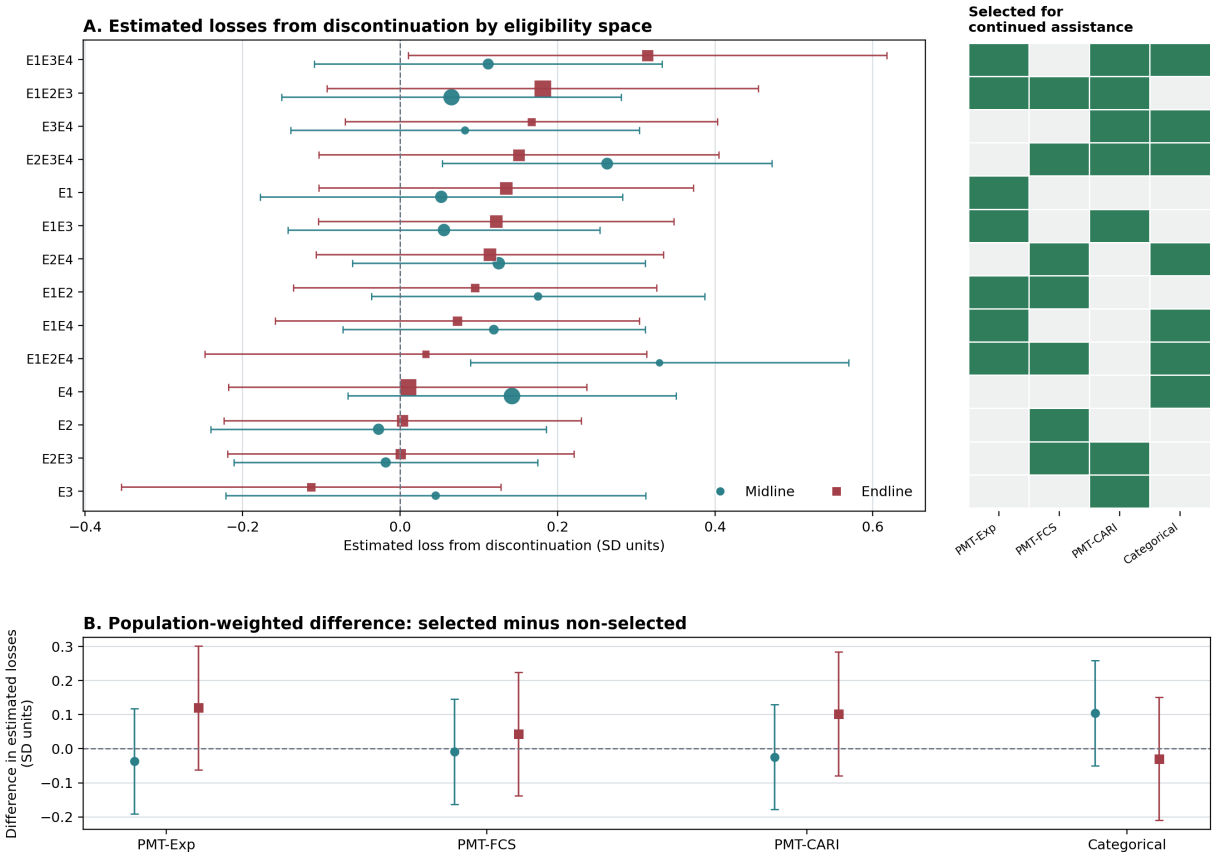
Across outcomes, the results do not point to a consistent ordering of the four prioritization methods. Panel B estimates are generally imprecise and vary across outcomes and survey waves. For the composite welfare index, the confidence intervals include zero for all four methods at both midline and endline. The results, therefore, confirm that no method systematically selects the eligibility spaces in which households would experience the largest welfare losses from assistance discontinuation.

Figure D.1: Prioritization Methods and Food Expenditure Responses to Assistance Discontinuation.



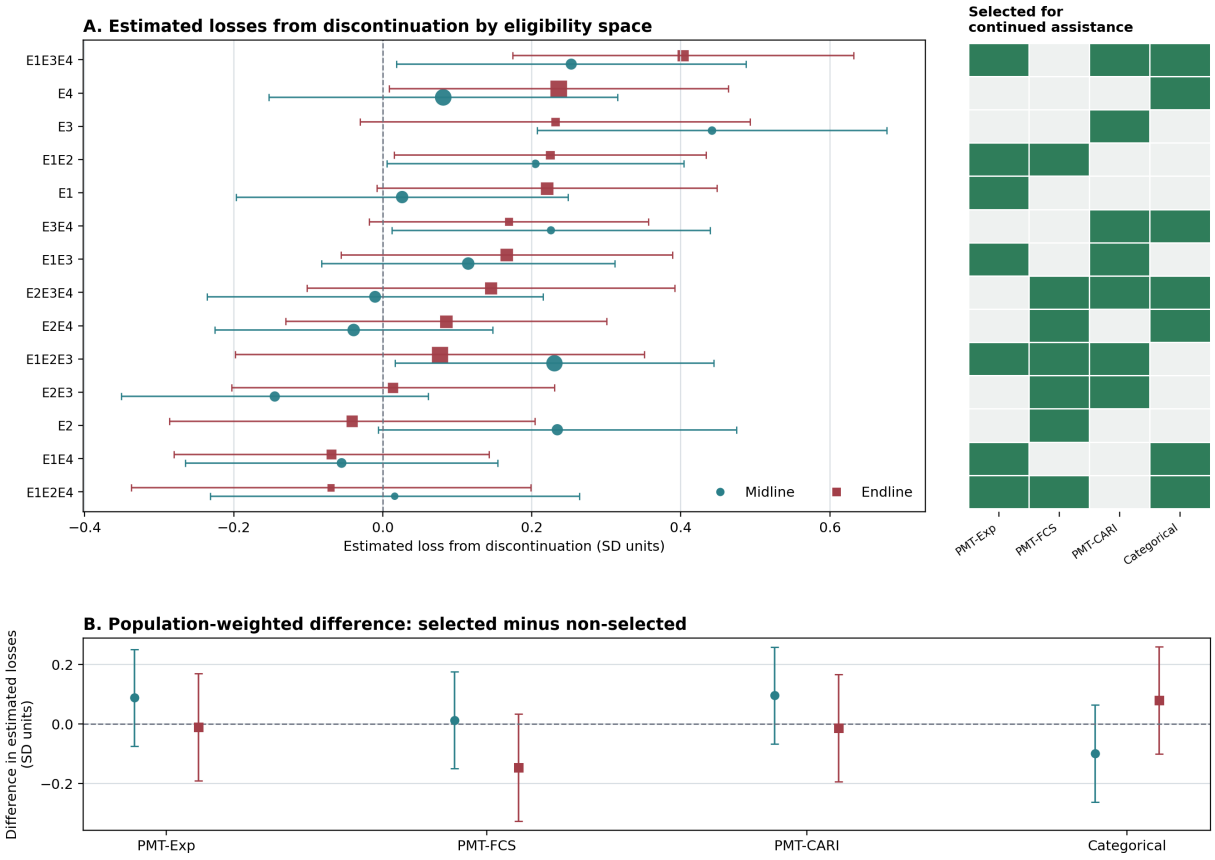
Notes: Panel A reports the estimated effects of assistance discontinuation on food expenditure for each of the 14 disagreement eligibility spaces, separately at midline and endline. Points show coefficient estimates and lines indicate 95 percent confidence intervals. Marker size is proportional to the number of households in each eligibility space in the full 2023 beneficiary population. Green cells in the adjacent matrix indicate the prioritization methods under which households in each space would be selected for continued assistance. Panel B summarizes these estimates by prioritization method. For each method and survey wave, it reports the population-weighted average discontinuation effect among eligibility spaces selected by the method minus the corresponding average among spaces not selected by the method. Weights are based on each space's share of the full beneficiary population. Positive values indicate larger increases in food expenditure following discontinuation among eligibility spaces selected by the method.

Figure D.2: Prioritization Methods and Food Security Losses from Assistance Discontinuation.



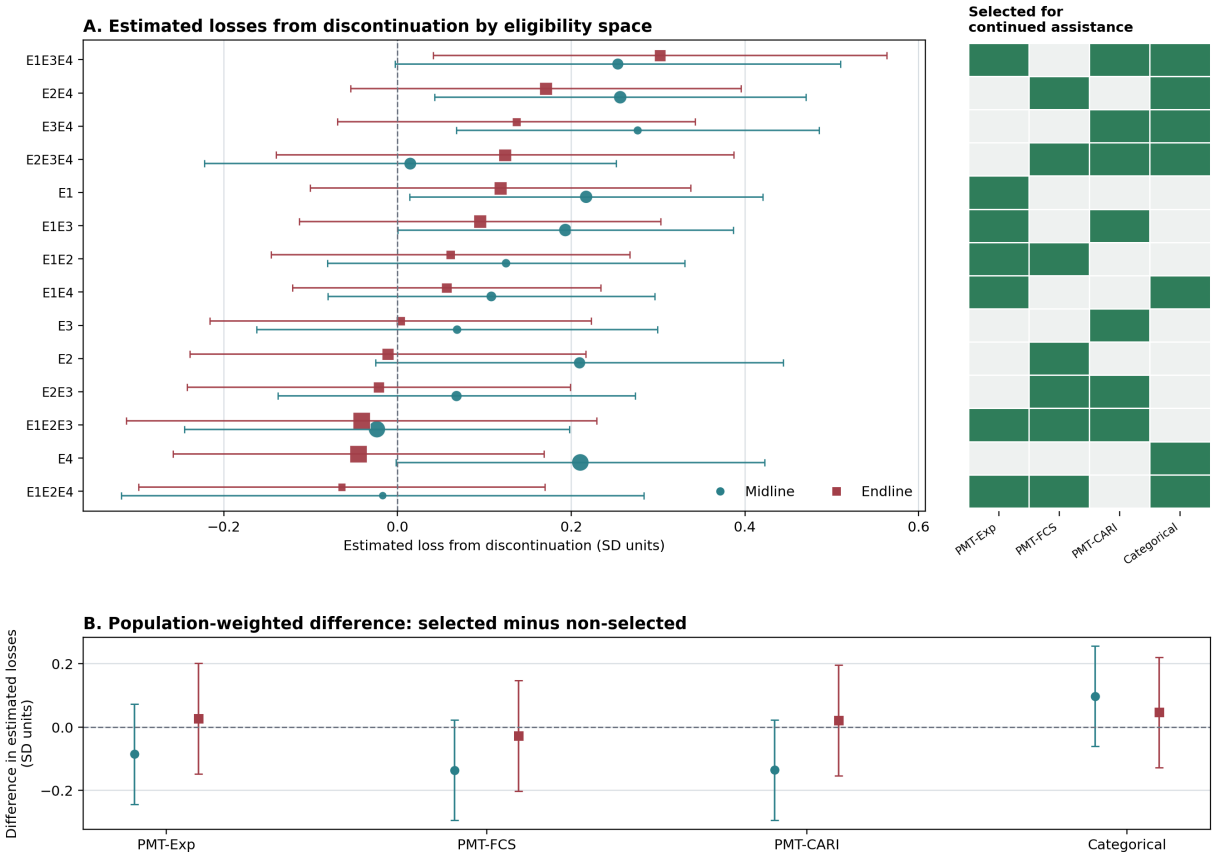
Notes: Panel A reports estimated losses in the food security index from assistance discontinuation for each of the 14 disagreement eligibility spaces, separately at midline and endline. Losses are defined as the negative of the estimated discontinuation effect, so that positive values indicate larger declines in food security. Points show estimated losses and lines indicate 95 percent confidence intervals. Marker size is proportional to the number of households in each eligibility space in the full 2023 beneficiary population. Green cells in the adjacent matrix indicate the prioritization methods under which households in each space would be selected for continued assistance. Panel B summarizes these estimates by prioritization method. For each method and survey wave, it reports the population-weighted average estimated loss among eligibility spaces selected by the method minus the corresponding average among spaces not selected by the method. Weights are based on each space's share of the full beneficiary population. Positive values indicate larger estimated food security losses among eligibility spaces selected by the method.

Figure D.3: Prioritization Methods and Psychological Well-being Losses from Assistance Discontinuation.



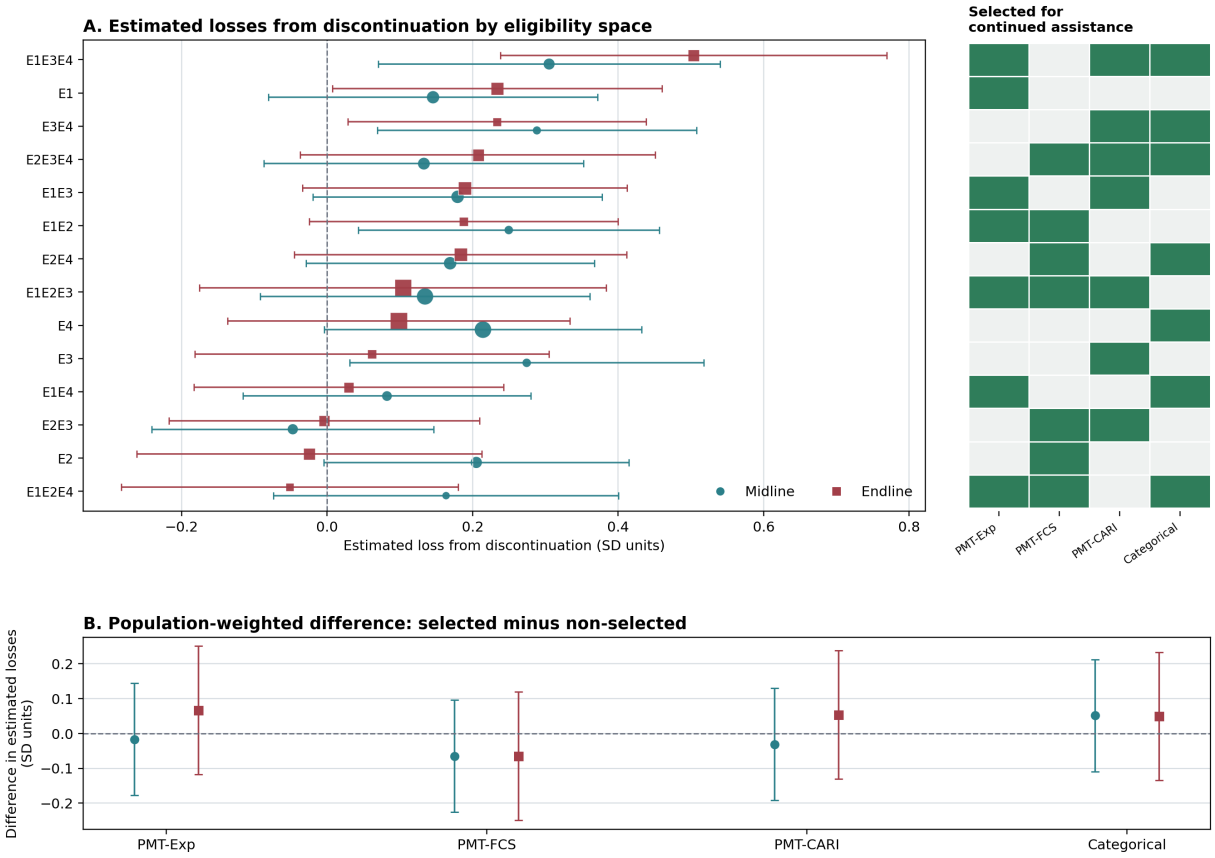
Notes: Panel A reports estimated losses in the psychological well-being index from assistance discontinuation for each of the 14 disagreement eligibility spaces, separately at midline and endline. Losses are defined as the negative of the estimated discontinuation effect, so that positive values indicate larger declines in psychological well-being. Points show estimated losses and lines indicate 95 percent confidence intervals. Marker size is proportional to the number of households in each eligibility space in the full 2023 beneficiary population. Green cells in the adjacent matrix indicate the prioritization methods under which households in each space would be selected for continued assistance. Panel B summarizes these estimates by prioritization method. For each method and survey wave, it reports the population-weighted average estimated loss among eligibility spaces selected by the method minus the corresponding average among spaces not selected by the method. Weights are based on each space's share of the full beneficiary population. Positive values indicate larger estimated psychological well-being losses among eligibility spaces selected by the method.

Figure D.4: Prioritization Methods and Trust Losses from Assistance Discontinuation.



Notes: Panel A reports estimated losses in the trust index from assistance discontinuation for each of the 14 disagreement eligibility spaces, separately at midline and endline. Losses are defined as the negative of the estimated discontinuation effect, so that positive values indicate larger declines in trust. Points show estimated losses and lines indicate 95 percent confidence intervals. Marker size is proportional to the number of households in each eligibility space in the full 2023 beneficiary population. Green cells in the adjacent matrix indicate the prioritization methods under which households in each space would be selected for continued assistance. Panel B summarizes these estimates by prioritization method. For each method and survey wave, it reports the population-weighted average estimated loss among eligibility spaces selected by the method minus the corresponding average among spaces not selected by the method. Weights are based on each space's share of the full beneficiary population. Positive values indicate larger estimated trust losses among eligibility spaces selected by the method.

Figure D.5: Prioritization Methods and Composite Losses from Assistance Discontinuation.



Notes: Panel A reports estimated losses in the composite index from assistance discontinuation for each of the 14 disagreement eligibility spaces, separately at midline and endline. The index is the unweighted average of the standardized food security, psychological well-being, and trust indices. Losses are defined as the negative of the estimated discontinuation effect, so that positive values indicate larger declines in welfare. Points show estimated losses and lines indicate 95 percent confidence intervals. Marker size is proportional to the number of households in each eligibility space in the full 2023 beneficiary population. Green cells in the adjacent matrix indicate the prioritization methods under which households in each space would be selected for continued assistance. Panel B summarizes these estimates by prioritization method. For each method and survey wave, it reports the population-weighted average estimated loss among eligibility spaces selected by the method minus the corresponding average among spaces not selected by the method. Weights are based on each space's share of the full beneficiary population. Positive values indicate larger estimated composite welfare losses among eligibility spaces selected by the method.