



WFP EVALUATION



**World Food
Programme**

**SAVING
LIVES
CHANGING
LIVES**

Targeting Humanitarian Aid in Crisis: Evidence from the reduction of WFP Assistance in Lebanon

OEV/2023/034

May 2026

Key personnel for the evaluation

WORLD FOOD PROGRAMME – OFFICE OF EVALUATION

Anne-Claire Luzot, Director of Evaluation

Jonas Heirman, Senior Evaluation Officer – Head of Impact Evaluation Unit

Felipe Dunsch, Impact Evaluation Officer

Ola ElToukhi, Impact Evaluation Analyst

Kristen McCollum, Impact Evaluation Analyst

Micah Cruz, Impact Evaluation Analyst

ISDC – INTERNATIONAL SECURITY AND DEVELOPMENT CENTER

Tilman Brück, Director

Ghassan Baliki, Research Director

Dorothee Weiffen, Researcher

Andrea Guariso, Research Affiliate

Jonathan Garcia, Research Affiliate

WORLD FOOD PROGRAMME – LEBANON COUNTRY OFFICE

Marco Principi, Head of Research, Assessment and Monitoring

Soha Moussa, Head of Monitoring and Evaluation

Catherine Saiid, Head of Vulnerability Assessment and Mapping

Elissar Samaha, Vulnerability Assessment and Mapping Data Manager

Abdulhaseeb Ayoubi, Monitoring and Evaluation Officer

Acknowledgements

WFP's learning partner on this impact evaluation is the ISDC - International Security and Development Center gGmbH.



Suggested citation: Baliki, G., Cruz, M., Dunsch, F., Garcia, J., Guariso, A., Heirman, J., & Weiffen, D. (2026). *Targeting Humanitarian Aid in Crisis: Evidence from the reduction of WFP assistance in Lebanon*. World Food Programme, Rome.

Disclaimer

The opinions expressed are those of the evaluation team, and do not necessarily reflect those of the World Food Programme (WFP) or International Security and Development Center (ISDC). Responsibility for the opinions expressed in this report rests solely with the authors. Publication of this document does not imply endorsement by WFP or ISDC of the opinions expressed.

The designations employed and the presentation of material in the maps do not imply the expression of any opinion whatsoever on the part of WFP or ISDC concerning the legal or constitutional status of any country, territory or sea area, or concerning the delimitation of frontiers.

Contents

Executive summary.....	i
1. Introduction.....	7
2. Country context.....	9
3. Programme description.....	10
3.1. Programme background.....	10
3.2. Targeting approaches.....	13
4. Impact evaluation motivation and design.....	15
4.1. Motivation.....	15
4.2. Evaluation questions.....	18
4.3. Evaluation design.....	19
5. Data and outcomes.....	23
5.1. Data sources.....	23
5.2. Outcomes.....	24
6. Main findings.....	25
Part 1: Prioritizing among households already being supported when scaling down.....	25
Part 2: The consequences for households excluded from assistance.....	40
Part 3: Whether and how households prepare for programme exit.....	56
7. Conclusions and considerations for future programming.....	62
7.1. Conclusions.....	62
7.2. Key considerations.....	63
8. Bibliography.....	65
9. Acronyms.....	70
10. Annexes.....	71
A1. Programme implementation additional details.....	71
A2. Targeting additional details.....	77
A3. Empirical strategy.....	87
A4. Randomization, sampling, balance, and attrition.....	91
A5. Limitations.....	103
A6. Additional details on outcomes.....	105
A7. Additional tables and figures on the main findings.....	115
A8. Data collection and ethical considerations.....	131
A9. Stakeholder analysis.....	133
A10. SMS reminder messages.....	134

List of figures

Figure 1. Distribution across Lebanon of households in the WFP in-kind food assistance programme	12
Figure 2. Random allocation of households into the four targeting approaches	20
Figure 3. Eligibility map.....	21
Figure 4. Timeline of programme implementation and data collection	23
Figure 5. Predictors of selection of targeting methods (using full census)	26
Figure 6. Appeal rate of discontinued households by targeting arm.....	28
Figure 7. Appeal rate of discontinued households by eligibility	29
Figure 8. FCS across targeting arms.....	31
Figure 9. Food expenditure (USD) across targeting arms	31
Figure 10. rCSI across targeting arms.....	32
Figure 11. Usage of livelihood stress coping strategies in the past six months across targeting arms	33
Figure 12. Life satisfaction across targeting arms	34
Figure 13. Mental wellbeing (WHO-5) across targeting arms.....	35
Figure 14. Worriedness across targeting arms	35
Figure 15. PTSD score across targeting arms	36
Figure 16. Descriptive differences in rCSI between discontinued and continued households by targeting method during survey round 1 (May 2024)	38
Figure 17. Comparison of the differential impacts of targeting on rCSI between PMT-Exp and Categorical.....	39
Figure 18. Comparison of the differential impacts of targeting on worriedness between PMT-Exp and Categorical.....	40
Figure 19. Impact of the discontinuation of assistance on FCS	42
Figure 20. Impact of the discontinuation of assistance on FCS components	42
Figure 21. Impact of the discontinuation of assistance on food expenditure (USD)	44
Figure 22. Impact of the discontinuation of assistance on food expenditure by components of the food basket (USD).....	44
Figure 23. Impact of the discontinuation of assistance on total expenditure (USD) by payment modality.....	46
Figure 24. Impact of the discontinuation of assistance on rCSI.....	47
Figure 25. Impact of the discontinuation of assistance on rCSI components.....	47
Figure 26. Impact of the discontinuation of assistance on the probability of adopting negative livelihood-related coping strategies in the past six months.....	48
Figure 27. Impact of the discontinuation of assistance on psychological well-being	49
Figure 28. Impact of the discontinuation of assistance on trust	50
Figure 29. Impact of the discontinuation of assistance on finances, employment, women’s agency, and general welfare	52
Figure 30. Impacts of discontinuation on food security by degree of eligibility	53

Figure 31. Impacts of discontinuation on psychological wellbeing by degree of eligibility	54
Figure 32. Differential impacts of discontinuation on food expenditure in USD by gender of the household head.....	55
Figure 33. Differential impacts of discontinuation on food expenditure in USD by wealth	55
Figure 34. Differential impacts of discontinuation on food expenditure in USD by presence of a household member with a disability	56
Figure 35. FCS for long notice, continued and discontinued households	57
Figure 36. Food expenditure (USD) for long notice, continued and discontinued households	58
Figure 37. rCSI for long notice, continued and discontinued households	58
Figure 38. Life satisfaction long notice, continued and discontinued households	59
Figure 39. Mental wellbeing (WHO-5) for long notice, continued and discontinued households..	59
Figure 40. Worriedness for long notice, continued and discontinued households	60

Annex Figures

Figure A1.1. Adherence to implementation/study protocol.....	74
Figure A1.2. Conflict events registered in between July and December 2024 at sub-district level	75
Figure A1.3. Cadaster-level prioritisation for cash distribution	76
Figure A4.1. Geographical distribution of study sample households by selection status	101
Figure A4.2. Geographical distribution of households by study arm.....	102
Figure A7.1. Correlation in selection between the four targeting approaches.....	122
Figure A7.2. Predictors of selection of levels of eligibility (using full census)	123
Figure A7.3. Distribution of FCS by targeting arm	124
Figure A 7.4. Distribution of food expenditure (USD) by targeting arm	124
Figure A 7.5. Distribution of rCSI by targeting arm.....	124
Figure A7.6. Social cohesion by targeting arm	125
Figure A7.7. Trust index by targeting arms.....	126
Figure A7.8. Impact of the discontinuation of assistance on social cohesion	127
Figure A7.9. Differential impacts of discontinuation on FCS by gender of the household head .	128
Figure A7.10. Differential impacts of discontinuation on life satisfaction by gender of the household head.....	128
Figure A7.11. Differential impacts of discontinuation on FCS by wealth.....	129
Figure A7.12. Differential impacts of discontinuation on life satisfaction by wealth.....	129
Figure A7.13. Differential impacts of discontinuation on FCS by presence of a member with a disability.....	130
Figure A7.14. Differential impacts of discontinuation on life satisfaction by presence of a member with a disability	130

List of tables

Table 1. Outcomes of interest	24
-------------------------------------	----

Annex Tables

Table A1.1. Distribution of in-kind assistance households across governorates and districts	72
Table A1.2. Composition of the in-kind food parcel distributed by WFP in Lebanon	73
Table A2.1. Indicators tested for PMT models	78
Table A2.2. Construction of the Food Consumption Score (FCS)	80
Table A2.3. Construction and classification of the CARI	81
Table A2.4. Performance of the PMT models.....	83
Table A2.5. Categorical score weighting.....	84
Table A4.1. Allocation of households across eligibility spaces and targeting methods	91
Table A4.2. Planned sample composition and rollout by eligibility space at rounds 1 & 2	93
Table A4.3. Survey sample and attrition	94
Table A4.4. Balance in baseline characteristics.....	95
Table A4.5. Balance across targeting methods	96
Table A4.6. Balance across assistance discontinuation status	98
Table A4.7. Balance across long notice and control	99
Table A4.8. Balance in randomisation across districts.....	100
Table A6.1. Details on outcomes of interest.....	105
Table A6.2. Details on the construction of composite indices	108
Table A7.1. Differences in selection in targeting across the targeting methods	115
Table A7.2. Targeting secondary outcomes – finances.....	116
Table A7.3. Targeting secondary outcomes – employment	116
Table A7.4. Targeting secondary outcomes – human capital & agency	117
Table A7.5. Heterogeneity in targeting by sex of household head	118
Table A7.6. Heterogeneity in targeting by age of household head (minor/elderly).....	118
Table A7.7. Heterogeneity in targeting by household head education (primary or lower).....	119
Table A7.8. Heterogeneity in targeting by presence of member with a disability	119
Table A7.9. Heterogeneity in targeting by household wealth (poorest 50%).....	120
Table A7.10. Heterogeneity in targeting by dependency ratio	120
Table A7.11. Heterogeneity in targeting by discontinuation status	121

Executive summary

Context

1. Over the past two decades, humanitarian organizations like the World Food Programme (WFP) have faced increasing budget constraints and volatile funding. This has led to difficult decisions about how to allocate limited resources.
2. Starting in 2020, in response to the country's severe economic crisis, the WFP Lebanon Country Office (LBCO) started providing food assistance to vulnerable Lebanese households. By 2023, they had reached more than 5 percent of the population. However, in 2024, a combination of funding cuts and shifting priorities required the LBCO to reduce its assistance coverage by more than half. This prompted a need for evidence-based strategies to prioritize the households that would continue to receive assistance.
3. The study focuses on the January retargeting for which the LBCO aimed to reduce the number of beneficiaries from 58,501 households to 25,160, or about 43 percent of the existing caseload^{1,2} – a target based on the LBCO's estimated number of households with consumption levels below the Minimum Expenditure Basket (MEB), which was set at LBP 3,238,315 or USD 71 per capita per month (January 2023 prices).

Impact evaluation objective

4. While there is substantial evidence to inform the design and roll-out of development and humanitarian programmes, there is virtually no rigorous evidence on how to effectively phase out such programmes. Policymakers receive little guidance on how to identify/target households that are most in need of continued assistance, while also minimizing the adverse consequences on vulnerable households whose assistance has been discontinued. To fill this gap, at the request of the LBCO, WFP's Office of Evaluation and the International Security and Development Center (ISDC) conducted a large-scale impact evaluation. The study focuses on three key evaluation questions:
5. **How should households be prioritized/retargeted to minimize welfare losses among the communities receiving food assistance? (Retargeting/prioritization):** The study does so by rigorously comparing four common strategies to target households for assistance. The population of assisted households is randomly divided into four groups of equal size, and a different targeting approach is applied to each group. Food assistance continues to be provided to households that are deemed most vulnerable according to the targeting approach assigned to that group, while other households are excluded from

¹ This estimation was originally based on extrapolation from the nationally representative Lebanon Household Survey (LHS) dataset. This figure aligned closely with findings from the LBCO's post-distribution monitoring survey in the first half of 2023 of a representative sample of households receiving in-kind food assistance, which indicated that 44 percent were below the MEB threshold.

² The impact evaluation period began in August 2023 and analysis concluded in January 2026, before the resumption of hostilities between Israel and Lebanon in February 2026.

assistance. The performance of the four methods relative to one another is then assessed by considering a rich set of household welfare measures.

6. **What are the consequences of ending assistance for households that were excluded by the targeting process? (Discontinuation):** The variation generated by the randomization of the targeting approaches allows the study to estimate the impact on household welfare of ending assistance – for those households that were eligible under some targeting methods but not others (61 percent of the population³).
7. **How do households prepare for the end of the assistance? (Long notice):** A random subset of households that were excluded from assistance by the prioritization exercise received a longer “notice window”: six months before the end of assistance compared to the standard of one month. By monitoring the behaviour of these “longer notice” households over time, the study is able to assess whether (and how) households undertake preventive actions to prepare for the end of assistance, and whether a longer warning period might result in better outcomes.

Targeting approaches

8. When considering the households most in need of food assistance, there is significant debate over how to assess or predict food insecurity. Many WFP country offices implement vulnerability-based (data-driven) methods, often choosing between two approaches: predictive modelling based on a proxy-means test (PMT) or categorical scoring. The LBCO had implemented both approaches in different phases of its in-kind food assistance programme. However, they were uncertain about which approach was most effective, and unsure about which PMT vulnerability criterion to use (“predict”): household expenditure; food consumption; or a combination of the two – see the next paragraph for details.
9. While per-capita expenditure is the most used proxy for vulnerability, this indicator is more associated with poverty. As such, it may not be well-calibrated to identify the most food insecure among the poor. Although the Food Consumption Score (FCS) is used by WFP as a key indicator of food insecurity, it is often considered a transient measure and may not be predictive of long-term food needs. WFP has also developed an indicator called the Consolidated Approach for Reporting Indicators of Food Insecurity (CARI) that combines current food security measures with long-term livelihood measures, but CARI has not yet been widely tested. Although the three measures differ in focus, they all rely on large-scale data collection to predict household vulnerability outcomes using a PMT methodology.
10. Because of the complexity of PMT-based approaches, there are those who favour simpler vulnerability-based methods. For example, a simpler categorical targeting approach, where households are scored based on the number of vulnerability criteria they meet (set by programme or sector experts). Categorical vulnerability targeting differs from the PMT approach as it is not centred on one vulnerability outcome. Criteria used for categorical targeting is often based on well-known indicators of vulnerability, which streamlines data

³ The 39 percent who were not included in the analysis are composed of households who were “always eligible” or “never eligible” under all four targeting methods.

collection. The downside is that this approach might miss more nuanced facets of deprivation that a simple count might not reflect.

11. As a result of this uncertainty, and at the request of the LBCO, the evaluation compares these four targeting approaches – PMT based on expenditure (PMT-Exp); PMT based on FCS (PMT-FCS); PMT based on CARI (PMT-CARI); and categorical scoring (Categorical) – to address the LBCO’s need to optimize targeting. The approaches tested in the evaluation are among the most commonly used to target development and humanitarian programmes across the world.

Impact evaluation design

12. The evaluation applies a different design to address each of its three research areas.
13. **Retargeting/prioritization:** To answer the evaluation questions on prioritization and targeting, the 58,501 households in the study were randomly allocated, through individual randomization, into one of four equally sized study arms (~14,626 households per study arm). Within each arm, one of the four retargeting methods was applied to prioritize assistance. The 43 percent of households deemed most vulnerable in each arm were then selected to continue receiving assistance (6,290 households from each study arm, for a total of 25,160 households), while the remaining 57 percent of households stopped receiving assistance in 2024.
14. **Discontinuation:** By definition, households that have been selected for continued assistance and households that have been discontinued are different and are therefore generally not comparable to each other. By *simulating* the application of each of the four targeting methods to households, it is possible to assess under which methods a household would have been deemed eligible for assistance – that is, where a household would have fallen within the 43 percent most vulnerable. This simulation produces groups of households – called “eligibility spaces” – based on the methods that they are eligible and ineligible under. Within these eligibility spaces, the random variation of the *actual* targeting method allows the rigorous comparison of continued and discontinued households.
15. **Long notice:** From the set of 33,341 households that were selected for discontinuation, 960 households were randomly selected to receive a six-month extension of the programme. This second randomization created three groups of households: (1) the “continued” assistance households, who continued to receive regular assistance throughout the study period; (2) the “discontinued-short notice (short notice)” assistance households who received their notice of discontinuation just before the end of assistance; and the “discontinued-long notice (long notice)” households who received a notice of discontinuation six months before assistance would end.

Main findings

Prioritizing among households already being supported when scaling down

16. **The four targeting approaches do not lead to substantial differences in terms of average food security, psychological well-being, social cohesion, and socioeconomic indicators among the target population based on their overall averages.** These

findings are consistent with the limited evidence from experimental assessments on targeting, which generally show that, while different approaches lead to the *selection* of different groups, such variation does not necessarily translate into differences in programme effectiveness as measured by the average well-being of the target population.

17. **Vulnerability scoring based on categorical targeting underperforms with respect to the three PMT-based approaches – that aim to predict food insecurity based on FCS, per capita expenditure, and CARI, respectively – in terms of higher appeal rates and their selection of more vulnerable households.** Appeal behaviour is used as a proxy for acceptance of the discontinuation decision and is observed for the entire population of interest. The appeal rate among households discontinued under *Categorical* targeting was substantially higher than the rates of the PMT methods. This indicates that reliance on coarse sociodemographic criteria, rather than the more granular information embedded in PMT methods, can generate greater targeting errors and reduce the perceived legitimacy of eligibility decisions. Comparisons of baseline wealth indicate that households discontinued under the Categorical approach are more likely to be asset-poor than those excluded under PMT-based methods. Consistent with this evidence, when comparing continued versus discontinued households, the Categorical method behaved differently from all three PMT approaches. The PMT methods appeared more likely to retain households that were worse-off – according to rCSI, food expenditure, and worriedness indicators – whereas the Categorical approach did not show this pattern.
18. **The probability of appealing the discontinuation decision increases with the number of methods a household qualifies under.** Discontinued households that were deemed eligible by multiple other methods were more likely to appeal their exclusion. In the reference group – discontinued households not identified as eligible by any method – exactly 50 percent submitted an appeal. This rate rises to 62 percent among discontinued households deemed eligible under three methods. This pattern indicates that combining targeting approaches may offer a promising path to improving the precision and perceived fairness of eligibility decisions.

The consequences for households cut from assistance

19. **In the short term (six months), the discontinuation of food assistance leads to lower food consumption, greater reliance on harmful coping strategies (including borrowing or purchasing food on credit), declines in life satisfaction and mental well-being, and reduced trust in aid agencies.** While some of these effects are reduced in the medium term (12 months), reductions persist in the consumption of specific food groups included in the WFP basket, such as pulses. Discontinued households partially compensate for the loss of assistance by increasing food expenditure, particularly on items previously provided by WFP. However, these increases are generally insufficient to fully offset the lost transfers. Consequently, over time, discontinued households increasingly resort to negative coping strategies. Psychological well-being also deteriorates, with lower life satisfaction persisting a year after discontinuation. Finally, trust in aid agencies fell sharply as a result of discontinuation and remained low 12 months after the end of assistance. Overall, while households exhibit some capacity to adjust, the withdrawal of assistance generates substantial and lasting adverse effects on socioeconomic conditions and mental well-being.

20. **More vulnerable households suffer the most from the end of assistance.** When comparing impacts on different segments of the population, the results show that more vulnerable households are also more severely affected by the discontinuation of assistance. For example, the impacts on economic well-being were greater for discontinued households that were deemed vulnerable under three targeting methods. These households are less able to increase spending on food to compensate for the end of assistance. They are more likely to rely on harmful coping strategies and have lower food security. The findings highlight how socioeconomic constraints limit the ability of households to adapt to the end of programmes.

Whether and how households prepare for programme exit

21. **Early notice does not mitigate the negative impact of the end of assistance.** The “long notice” analysis shows no clear evidence that a longer warning window helped households prepare for the end of assistance and improve later outcomes. While long notice households behaved similarly to the households that continued to receive assistance in the short run, they converged with discontinued households by the end of the study period – when their assistance ended, they exhibited food security, expenditure, coping behaviour, and psychological well-being outcomes very similar to households who lost assistance without advance notice. These results indicate that households did not manage to smooth consumption in anticipation of the shock. This was likely due to limited ability to save or store resources. This underscores the need for more effective strategies to support households transitioning off humanitarian assistance in fragile settings, such as in Lebanon.

Considerations for programming

22. Based on the evidence, the study provides four key considerations for future humanitarian programming: combining methods for more precise targeting; assessing trade-offs of targeting precision; developing exit strategies; and more careful support for phasing out assistance.
23. **Combining targeting methods to improve accuracy and fairness:** The analysis shows that greater targeting accuracy can be achieved by combining multiple targeting methods and prioritizing households that are consistently identified as vulnerable across approaches. Institutionalizing this approach in WFP operations – as already piloted in Lebanon with new programmes based on these results – could enhance both fairness and effectiveness in future scale-up and scale-down efforts.
24. **Assessing trade-offs between precision with practicality:** In the context of a global contraction in humanitarian funding, accurately targeting the most vulnerable households is particularly critical. While the evaluation findings indicate that combining multiple targeting methods can improve selection precision, this may be costly to implement. Categorical targeting is generally easier to apply in contexts where extensive survey data is lacking or data collection is infeasible. However, the study finds that this approach produces larger well-being gaps between continued and discontinued households, and it is associated with higher rates of appeals. Policymakers need to balance the lower precision of categorical targeting with the potentially higher cost of conducting PMT-based targeting.

When more detailed household data is available, combining a prediction model approach with the simpler categorical scoring is expected to improve selection precision.

25. **Developing more deliberate and evidence-based exit strategies to maintain household well-being:** Even when assistance is scaled back after the peak of an emergency, and efforts are made to retain the most vulnerable households, discontinued households still experience significant short-term and medium-term adverse effects. This points to a broader structural issue in the design of humanitarian (and development) programmes, which wind down once the emergency subsides under the assumption that households can sustain their well-being independently. In practice, however, most programmes lack clear exit strategies. Developing more deliberate and evidence-based exit strategies could reduce the risk of harm when assistance ends and help to ensure that the gains made during the programme are not reversed.

26. **More careful support to transition households towards the end of assistance:** The evaluation finds that simply providing advance notice of discontinuation does little to help households smooth consumption or adjust effectively. This indicates the need for alternative approaches that minimize the disruption caused by programme exit. Also, when assistance ends abruptly, trust in humanitarian actors deteriorates sharply. Implementing interventions to support the transition may help to avoid this.

1. Introduction

27. Over the past two decades, humanitarian assistance has faced persistently large, and increasingly widening, funding gaps. In 2024, the United Nations raised only 46 percent of the USD 49.6 billion required to meet global humanitarian needs, leaving millions without essential support.⁴ Funding for humanitarian programmes is also volatile, subject to abrupt reallocation due to emerging crises or political shifts, as recently illustrated by dramatic funding cuts from the United States Government. Against this background, the World Food Programme (WFP) and other humanitarian actors are regularly confronted with critical decisions regarding how to best allocate/reallocate and prioritize their scarce resources.
28. The WFP Lebanon Country Office (LBCO) exemplifies this challenge. Since the beginning of Lebanon's socioeconomic crisis in late 2019 – which led to skyrocketing inflation, currency depreciation, and rising food insecurity – the LBCO has been providing food assistance to vulnerable Lebanese households, reaching more than 5 percent of the Lebanese population in 2023.⁵ However, funding shortages paired with shifting priorities led the LBCO to reduce assistance coverage by over half in 2024. Like other actors facing similar challenges, the LBCO needed to identify an effective way to prioritize assisted households to select the most vulnerable and minimize the negative consequences of the programme cuts.
29. There is a rich body of research that focuses on the design and initial implementation of development and (to a smaller extent) humanitarian programmes.⁶ This literature often relies on rigorous experiments to test variations in the programmes (e.g., providing cash vs. in-kind; lump-transfer vs. regular transfers; targeting men vs. women) and to assess their impact on a wide range of household welfare measures.^{7,8} A subset of this literature focuses on testing alternative strategies to select the initial target population. Taken together, these studies provide helpful insights that can inform policymakers on how to best design and implement *new* assistance programmes. However, there is very little evidence overall – and no experimental evidence – on how to approach the *end* of assistance programmes. How should organizations implement the cuts in a way that minimizes disruption to household well-being? What are the consequences on households of the discontinuation of assistance? How do households cope with such cuts? These are questions that are still without clear answers.

⁴ UNOCHA. 2024. *Global Humanitarian Overview 2024*. <https://www.unocha.org/attachments/b9488007-b44e-4a35-9827-6a9f8488194e/10%20GHO%202024%20October%20Update.pdf>

⁵ Author's estimation based on the number of households in the programme and average household size.

⁶ Jeong, D., & Trako, I. 2022. *Cash and In-Kind Transfers in Humanitarian Settings: A Review of Evidence and Knowledge Gaps* (No. 10026; Policy Research Working Paper). World Bank Group. <https://openknowledge.worldbank.org/entities/publication/2152a365-c73f-5bfd-9070-6f1fa3a81f03>

⁷ Crosta, T., Karlan, D., Ong, F., Rüschenpöhler, J., & Udry, C. R. 2024. (Working Paper No. 32779). National Bureau of Economic Research. <https://doi.org/10.3386/w32779>

⁸ Kondylis, F., & Loeser, J. 2024. Unconditional Cash Transfers: A Bayesian Meta-Analysis of Randomized Evaluations in Low and Middle Income Countries. NBER Working Paper Series. National Bureau of Economic Research. <http://johnloeser.github.io/assets/kl.pdf>

30. To shed light on these questions, the Office of Evaluation and the International Security and Development Center (ISDC) conducted an impact evaluation in partnership with the LBCO. The study follows a randomized controlled trial (RCT) design and is structured in three key parts:
- The **first** part tackles the key question of how to implement programme cuts in a way that minimizes welfare losses among the communities receiving food assistance.
 - The **second** part of the study investigates the consequences of ending assistance for households that were excluded by the prioritization implemented in the first part.
 - The **third** step of the analysis investigates whether and how households prepare for the end of the assistance, by providing a longer “notice window” to a random subset of households that were excluded from assistance by the prioritization exercise.

2. Country context

31. The first sign of what would become Lebanon's complex economic, political, and social crisis occurred in August 2019 when banks withheld deposits in US dollar accounts. Set in motion by long-standing financial mismanagement, the crisis was severely compounded by two major shocks: the COVID-19 pandemic, and the Beirut Port explosion in August 2020. The port explosion killed nearly 200 people, injured more than 6,500, and displaced 300,000 individuals. It also destroyed critical infrastructure, including the country's primary grain silos.⁹ Concurrently, the COVID-19 pandemic led to stringent containment measures that further deepened the economic downturn. This resulted in widespread business closures, job losses, and a severe strain on Lebanon's already fragile health care system.¹⁰
32. This extraordinary combination of adverse events led to the most devastating, multipronged crisis in Lebanon's modern history. This heightened social tensions and exacerbated existing inequalities. In 2020, as a response to these shocks, the LBCO started providing emergency assistance in the form of monthly in-kind food parcels to Lebanese households in need. The first households started receiving their parcels in July 2020 and the pool of households progressively expanded to reach about 5 percent of the Lebanese population by 2023.
33. During the evaluation period, the country experienced yet another dramatic shock with the escalation of tensions with Israel along the southern border. Between September and November 2024, the conflict resulted in more than 3,900 fatalities and 16,500 injuries.¹¹ The conflict also resulted in extreme levels of displacement with more than a quarter of the population in Lebanon evacuated – both Syrians and Lebanese.^{12,13} Although a ceasefire agreement was signed on 29 November 2024, the tensions along the southern border continue.¹⁴

⁹ WFP. 2021. *Lebanon Annual Country Report 2020*. <https://www.wfp.org/publications/annual-country-reports-lebanon>

¹⁰ WFP. 2023c. *Lebanon Annual Country Report 2022*. <https://www.wfp.org/publications/annual-country-reports-lebanon>

¹¹ Integrated Food Security Phase Classification. 2025. [IPC Acute Food Insecurity Analysis October 2024–March 2025: Lebanon](#).

¹² Ibid.

¹³ International Organization for Migration. 2024. Mobility Snapshot - Round 67 - 12-12-2024. Global Data Institute Displacement Matrix. <https://dtm.iom.int/datasets/mobility-snapshot-round-67-12-12-2024>

¹⁴ The impact evaluation period began in August 2023 and analysis concluded in January 2026, before the resumption of hostilities between Israel and Lebanon in February 2026.

3. Programme description

3.1. Programme background

34. The WFP LBCO's food assistance programme started in 2020. It was designed to help Lebanese households mitigate the effects of the severe socioeconomic crisis, which eroded their ability to meet basic needs. The food assistance programme is implemented in close collaboration with the national social protection system in Lebanon, which includes the Emergency Social Safety Net and the National Poverty Targeting Programme.¹⁵
35. To identify target beneficiaries, the LBCO initially relied on lists of vulnerable households provided by community partners and the government, along with households that self-registered through a dedicated online portal or hotline. The LBCO then validated the inclusion of each household by conducting a survey. Inclusion in the programme was initially based on the household's Food Consumption Score (FCS) and a phone survey of reliance on coping strategies. For later rounds, inclusion was based on a categorical index designed to capture multidimensional deprivation, collected through an in-person survey.
36. Starting in 2020, the programme was progressively expanded to new households, and by 2023 the LBCO was providing food assistance to 73,520 Lebanese households across the country. (Appendix A1 presents more detail on the in-kind programme with Table A1.1 providing the number of households per governorate and per district enrolled in the programme.)
37. Households included in the programme are guaranteed the provision of monthly food assistance at least for the duration of an "assistance cycle" (12 months). Assistance is then renewed based on funding availability.¹⁶ Each household receives one food parcel for every two household members (up to a maximum of three parcels). Each parcel includes about 20 kg of commodities and is worth about USD 32 (based on January 2023 market prices). This corresponds to about 30 percent of the monthly food expenditure and 13 percent of total expenditure for the median Lebanese household.¹⁷ Each parcel includes ten items that are meant to provide at least 1,300 kcal per person per day (details on the parcel composition are available in Table A1.2).

Funding constraints and retargeting

38. Funding constraints and a shift in priorities led the LBCO to reduce its assistance coverage for the 2024 cycle. Consequently, it decided to conduct a retargeting of households in two phases: the first phase involved 80 percent of the beneficiary population and took place in January 2024; the second phase involved the remaining 20 percent of the population and took place later in July.¹⁸

¹⁵ These two initiatives were merged under Lebanon's national cash assistance AMAN programme in 2025.

¹⁶ Between 2020 and 2023 coverage of the programme progressively expanded and assistance was continuously renewed.

¹⁷ Authors' estimation using the nationally representative 2023 Lebanon Household Survey (LHS) data.

¹⁸ The July re-targeting comprised 15,019 households which qualified for a longer assistance period due to their later assessment by the LBCO.

39. The evaluation focuses on the January retargeting for which the LBCO aimed to reduce the number of beneficiaries from 58,501 households to 25,160, or about 43 percent of the existing caseload¹⁹ – a target based on the LBCO’s estimation of the number of households with consumption levels below the Minimum Expenditure Basket (MEB), which was set at LBP 3,238,315 or USD 71 per capita per month (January 2023 prices). Figure 1 shows the geographic distribution of the 58,501 households across Lebanon.
40. At the end of the 2023 cycle (December 2023), all households included in the January retargeting phase received an update on their status for the 2024 cycle (which started in January 2024). Communication was conducted entirely via SMS to ensure timely, direct, and consistent messaging across all households. The message either confirmed that assistance would continue (at least) throughout 2024 or that assistance would end due to limited funding.^{20,21} Within the latter group, the large majority of households (97 percent) were informed that assistance would end by December 2023, while a subset of 960 households (who were part of the “long notice” sample – see below for more details) were informed that assistance would continue until June 2024.²² These 960 long notice households were reminded through a second SMS in April 2024 that their last round of assistance would be in June 2024.²³ Following WFP practice, the message also mentioned the possibility to appeal the decision through a dedicated call centre or through an online form. When an appeal was submitted, the LBCO reviewed the case to determine whether the household should be reassessed and readmitted.²⁴

¹⁹ This estimation was originally based on extrapolation from the nationally representative LHS dataset. This figure aligned closely with findings from the post-distribution monitoring survey that the LBCO conducted in the first half of 2023 among a representative sample of households receiving in-kind food assistance, which indicated that 44 percent were below the MEB threshold.

²⁰ The message followed standard practice within the LBCO and contained short generic information on the households’ inclusion or exclusion from the programme. Households excluded from assistance stopped receiving the regular SMS to collect parcels and were not included in any alternative programme. The exact wording of the messages is reported in Annex A10.

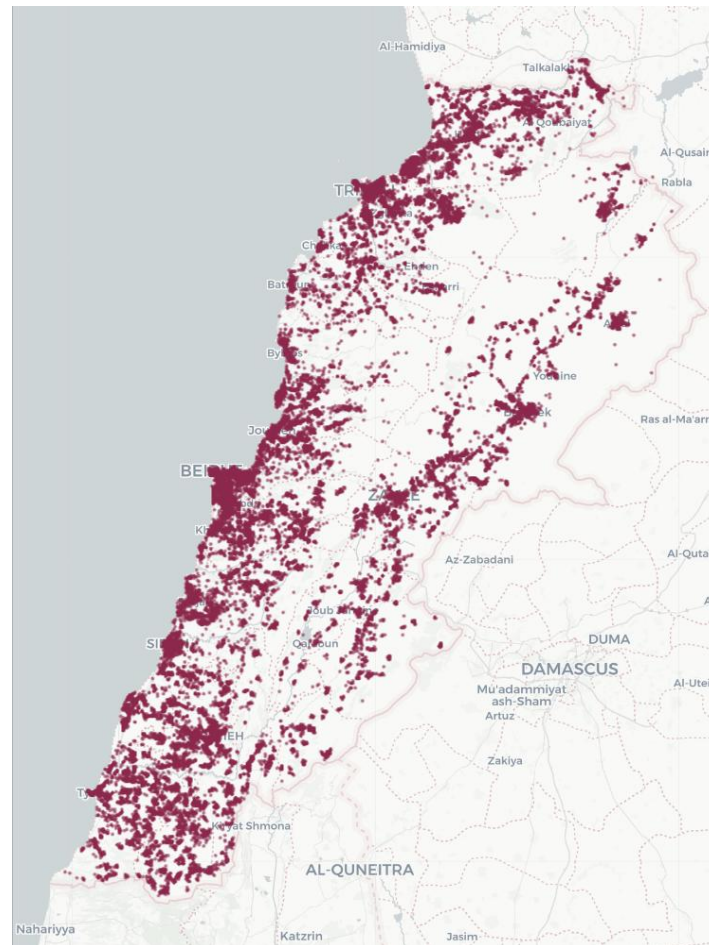
²¹ Adherence to the study protocol is confirmed by the administrative data collected by the LBCO and the household survey collected through the study (see Figure A1.1 in Annex A1).

²² For the message wording, see Annex A10.

²³ For the message wording, see Annex A10.

²⁴ About 52 percent of excluded households submitted an appeal, but the LBCO had only changed its decision in less than 1 percent of the cases by the time of the June data collection.

Figure 1: Distribution across Lebanon of households in the WFP in-kind food assistance programme



Source: Evaluation team

2024 Conflict

41. During the study period, the conflict between Lebanon and Israel along the southern Lebanese border escalated. This led to a ground invasion in September 2024, intense bombing across the country, and population displacement.²⁵ The conflict resulted in thousands of casualties and large-scale displacement.^{26,27}
42. In response to this crisis, the LBCO started providing emergency cash transfers to households located in conflict-affected areas, (as defined by conflict or displacement intensity), that were not already supported through its food assistance programme.²⁸ Each household selected for cash assistance was entitled to receive food assistance amounting to USD 20 per household member, for a maximum of six household members, and non-

²⁵ Figure A1.2 in Annex A1 provides a map of sub-districts affected by the conflict between July to December 2024.

²⁶ Raleigh, C. et al. 2010. Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research*, 47(5), 651-660. <https://academic.oup.com/jpr/article-abstract/47/5/651/8366127?redirectedFrom=fulltext>

²⁷ Figure A1. 2 in the Annex visualizes the geographic spread of the conflict.

²⁸ A given area (cadastre) was deemed priority area if it fulfilled any of the following three criteria: (i) it has a value of 3 or higher on the LBCO's Conflict Impact Index; (ii) it hosts at least 500 internally displaced persons; or (iii) more than 20 percent of the local population are internally displaced. Figure A1.3 in Annex A1 provides a map of the cadastre-level prioritization in practice.

food assistance of USD 25 for the whole household.²⁹ Emergency cash assistance began in October 2024 and continued for at least three months, with some locations receiving it as late as May 2025. Because of this situation, 74 percent of the households assigned to stop receiving WFP food assistance ended up receiving three months of emergency cash support by the time of the final data collection. While the emergency cash transfer is not part of this evaluation, it affected the original study design by likely lowering the contrast between continued and discontinued assistance households in the second part of the study, which is important to keep in mind when interpreting the medium-term results.

3.2. Targeting approaches

43. The LBCO had relied on different approaches for initial programme targeting – outcome-based targeting using FCS and the Multidimensional Deprivation Index (MDDI) – but was still uncertain about the optimal way to assess and target specific vulnerabilities. Addressing these questions was critical as the LBCO aimed to minimize welfare losses in the population. In practice, this meant ensuring continued assistance to the most vulnerable households, who were expected to be (relatively) less able to cope with the end of assistance.
44. To operationalize this objective, the Office of Evaluation, ISDC, and the LBCO jointly identified and selected four commonly used targeting methods built on the vulnerability-based targeting that WFP uses and that had been implemented by the LBCO. These targeting approaches are designed to identify the most vulnerable households within a given population. While they are mainly used to determine initial eligibility when a programme is introduced, they can also be applied in retargeting or prioritization exercises to decide which households should continue to receive assistance or exit the programme. These four approaches are described in detail below.
45. **PMT targeting based on per-capita expenditure vulnerability (PMT-Exp or method 1):** Eligible households are identified through a PMT model that determines households with lower per-capita monthly expenditure. This information is gathered through a detailed set of more than 100 potential questions, including food and non-food related expenditure. Food-related questions are designed to record purchased food but also any other source of food, including transfers and own production. Targeting based on per-capita expenditure is one of the most common approaches used across low-and middle-income settings, where direct measures of poverty are difficult to gather. This PMT model is therefore one of the most popular targeting approaches and is widely adopted by WFP and other organizations. It will later be referred to as the reference method in the empirical section of this study.
46. **PMT targeting based on food consumption vulnerability (PMT-FCS or method 2):** Eligible households are identified through a PMT model that focuses on households with lower food consumption, as measured through the FCS. The FCS, commonly used by WFP and other international organizations, is built using nine survey questions on the frequency of food groups consumed by the households in the past seven days and has a value

²⁹ The maximum market value of the in-kind food assistance package in 2024 for a household of six people was USD 95.01, and the maximum value of the emergency cash assistance was USD 145.

between 0 and 112. This targeting approach might be considered the most obvious choice for food assistance programmes, given that it targets food-related vulnerability.

47. **PMT targeting based on a multidimensional vulnerability index (PMT-CARI or method 3):** Eligible households are found through a PMT model designed to identify households with a combination of low expenditure per-capita, low food security, and low capacity to cope with shocks (as captured through a set of 16 survey questions). These three dimensions are combined into a single measure, called the Consolidated Approach for Reporting Indicators of Food Security (CARI). In practice, CARI, commonly used across WFP and other international organizations, assigns households to one of four food security levels: food secure; marginally food secure; moderately food insecure; and severely food insecure. This study groups together the moderately and severely food insecure categories to identify food insecure households. This targeting approach largely combines the first two approaches, accounting for both permanent (economic capacity) and transient (food consumption) dimensions of vulnerability.
48. **Categorical vulnerability scoring (Categorical or method 4):** Eligible households are identified using a score based on vulnerability criteria recorded through a survey administered to 28 local experts and practitioners from Lebanon. The experts were asked to indicate (in order of importance) their top-10 criteria for identifying and selecting Lebanese households in need of food assistance. The list of potential criteria mirrored the variables available for the PMT models. The answers were then combined, and weights were attributed to each criterion, based on their frequency and ranking order. Each household was then assigned a score based on these weights.
49. The three PMT models were developed (“trained”) using data from the Lebanon Household Survey (LHS), collected between January and May 2023. This dataset is representative for most of the country³⁰ and includes hundreds of variables, including those required to construct the three PMT outcomes of interest: monthly per capita expenditure, FCS, and CARI. The models were then built using a pool of 126 different potential variables (“predictors”).³¹ These variables were selected based on their inclusion in the WFP household census,³² which covered the full sample of 58,501 households and was conducted by the LBCO just a few months later in August and September 2023. Although the census did not include the variables necessary to compute the PMT outcomes, it served as the foundation for implementing the prioritization process. Further details on the four targeting approaches, along with simulations of targeting errors and predictions using the LHS dataset, can be found in Appendix A2.

³⁰ The LHS did not include in-person interviews in three governorates – Baalbek El-Hermel, El Nabatieh, and South Lebanon – due to security issues during data collection.

³¹ Considering interactions and squared terms, the set of potential predictors used to develop the PMT models reached close to 300 variables.

³² The LBCO refers to the census as the Census Phase 3 (CP3) survey.

4. Impact evaluation motivation and design

4.1. Motivation

50. This evaluation aims to contribute to two evidence gaps related to the prioritization and effectiveness of humanitarian assistance: the targeting and prioritization of households: and the consequences of ending assistance.

Targeting/prioritization of households

51. How to best identify and select households that need assistance is a constant question for assistance programmes. Over recent decades, PMT has become the most popular targeting approach across low-capacity and fragile settings where direct measures of poverty and vulnerability are typically unavailable. The PMT approach builds on a statistical model of proxy indicators to predict the desired vulnerability measure. It gained increasing popularity on the premise of being objective and difficult to manipulate.³³
52. However, its accuracy and reliability heavily depend on regular, large-scale data collection to supply the underlying data used to develop the model,³⁴ which makes it complicated to implement and communicate.³⁵ In recent years, the traditional PMT approach has been enriched with innovative techniques leveraging machine learning algorithms, which has led to significant improvements in its ability to predict target outcomes.^{36,37,38,39,40,41,42} Despite its popularity, few experimental studies exist on how PMT compares with other targeting

³³ Follett, L., & Henderson, H. 2023. A Hybrid Approach to Targeting Social Assistance. *Journal of Development Economics*, 160, 103002. <https://doi.org/10.1016/j.jdeveco.2022.103002>

³⁴ Aiken, E., Ohlenburg, T., & Blumenstock, J. 2023. *Moving Targets: When does a Poverty Prediction Model need to be Updated?* Proceedings of the 6th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies, 117–117. <https://doi.org/10.1145/3588001.3609369>

³⁵ Cameron, L., & Shah, M. 2014. Can Mistargeting Destroy Social Capital and Stimulate Crime? Evidence from a Cash Transfer Programme in Indonesia. *Economic Development and Cultural Change*, 62(2), 381–415. <https://doi.org/10.1086/674102>

³⁶ Follett, L., & Henderson, H. 2023. A Hybrid Approach to Targeting Social Assistance. *Journal of Development Economics*, 160, 103002. <https://doi.org/10.1016/j.jdeveco.2022.103002>

³⁷ 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.

³⁸ Araar, A., & Tiberti, L. 2020. *Optimal Targeting and Poverty Reduction with Fixed Budget and under Imperfect Information*. https://dasp.ecn.ulaval.ca/dasp3/refs/Optimal_Targeting_and_Poverty_Reduction_april2020.pdf

³⁹ Gram-Hansen, B. J., Helber, P., Varatharajan, I., Azam, F., Coca-Castro, A., Kopackova, V., & Bilinski, P. 2019. Mapping Informal Settlements in Developing Countries using Machine Learning and Low Resolution Multi-spectral Data. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 361–368. <https://doi.org/10.1145/3306618.3314253>

⁴⁰ Caria, S. et al. 2020. An Adaptive Targeted Field Experiment: Job Search Assistance for Refugees in Jordan. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3689456>

⁴¹ Aiken, E. et al. 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>

⁴² Hall, O. et al. 2023. A Review of Machine Learning and Satellite Imagery for Poverty Prediction: Implications for Development Research and Applications. *Journal of International Development*, 35(7), 1753–1768. <https://doi.org/10.1002/jid.3751>

methods when selecting recipients. Even less is known about how to reduce recipients, or how different PMT models measure differences in programme impact.^{43, 44}

53. The studies that have investigated using alternative targeting approaches on the well-being of the target population have typically compared PMT against other approaches, but not different PMT modelling choices against each other. PMT methods exhibit slightly greater precision in identifying vulnerable households compared to the most common alternatives – categorical or community-based targeting approaches – but this does not create significant differences in overall programme impact.^{45,46,47,48,49}
54. Two studies compare PMTs targeting poverty to alternative data-driven targeting methods. The first study compares targeting household expenditure (predicted through a standard PMT model), to targeting FCS (directly recorded through survey data), and finds no difference in programme impact.⁵⁰ The second study compares a standard PMT approach targeting households' level of deprivation/poverty to a new PMT approach that targets households' expected benefits from the programme, as identified through panel data and a "random forest" algorithm.⁵¹ It finds that the latter approach does lead to higher programme impact, although this might come at the cost of excluding some of the most vulnerable households.

The consequences of ending assistance

55. There is a rich body of studies that investigate the effectiveness of cash and in-kind assistance across a variety of settings. A large portion of this evidence is based on experimental studies that randomize enrolment into assistance programmes across locations and/or populations. A recent review of cash assistance programmes identifies as many as 114 experimental studies from low- and middle- income settings.⁵² There is also some – albeit more limited – evidence coming from cash or in-kind assistance programmes in humanitarian settings, with a recent review identifying a total of nine experimental and

⁴³ Altındağ, O., O'Connell, S. D., Şaşmaz, A., Balçioğlu, Z., Cadoni, P., Jerneck, M., & Foong, A. K. 2021. Targeting Humanitarian Aid using Administrative Data: Model Design and Validation. *Journal of Development Economics*, 148, 102564. <https://doi.org/10.1016/j.jdeveco.2020.102564>

⁴⁴ In response to this evidence gap, the Office of Evaluation has made generating insights on how to optimize the targeting and prioritization of beneficiaries a central focus of its [Humanitarian Workstream](#). See also initial results from a [targeting impact evaluation conducted in partnership with the World Bank in the DRC](#).

⁴⁵ Alatas, V. et al. 2012. Targeting the Poor: Evidence from a Field Experiment in Indonesia. *American Economic Review*, 102(4), 1206–1240. <https://doi.org/10.1257/aer.102.4.1206>

⁴⁶ Beaman, L. et al. 2021. Can Network Theory-Based Targeting Increase Technology Adoption? *American Economic Review*, 111(6), 1918–1943. <https://doi.org/10.1257/aer.20200295>

⁴⁷ 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>

⁴⁸ Stoeffler, Q., Mills, B., & Del Ninno, C. 2016. Reaching the Poor: Cash Transfer Program Targeting in Cameroon. *World Development*, 83, 244–263. <https://doi.org/10.1016/j.worlddev.2016.01.012>

⁴⁹ Hillebrecht, M., Klonner, S., & Pacere, N. A. 2023. The Dynamics of Poverty Targeting. *Journal of Development Economics*, 161, 103033. <https://doi.org/10.1016/j.jdeveco.2022.103033>

⁵⁰ Schnitzer, P., & Stoeffler, Q. 2024. Targeting Social Safety Nets: Evidence from Nine Programs in the Sahel. *The Journal of Development Studies*, 60(4), 574–595. <https://doi.org/10.1080/00220388.2023.2291325>

⁵¹ Haushofer, J. et al. 2025. Targeting Impact versus Deprivation. *American Economic Review*, 115(6), 1936–1974. <https://doi.org/10.1257/aer.20221650>

⁵² Crosta, T. et al. 2024. Unconditional Cash Transfers: A Bayesian Meta-Analysis of Randomized Evaluations in Low and Middle Income Countries (Working Paper No. 32779). National Bureau of Economic Research. <https://doi.org/10.3386/w32779>

12 non-experimental studies.⁵³ These studies provide fairly rich evidence on the expected benefits of transfer programmes, and on the consequences of different programme design choices.

56. However, the existing evidence provides little guidance for contexts where programmes are to be cut. There are no experimental studies that assess the consequences of ending assistance on households, or investigate whether and how households prepare for the end of assistance. Even among non-experimental studies, the evidence is limited. WFP's own programme reviews and reports discuss the consequences of reductions in assistance only in a few instances. A summary of evaluation evidence found negative consequences in food and nutrition outcomes among those whose assistance was discontinued in four cases (Kenya, Rwanda, Benin, and the Central African Republic),⁵⁴ while in one instance (Nigeria), a 70-percent cut in rations was reported to have had no effect on similar food security outcome indicators.⁵⁵ Some studies have also simulated the impact of the recent funding cuts on household well-being and the local economy in countries such as Uganda,⁵⁶ Sudan,⁵⁷ Somalia,⁵⁸ and the Democratic Republic of the Congo.⁵⁹ There is only one (non-experimental) study that assesses the impact of a cut in cash assistance: by taking advantage of variation in the timing of the household surveys, the study finds a large drop in food security and expenditure among refugees in Kenya.^{60,61,62} The overall lack of rigorous evidence on the consequences of cutting assistance programmes is concerning, as organizations are regularly scaling down or ending their programmes in response to changes in priorities and funding. This study contributes to filling this gap by providing the first experimental evidence of the consequences of ending a humanitarian food assistance programme on household well-being.

⁵³ Jeong, D., & Trako, I. 2022. Cash and In-Kind Transfers in Humanitarian Settings: A Review of Evidence and Knowledge Gaps (No. 10026; Policy Research Working Paper). World Bank Group.

<https://openknowledge.worldbank.org/entities/publication/2152a365-c73f-5bfd-9070-6f1fa3a81f03>

⁵⁴ WFP. 2025. *Summary of Evaluation Evidence: Prioritization (2021–2023)* [WFP Evidence Summary].

<https://docs.wfp.org/api/documents/WFP-0000165455/download/>

⁵⁵ WFP. 2023a. *Evaluation of Nigeria WFP Country Strategic Plan 2019–2022* [Centralized Evaluation Report].

<https://docs.wfp.org/api/documents/WFP-0000146137/download/>

⁵⁶ Kagin, J. et al. 2024. *The Cost of Inaction: Impacts of WFP Refugee Assistance Shortfalls on Food Security Outcomes in Uganda*. World Food Programme. <https://hdl.handle.net/10919/121144>

⁵⁷ WFP. 2023b. *Impacts of the Cost of Inaction on WFP Food Assistance in Sudan (2021 & 2022)*.

<https://reliefweb.int/report/sudan/impacts-cost-inaction-wfp-food-assistance-sudan-2021-2022-april-2023>

⁵⁸ Kagin, J. et al. 2024. *The Cost of Inaction: Impacts of WFP Assistance Shortfalls on Food Security Outcomes in Somalia*. World Food Programme. <https://hdl.handle.net/10919/121145>

⁵⁹ WFP. 2023d. *The Cost of Inaction: The Consequences of Not Being Able to Reach Those Most in Need in Eastern DRC*.

<https://reliefweb.int/report/democratic-republic-congo/cost-inaction-consequences-not-being-able-reach-those-most-need-eastern-drc?lang=ru>

⁶⁰ Sterck, O., & Bruni, V. (2025). *The Welfare and General-Equilibrium Impacts of Aid Cuts* (SSRN Scholarly Paper No. 5270846). Social Science Research Network. <https://doi.org/10.2139/ssrn.5270846>

⁶¹ A related study from the same setting uses the same variation to investigate the consequences of delays in the disbursement of assistance, also finding negative effects on various dimensions of households' well-being: Bruni, V., & Sterck, O. 2025. *The Welfare and Market of Effects of Delays in Humanitarian Assistance* (SSRN Scholarly Paper No. 5190878). Social Science Research Network. <https://doi.org/10.2139/ssrn.5190878>

⁶² In additional research from high-income countries, rich administrative data assesses the impact of social welfare cuts. Findings suggest that reductions or discontinuation of transfers influence labour force participation only in the short term, but have negative long-term effects on children's well-being and human capital accumulation, and may also lead to an increase in crime: Dustmann, C., Landersø, R., & Andersen, L. H. 2024b. [Refugee Benefit Cuts](#). *American Economic Journal: Economic Policy*, 16(2), 406–441.; Dustmann, C., Landersø, R., & Andersen, L. H. (2024d). Unintended Consequences of Welfare Cuts on Children and Adolescents. *American Economic Journal: Applied Economics*, 16(4), 161–185. <https://doi.org/10.1257/app.20230519>

57. Understanding the consequences of these cuts should provide an important input for the definition of appropriate “exit strategies” that could sustain improvements in household well-being and minimize disruption for households. A policy priority is to find ways to help households transition out of assistance while sustaining their well-being and consumption. This study takes advantage of the provision of a longer notice window (i.e. the time between the notice that assistance will end and its actual end) to understand whether and how households prepare for the drop in resources and try to smooth their consumption over time.⁶³

4.2. Evaluation questions

58. The immediate aim of this impact evaluation is to support the LBCO in identifying the best way to prioritize food assistance among households that were already part of its in-kind food assistance programme. The broader aim is to shed light on the consequences of cutting humanitarian food assistance and to understand how to best minimize disruption for discontinued households. To achieve these aims, this study is structured in three core parts, each designed to answer a different set of questions.

Part 1: Prioritizing among households already being supported when scaling down

59. The first part of the study compares the four targeting approaches and answers the following evaluation questions:
- Q1:** What are the differences across the four targeting approaches in terms of who gets selected to continue receiving assistance?
 - Q2:** Which targeting approach leads to higher acceptance (as proxied by the lowest number of formal appeals)?
 - Q3:** Which targeting method leads to overall higher well-being among the target population?
 - Q4:** Which types of households benefit the most from each targeting approach?

Part 2: The consequences for households excluded from assistance

60. The second part of the study compares households that keep receiving assistance and those that stop receiving it. It answers the following evaluation questions:
- Q5:** What is the impact of ending humanitarian food assistance on households’ well-being?
 - Q6:** Which types of households are most affected by the discontinuation of assistance?

Part 3: Whether and how households prepare for programme exit

61. The third part of the study compares the subset of households that received an extended notice window before assistance was ended and the other two household groups – those

⁶³This analysis relates to the permanent income hypothesis (PIH), which proposes that individuals base consumption on expected long-term income rather than current income. This implies that temporary changes in available resources should have limited effects on consumption if credit or savings are available. (Friedman, M. 1957, *The Permanent Income Hypothesis*, NBER Chapters, 20–37); -

that received continued assistance and those whose assistance ended (with a standard short notice period). It answers the following evaluation question:

Q7: Do households prepare for the end of assistance when they receive notice?

4.3. Evaluation design

62. This section provides a description of the specific design and analysis method used to address each evaluation question. The evaluation design was [pre-registered in the American Economic Association's registry](#) for randomized controlled trials. Details on the estimation methods are provided in Annex A3.

Part 1: Prioritizing among households already being supported when scaling down

Questions

Q1: What are the differences across the four targeting approaches in terms of who gets selected to continue receiving assistance?

Q2: Which targeting approach leads to higher acceptance (as proxied by the lowest number of formal appeals)?

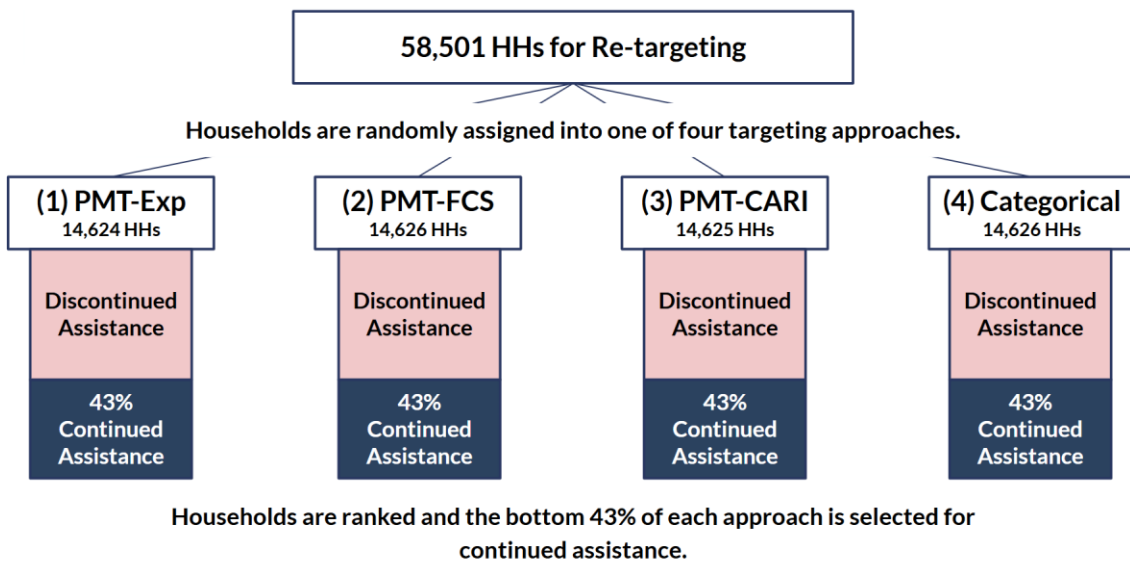
Q3: Which targeting method leads to overall higher well-being among the target population?

Q4: Which types of households benefit the most from each targeting approach?

63. To address the evaluation questions on **prioritization and targeting**, the 58,501 households in the study were individually randomly allocated into one of four equally sized study arms (~14,626 households per study arm). Within each arm, one of the four retargeting methods was applied to prioritize assistance.⁶⁴

⁶⁴ The key input was the rich census of all 58,501 beneficiaries that was collected by the LBCO between August and September 2023.

Figure 2: Random allocation of households into the four targeting approaches



Note: Own visualization. The figure displays how households were first randomly allocated into one of the four targeting approaches and then selected into (dis)continuation based on their vulnerability.

64. The 43 percent of households deemed most vulnerable in each arm were then selected to continue receiving assistance (6,290 households from each study arm, for a total of 25,160 households), while the remaining 57 percent of households stopped receiving assistance in 2024.⁶⁵ Figure 2 visualizes the design with the random allocation into the four targeting approaches.⁶⁶

Part 2: The consequences for households excluded from assistance

Questions

- Q5: What is the impact of ending humanitarian food assistance on households' well-being?
- Q6: Which types of households are most affected by the discontinuation of assistance?

65. The randomization can also be leveraged to address questions related to the impact of the **discontinuation (or cuts) of assistance**. For each household in the study population, the study *simulates* the application of each of the four targeting methods – the one the household was actually assigned to, plus the other three. In this way, it is possible to assess which of the four methods would have deemed a household eligible for assistance – that is, fallen within the 43 percent most vulnerable.

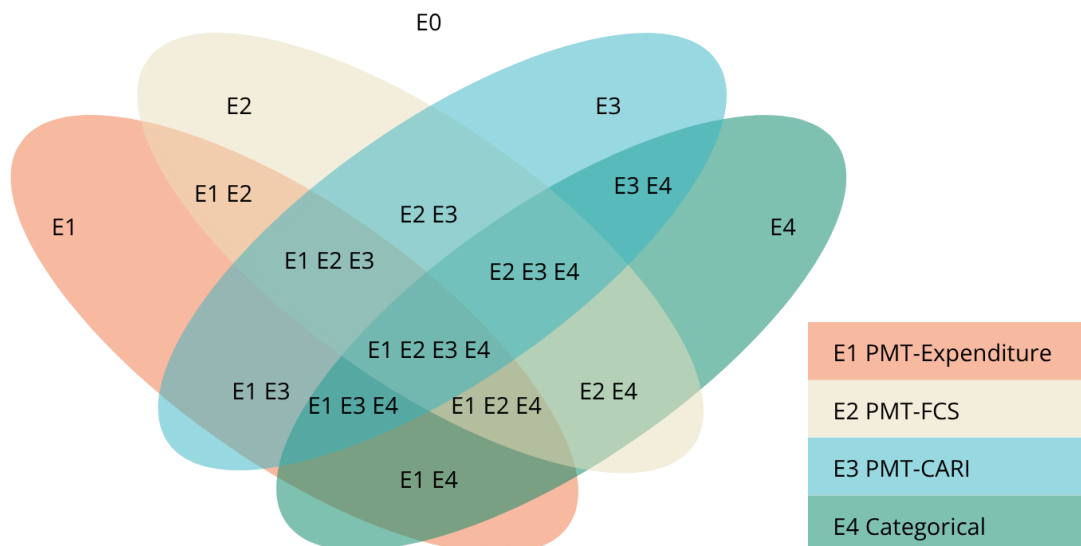
66. This simulation allows the placement of each household into an “eligibility space” as shown in the “eligibility map” in Figure 3, with the letter “E” signifying eligibility for continued assistance under the specified method. For instance, the eligibility space E1E2 represents

⁶⁵ Most households excluded from assistance stopped receiving it in January 2024. An exception to this was a randomly selected subset of 960 households, which received an additional six months of assistance (until June 2024). More details are reported when discussing Part 3 of the design.

⁶⁶ Figure A4.1 in Annex 4 shows the geographical distribution of households in the targeting arms, and Figure A4.2 shows the geographical distribution of households selected for continuation of in-kind assistance.

the set of households that would have been eligible for continued assistance if targeting method 1 (PMT-Exp) or 2 (PMT-FCS) had been applied, but that would have stopped receiving assistance if methods 3 (PMT-CARI) or 4 (Categorical) had been used instead. Overall, there are 16 eligibility spaces⁶⁷ generated by all possible combinations of outcomes, including E0 which denotes the space for households that are never deemed eligible regardless of the method applied.

Figure 3: Eligibility map



Note: Own visualization. The figure illustrates all possible combinations of eligibility a household could fall under.

67. The random allocation of the *actual* targeting method across the population, thus generates random variation in the discontinuation of assistance for households located within the same eligibility space. For example, the random allocation of households into one of the four methods implies that households in space E1E2 that were randomly allocated to targeting method 1 or 2 will keep receiving assistance, while those randomly allocated to method 3 or 4 will not.⁶⁸ Such random variation provided the opportunity to study the consequences of the end of assistance, by comparing similar households that belong to the same eligibility space.

Part 3: Whether and how households prepare for programme exit

Question

Q7: Do households prepare for the end of assistance when they receive notice?

⁶⁷ Households are spread across all spaces, with E1E2E3 being the smallest one with 458 households. Eligibility spaces E0 (never eligible households) and E1E2E3E4 (always eligible households) combined cover about 39 percent of the population. Annex 4 Table A4.1 shows the distribution of all households across the 16 eligibility spaces and four targeting methods.

⁶⁸ Random allocation into the four study arms was stratified by eligibility space to ensure balance across this dimension, given that some eligibility spaces were relatively small. However, one should expect exactly a fourth of the households within each eligibility space to be allocated to each targeting arm, even in the absence of stratification.

68. To answer the last question **on how households prepare for discontinuation**, a second randomization was performed within the set of 33,341 households that were excluded from assistance. These households were expected to be informed in December 2023 (i.e. at the end of the 2023 cycle) that their assistance would not continue in January 2024. A random subset of 960 households was selected to receive a six-month extension of the programme until June 2024.⁶⁹ These 960 “long-notice” households were informed at the end of the 2023 cycle that they would stop receiving assistance after June 2024. A reminder message was sent to them in April 2024 with the same notice.
69. This second randomization creates three groups of households:
- the “continued” assistance households, who continued to receive regular assistance throughout the study period;⁷⁰
 - the “discontinued-short notice (short notice)” assistance households who stopped receiving assistance in January 2024; and
 - the “discontinued-long notice (long notice)” households who stopped receiving assistance in July 2024.⁷¹

⁶⁹ Randomization was stratified by each combination of eligibility space and targeting method – 30 households were randomly selected from within each one of the 32 cells indicated in red in Table A4.1 in Annex 4.

⁷⁰ Assistance for this group continued until at least June 2025, when the LBCO conducted a new prioritization phase.

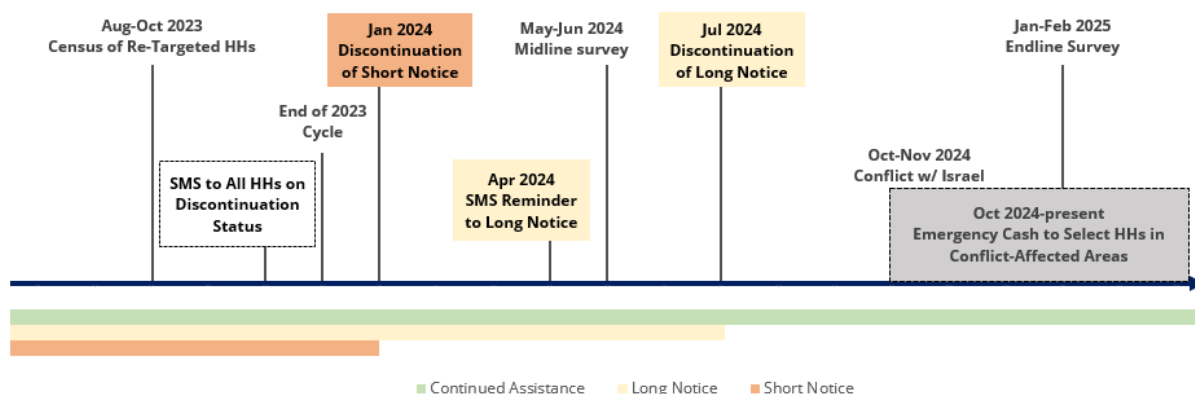
⁷¹ See Appendix A5 for the study limitations and implications for the study's analysis and findings.

5. Data and outcomes

5.1. Data sources

70. The evaluation combines the following main sources of data:
71. A **baseline census** of the 58,501 households that were part of the retargeting process: This was conducted by the LBCO between August and September 2023. It serves as the basis for the implementation of the study randomization as well as the source of baseline information for the main analysis.
72. **WFP administrative data** on programme implementation and appeals: The evaluation uses (1) monthly records of households that collect the food parcels to assess adherence to the study design, and (2) information on appeals submitted through phone calls or an online form.
73. **Two rounds of household survey data**⁷² across a representative study sample of 6,400 households: The first round of data collection took place between May to June 2024, just before the end of assistance for the long notice households. This allowed the evaluation to investigate the impact of discontinuation and the differences in outcomes and behaviours between short notice and long notice households. The second round of data collection took place between January and February 2025, which provided the evaluation the opportunity to study impacts in the longer run. Figure 4 shows how data collection was timed in relation to programmed implementation.

Figure 4: Timeline of programme implementation and data collection



74. The **LHS**⁷³ for PMT modelling: The LHS is structured as a nationally representative survey of approximately 3,500 households. The dataset records detailed information on poverty levels, food security, employment status, access to essential services, and coping mechanisms among households. The LHS was collected between January to May 2023, just a few months before the

⁷² Annex A4 provides further details on survey sampling, attrition and balance.

⁷³ The LHS is a collaboration between the World Bank, United Nations Refugee Agency (UNHCR), and WFP – with support from the Joint Data Center on Forced Displacement.

census. The evaluation makes use of it to train the three PMT targeting models by considering the set of 126 variables that overlapped with the census dataset.⁷⁴

5.2. Outcomes

75. Table 1 outlines the key outcomes included in the study as well as their related data sources.⁷⁵ More details on these outcomes are provided in Appendix A6 with Table A6.1 describing the indicators and Table A6.2 explaining the construction of outcomes based on composite indices.
76. **Cost-effectiveness:** While WFP impact evaluations typically include cost-effectiveness as a key outcome of interest, the nature of the evaluation design resulted in costs that are equal across all four data-driven targeting methods. Thus, cost-effectiveness was not analysed as part of the evaluation.

Table 1: Outcomes of interest

Outcome group	Outcome	Admin data	Survey round 1	Survey round 2
Satisfaction with targeting	Appeal rate	X		
Food security	Food Consumption Score (FCS)		X	X
	Household expenditure for essential food and non-food items		X	X
	Reduced Coping Strategies Index (rCSI)		X	X
	Percentage of households using livelihood coping strategies		X	X
Psychological well-being	Life Satisfaction Score		X	X
	Worriedness Score		X	X
	World Health Organization-Five Well-Being Index (WHO-5 Index)		X	X
	Post-Traumatic Stress Disorder (PTSD) Score		X	X
Community cohesion	Social Cohesion Index		X	X
	Trust Index		X	X
Household finances	Labour Income		X	X
	Percentage of households with debt		X	X
	Percentage of households receiving remittances		X	X
	Perceived financial security			X
Employment	Overall employment rate		X	X
	Male employment rate		X	X
	Female employment rate		X	X
Women's agency	Household decision-making related to food		X	X
	Household decision-making related to major purchases		X	X
	Household decision-making related to employment			X
Human capital development	Enrolment rate (compulsory schooling)		X	X
	Health-seeking behaviour		X	X

⁷⁴ To increase similarity with the target population, non-Lebanese households and the richest 25 percent (based on per capita expenditure) were excluded from the Lebanon Vulnerability and Poverty Assessment sample before using it with the PMT models.

⁷⁵ While WFP impact evaluations typically include cost-effectiveness as a key dimension of interest, the nature of this evaluation implies very similar costs across the different targeting methods, making such analysis uninformative.

6. Main findings

77. This chapter presents the study results in three parts, organized by evaluation question. Sections 6.1 to 6.4 address the first component of the study, which compares different prioritization methods: 6.1 discusses the profiles of households selected by each method; 6.2 compares the targeting methods by appeal rate; 6.3 looks at average impacts; and 6.4 analyses whether there are differential impacts due to the targeting approach applied. Sections 6.5 and 6.6 then analyse the consequences of ending the assistance, with 6.5 detailing average impacts of discontinuation and 6.6 investigating whether certain types of households are more affected. Section 6.7 presents findings on whether and how households prepare for programme exit.

Part 1: Prioritizing among households already being supported when scaling down

78. Section 6.1 uses the *full* population of households receiving in-kind assistance to study prioritization. Section 6.2 examines acceptance using administrative appeal data covering the entire population of *excluded* households. All remaining outcomes and analyses in sections 6.2 and 6.3 rely on data from the two survey rounds, which cover the stratified random sample described in Section 5. Sampling weights are applied to correct for the overrepresentation of smaller eligibility spaces, ensuring that estimates are representative of the full population.

6.1 What are the differences across the four methods in terms of who gets selected to continue receiving assistance?

Summary of findings: The four targeting approaches lead to substantial differences in the profiles of the households selected for continued assistance. The three PMT-based methods tend to select households with more comparable profiles relative to the categorical approach, though there are notable differences in their selection as well. Statistical tests confirm that the differences in the profiles selected by the four methods are systematically different.

79. The analysis begins by comparing the selections made by the four different targeting methods. Although the four methods overlap to some extent in the way they prioritize household characteristics, they do not select the same set of households. The three PMT-based methods have more overlaps, but the correlation in their selection – estimated by simulating the three methods over the entire population – is only 0.55 on average (see Figure A7.1 in Annex A7).

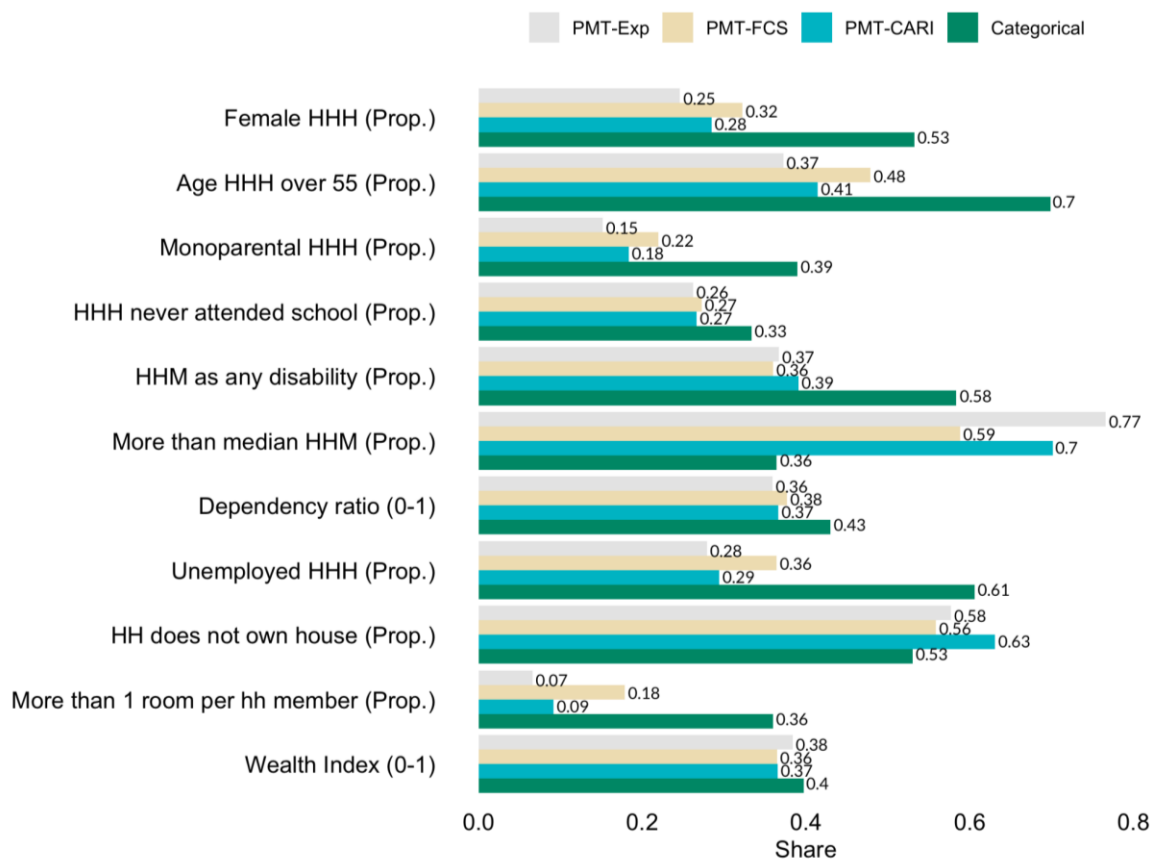
80. Figure 5 shows a selection of key differences in characteristic for the households selected by each method.⁷⁶ Consistent with the finding on the correlation between methods, the Categorical approach stands out in selecting a notably different profile of households. For

⁷⁶ Table A7.1 in Annex A7 provides a richer comparison, which includes a formal test for equality of the means, across 30 key baseline socioeconomic and demographic characteristics.

example, 53 percent of households selected to continue receiving assistance under the Categorical approach were headed by women,⁷⁷ compared to only 25–32 percent under the PMT-based methods. Similarly, 61 percent of household heads selected by the Categorical approach were unemployed, versus 28–36 percent under the PMT-based approaches.

Although it is difficult to neatly label the profiles selected by the different methods, the figure illustrates that the Categorical approach is more likely to select households with “common” markers for vulnerability (such as gender of the household head, disability, or unemployment). Possible implications of this finding are further discussed in section 6.3 on the heterogeneous impacts of the different targeting methods.

Figure 5: Predictors of selection of targeting methods (using full census)



Note: Own visualization. The figure shows the proportion of households interviewed at baseline that fulfil the characteristic reported among households selected to continue receiving food assistance, disaggregated by targeting group.

81. Nevertheless, meaningful variation also exists among the three PMT methods. When comparing the selection along the full set of 142 baseline characteristics recorded in the census, even when excluding the Categorical approach and restricting the analysis to only the three PMT approaches, 75 percent of the variables show statistically significant differences at the 5 percent level, confirming that substantial variation exists between the PMT approaches.

⁷⁷ Differential impacts of targeting on key subgroups are discussed in section 6.4.

6.2 Which targeting approach leads to higher acceptance (as proxied by the lowest number of formal appeals)?

Summary of findings: Appeal data reveal that excluded households are more likely to contest their exclusion under the Categorical approach. By relying on coarse sociodemographic criteria – such as having a member with disability, or household head being elderly – this method may be less effective than data-based approaches in capturing the complexities of vulnerability, thus leading to lower acceptance. The results also show that households that were eligible under more methods had significantly higher appeal rates: households eligible under three of the four methods but that were excluded appealed 62 percent of the time versus 50 percent for those that were never eligible ($p < 0.01$). To the extent that more vulnerable households are more likely to contest their exclusion, these findings indicate that combining methods can lead to improved accuracy.

82. Households that were excluded from assistance had the right to appeal through WFP's complaint and feedback mechanism (through a telephone hotline or an online platform). This option was communicated to households in the same message that notified them of the termination of assistance. The message was standardized across all excluded households and did not specify the particular targeting method that resulted in their discontinuation.
83. This analysis makes use of administrative data on appeals collected by the LBCO for the entire population of excluded households.⁷⁸ By using appeals as a proxy for acceptance of the targeting decision, it is possible to assess the level of acceptance of the selection resulting from the different targeting methods.
84. Figure 6 shows the share of discontinued households that submitted an appeal, disaggregated by targeting method. In the figure, the PMT-Exp group is taken as reference. Among the 32,361 households whose assistance ended, an average of 52 percent filed a formal appeal.^{79,80} In the reference group, exactly 50 percent of discontinued households contested their exclusion from the programme. By comparison, appeal rates were significantly higher under PMT-FCS targeting at 52 percent (p -value < 0.05) and Categorical targeting at 56 percent (p -value < 0.01).
85. The higher rate of appeals from households excluded under the Categorical approach is notable. Although this method prioritizes groups that are typically considered to be structurally vulnerable, it may be less suitable to capture the complexity of household-level needs when compared to other data-driven methods.
86. This is confirmed when considering the proportion of poor households (as measured by the asset index, which is the only proxy available in the census) excluded under the

⁷⁸ Results remain qualitatively similar when restricting the focus to the study sample.

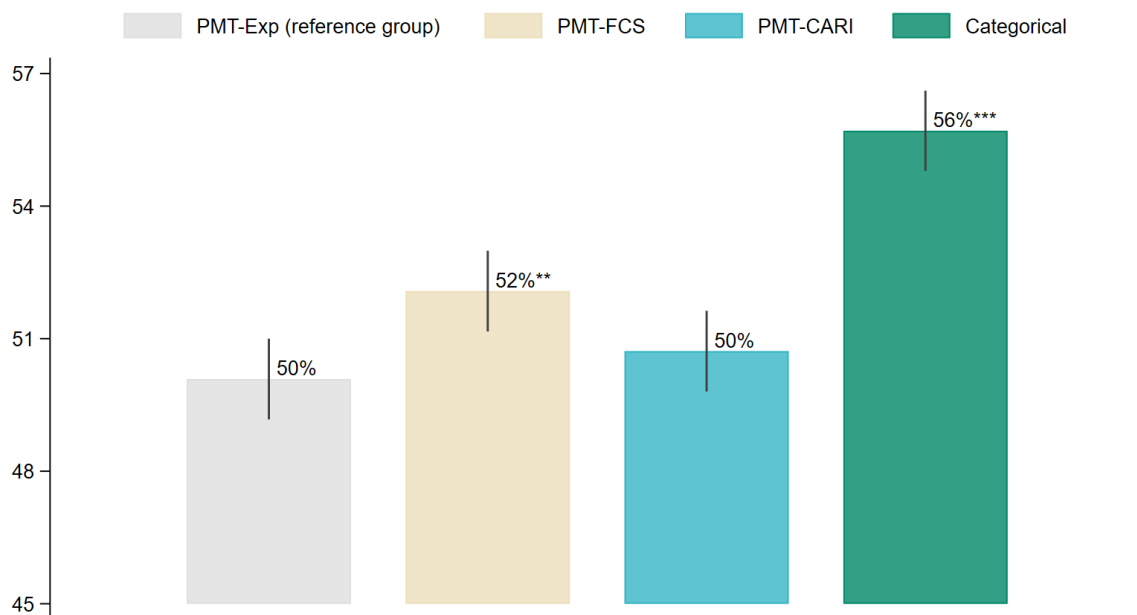
⁷⁹ The 980 households assigned to the long-notice window are excluded from this computation.

⁸⁰ By the first round of data collection in June 2024, less than 1 percent of appeals had resulted in a change of status. After July 2024, another 2 percent of appeals were deemed successful and their assistance was resumed.

Categorical versus PMT-based methods: 38 percent of households excluded by the Categorical approach are below median wealth,⁸¹ compared to 30 percent for the PMT methods.⁸²

87. This finding is consistent with existing evidence that richer data-driven approaches, such as PMT, tend to be more precise in identifying household-level vulnerability than categorical approaches.⁸³ By relying on coarse demographic or socioeconomic categories, there is a higher risk of incurring inclusion and exclusion errors which can undermine the satisfaction and acceptance of targeting decisions.

Figure 6: Appeal rate of discontinued households by targeting arm



Note: The figure shows the percentage of households excluded from assistance that submitted an appeal, disaggregated by targeting method. The reported estimates represent weighted means and the PMT-Expenditure group serves as the reference category. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Estimates use robust standard errors.

Combining methods

88. While the primary objective of this section is to compare the effectiveness of the four distinct targeting methods, the design also enables a complementary analysis of *multiple eligibility*: that is, whether being identified as eligible by multiple methods correlates with higher chances of appealing an exclusion decision.
89. Since each targeting approach focuses on a different definition of vulnerability, combining the selections across methods may increase the likelihood of identifying the most

⁸¹ Poverty is proxied by a binary indicator which is coded as 1 if a household is below the median of the wealth index, which itself is a combination of key assets owned by a household.

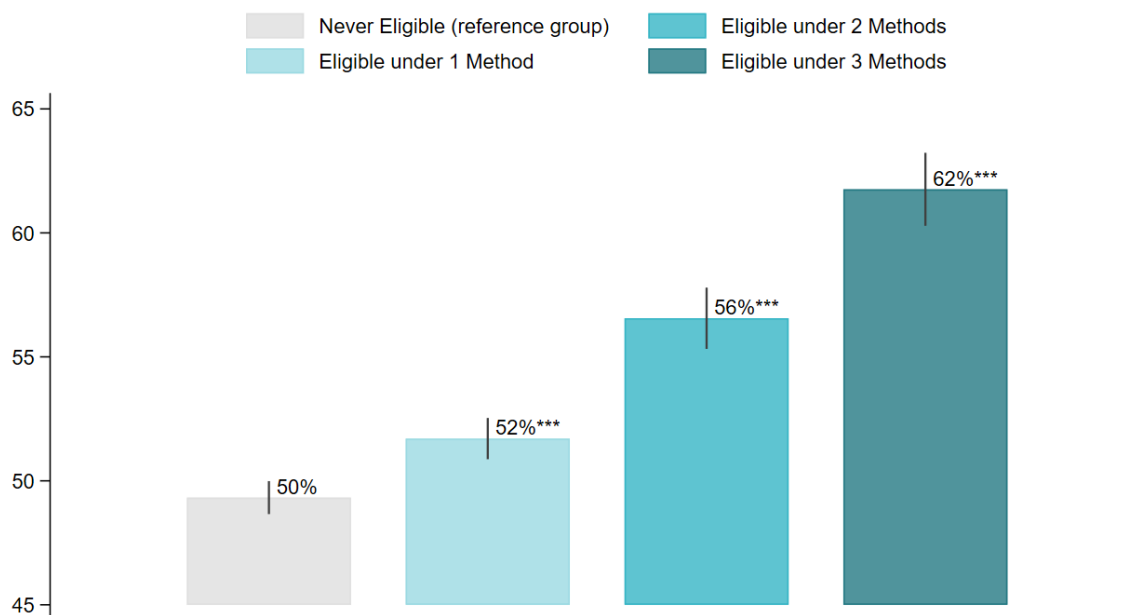
⁸² The analysis also explored whether the much higher rate of appeals under the Categorical method could be attributable to gender-based differences in appealing, but this did not appear to be a significant factor.

⁸³ 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>

vulnerable households. Intuitively, households deemed eligible by several methods are more likely to be truly vulnerable than those identified by only one or none. If appeal behaviour reflects underlying need, one would expect higher appeal rates among excluded households that were eligible under three of the four methods, and lower rates among those deemed eligible under none of the methods.

90. Figure 7 supports this hypothesis by showing how the probability of appealing a discontinuation decision increases with the number of methods under which a household discontinued household would have qualified. In the reference group – households not identified as eligible by any method (i.e. eligibility space E0 in Figure 3) – exactly 50 percent submitted an appeal. This rate rises to 62 percent among households deemed eligible under three methods.⁸⁴

Figure 7: Appeal rate of discontinued households by eligibility



*Note: The figure shows the percentage of households excluded from assistance who submitted an appeal based on the full census sample, disaggregated by the number of targeting methods they are eligible under. Estimates reflect simple means, and the Never Eligible group serves as the reference category. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.*

91. The figure reveals a clear, positive gradient: the more targeting methods that identified a household as eligible, the more likely the household was to contest its exclusion from assistance. This pattern indicates that combining targeting approaches can be a promising path to improving the accuracy and perceived fairness of eligibility decisions. Figure A7.2 in Appendix A7 shows that, consistent with this interpretation, households deemed eligible under multiple targeting methods tend to exhibit lower socioeconomic well-being, including lower levels of wealth and educational attainment.

⁸⁴ When restricting the sample to only households interviewed during round 2, the same analysis suggests similar results with 47 percent complaining under no eligibility, 59 percent complaining when eligible under one method, 61 percent complaining when eligible under two methods, and 61 percent complaining when eligible under three methods.

6.3 Which targeting method leads to overall higher well-being among the target population?

Summary of findings: Analysis of key well-being outcomes – including food security (FCS, food expenditure, coping strategies), psychological well-being (life satisfaction, mental well-being, worriedness, and PTSD), household and human capital (finances, employment, agency, education, and health), and community indicators (social cohesion and trust) – shows that, although the different targeting methods select distinct profiles, no one targeting method consistently outperforms the others in improving average outcomes among the target population (when averaging over continued and discontinued households). This aligns with previous evidence indicating that greater targeting precision does not necessarily translate into measurable gains in aggregate well-being.

92. The analysis for this section is based on survey data collected from the study sample.⁸⁵ The analysis examines whether the differences in selection and acceptance between targeting methods described above translate into differences in programme effectiveness or improvements in the well-being of the overall target population. Outcomes of households that continued to receive assistance and discontinued households are combined here to provide a holistic view of how the application of one targeting method rather than another affects the overall well-being of the target population. The following discussion of the results is organized by outcome. The next section provides a more nuanced analysis on the differences between continued and discontinued households.

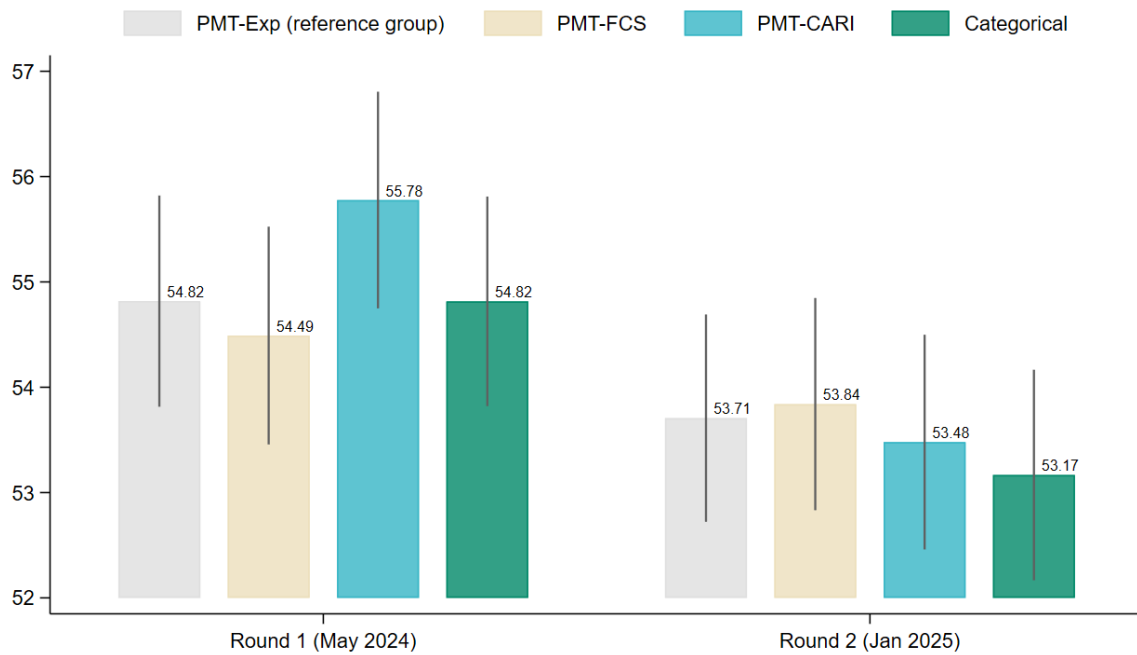
Food security

Food Consumption Score (FCS)

93. Figure 8 compares FCS across the four targeting methods, disaggregating the results by short (about 6 months) and medium (about 12 months) term, and keeping PMT-Exp as the reference group. This analysis combines continued assistance households and discontinued households.
94. No statistically significant differences across the four targeting methods are detected in either the short or the medium term. Although households under PMT-CARI appear to have slightly augmented FCS in round 1, this difference dissipates by round 2, where the average FCS was nearly identical across all four targeting arms. Overall, these results show that no targeting method outperforms the others in terms of average food consumption in the target population.
95. A surprising finding from this analysis is that PMT-FCS – despite being explicitly designed to support households with low FCS – did not perform better than the other methods in improving the FCS for the target population. Figure A7.3 in Appendix A7 shows that the average FCS and its distribution across the population is very similar across groups.

⁸⁵ As discussed, the study sample is a stratified random sample of the overall population.

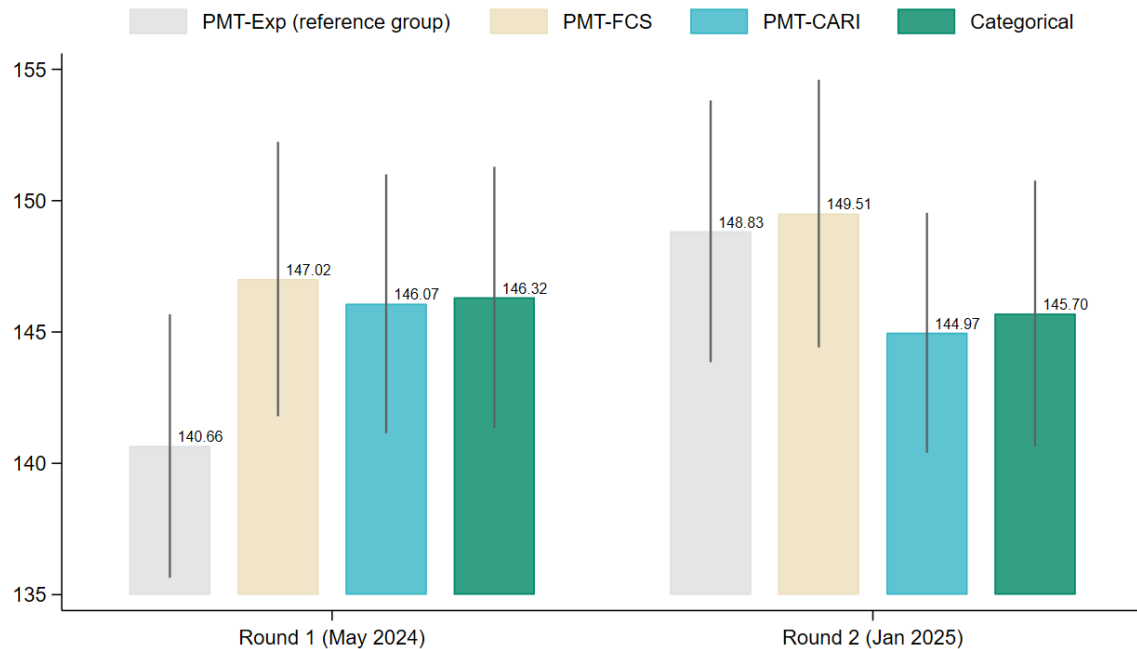
Figure 8: FCS across targeting arms



Note: Food Consumption Score (FCS) is a continuous variable ranging from 0 to 112, with higher values indicating better food security. The reported Estimates represent weighted means and the PMT-Expenditure group serves as the reference category. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

Food expenditure

Figure 9: Food expenditure (USD) across targeting arms



Note: Food expenditure includes food purchased with cash/credit and the value of food from own production, but it excludes in-kind food assistance received by the household. Expenditure was winsorized at the 95th percentile to reduce the influence of outliers. The reported estimates represent weighted means and the PMT-Expenditure group serves as the reference category. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

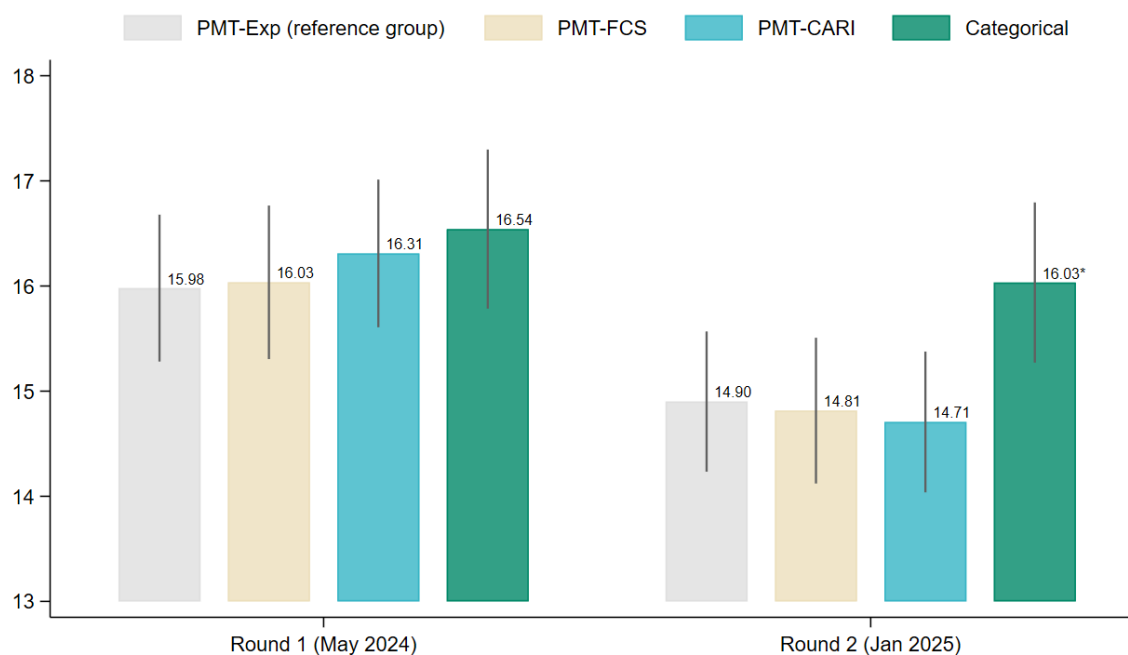
96. Figure 9 compares the four arms in terms of household food expenditure (in USD). The outcome includes the value of any food purchased with cash/credit as well as the value of food from own production, but it excludes in-kind food assistance. No statistically significant differences were detected between the four targeting methods. Although food expenditure under PMT-Exp (the reference group) appears somewhat lower than the others in round 1, this difference disappears by round 2.

Coping Strategies (rCSI and LCSi)

97. Figure 10 compares the rCSI across the four study arms. The index, which ranges from 0 to 56, captures the reliance of households on harmful food-related coping strategies, such as reducing the number of meals or decreasing portion sizes. A higher index indicates greater reliance on negative coping behaviours.

98. The only significant difference across the four methods emerges in round 2, where the Categorical arm recorded somewhat higher (worse) rCSI compared to the reference group (16.03 versus 14.90, $p < 0.10$). However, the results overall show no clear or consistent differences in average rCSI across targeting arms. Figure A7.5 in Appendix A7 further confirms that the means and the distributions of rCSI are very similar across the four groups.

Figure 10: rCSI across targeting arms



*Note: The reduced Coping Strategies Index (rCSI) ranges from 0 to 56 and captures the frequency and severity of harmful behaviours that households might adopt to cope with food shortages. The reported estimates represent weighted means and the PMT-Expenditure group serves as the reference category. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.*

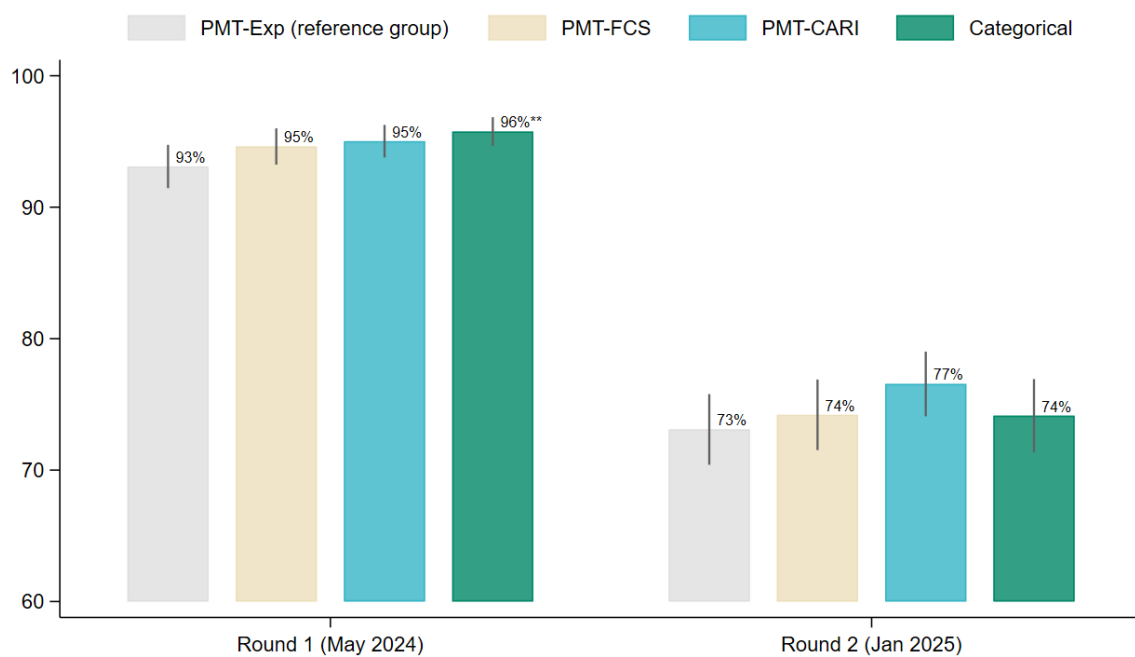
99. Figure 11 examines adoption of stress-related coping strategies in the six months before data collection, as measured by the Livelihood Coping Strategies Index (LCSI). These strategies are the most commonly reported in the sample. They are generally interpreted

as early indicators of household difficulties in accessing food or income, occurring before longer-term productive capacity is compromised.⁸⁶

100. The strategies examined include reducing or ceasing payments for essential utilities and bills, selling household assets, cutting education expenditures, and limiting non-essential health spending. Approximately 95 percent of households report having adopted at least one of these strategies in round 1, and slightly less than 75 percent in round 2. The figure shows that in round 1, households selected under the Categorical approach are the most likely to resort to stress coping strategies (96 percent, $p < 0.05$). However, such difference disappears by round 2.

101. Taken together, the analysis on food security shows that the different targeting approaches do not lead to substantially different outcomes across the target population in either the short or medium run.

Figure 11: Usage of livelihood stress coping strategies in the past six months across targeting arms



Note: The LCS-stress indicator captures whether households resorted to using stress-related coping strategies, such as borrowing food or money or spending savings. The outcome is expressed as a binary variable that takes a value of 1 if the household adopted at least one strategy. The values reported in the graph represent the share of households adopting at least one strategy approximately six months before data collection in each study arm. Estimates reflect simple means, and the PMT-Expenditure group serves as the reference category. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

Psychological well-being

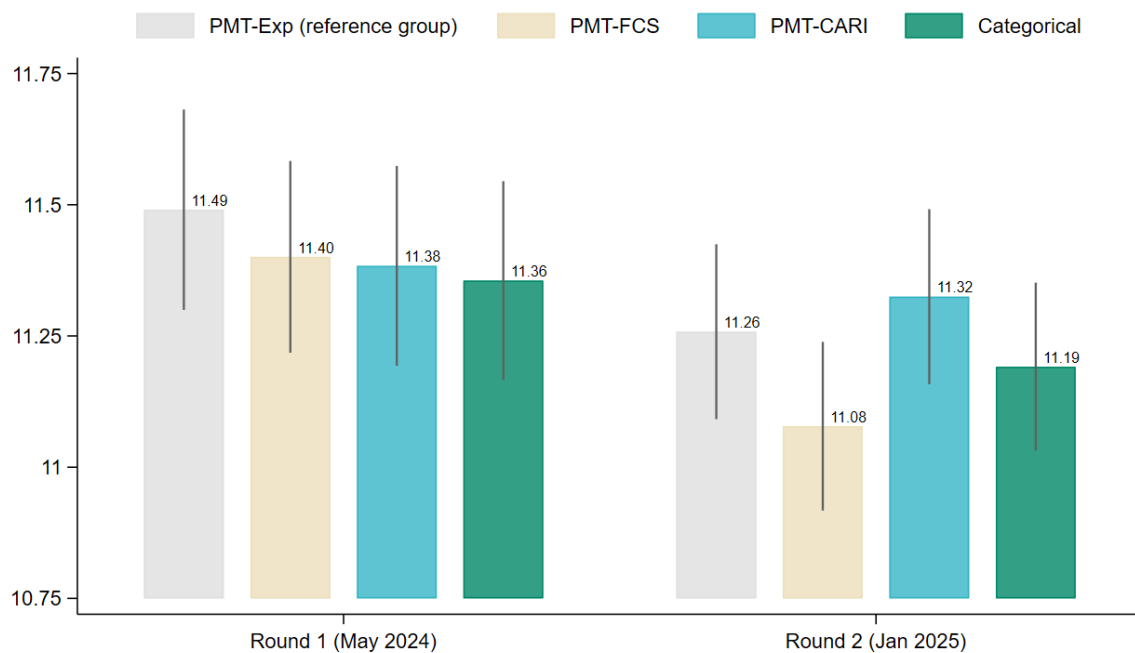
102. No statistically significant differences across the four targeting methods are observed in any of the psychological well-being measures in either survey round. Figure 12 presents

⁸⁶ Vaitla, B., Coates, J., & Maxwell, D. 2015. *Comparing Household Food Consumption Indicators to Inform Acute Food Insecurity Phase Classification*. FHI 360/Food and Nutrition Technical Assistance III Project (FANTA).

details of the comparison of life satisfaction scores, which range from 0 to 20 and are defined as the sum of responses to Likert-scale items assessing satisfaction with household nutrition, health, finances, community relations, and overall life. Figure 13 presents results for the WHO-5 index, which shows averages ranging from about 20.89 to 24.52, well below the standard threshold of 50 for poor mental well-being. Figure 14 reports results for the “worriedness index”, which was built following Hidrobo et al. (2023).⁸⁷ Overall, results show little variation across targeting arms and rounds, with no statistically significant differences observed.

103. Figure 15 compares Post-Traumatic Stress Disorder (PTSD) screening scores across the four targeting methods. It shows no significant differences. Unlike the other psychological well-being outcomes, results are only available for round 2 of the survey for this outcome as the PTSD module was only added for round 2 after the escalation of the conflict with Israel in October 2024. While PTSD scores and their prevalence do not appear to differ between targeting methods, it is notable that almost half of the households surveyed (45 percent) met the criteria for having PTSD. This is much higher than the 16 percent prevalence rate found in other studies of conflict-affected populations that used the International Trauma Questionnaire.⁸⁸

Figure 12: Life satisfaction across targeting arms

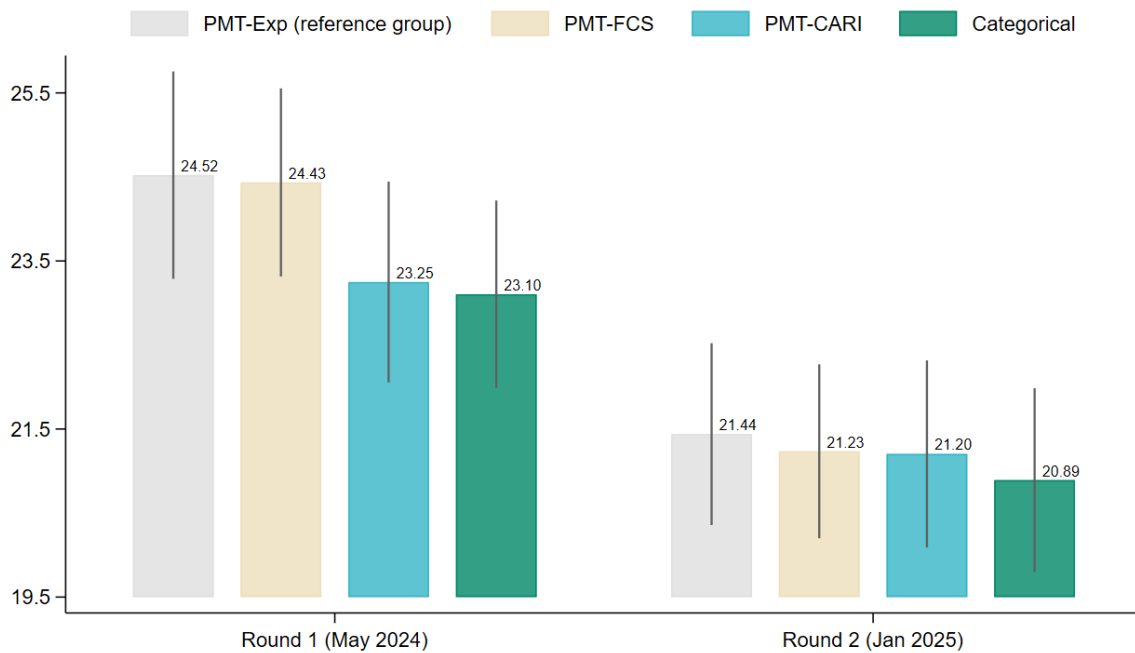


*Note: The life satisfaction index is presented here as a composite score based on Likert-scale questions capturing satisfaction with household nutrition, health, finances, community relations, and overall life. A higher score indicates better life satisfaction. Estimates represent weighted means and the PMT-Expenditure group serves as the reference category. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.*

⁸⁷ Hidrobo, M., Karachiwalla, N., & Roy, S. 2023. The Impacts of Cash Transfers on Mental Health and Investments: Experimental evidence from Mali. <https://doi.org/10.1016/j.jebo.2023.10.016>

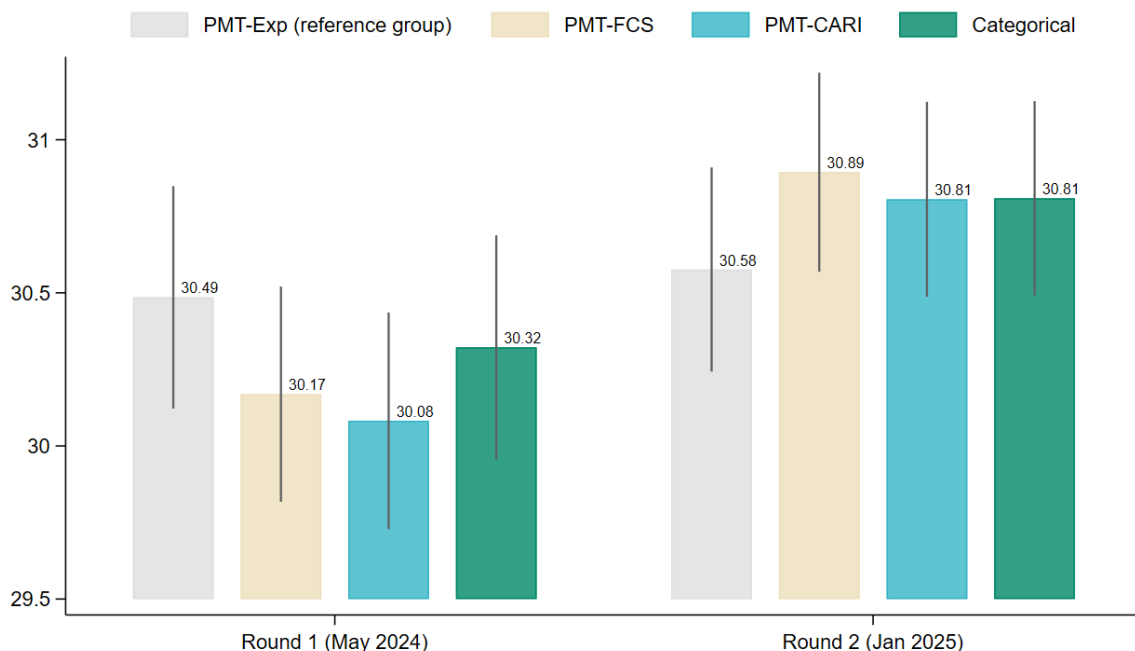
⁸⁸ Fung, Hong Wang et al. Prevalence of ICD-11 Post-Traumatic Stress Disorder (PTSD) and Complex PTSD in the General Populations: A Systematic Review and Meta-Analysis. *Asian Journal of Psychiatry* 110 (August 2025): 104610. <https://doi.org/10.1016/j.ajp.2025.104610>

Figure 13: Mental well-being (WHO-5) across targeting arms



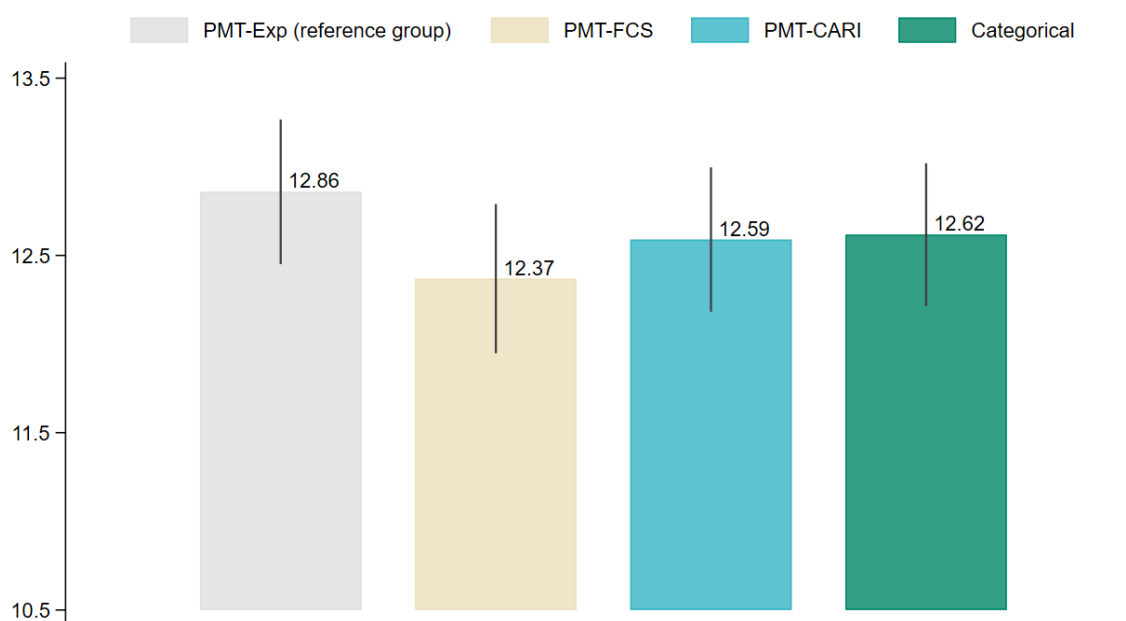
Note: The adapted WHO-5 mental well-being index captures mental well-being over the past week (as opposed to the standard two weeks) through five items on cheerfulness, calmness, rest, energy, and interest in life. Each is rated from 0 ("At no time") to 5 ("All of the time"), yielding a total from 0 to 25, which is multiplied by 4 to obtain a score out of 100. Estimates represent weighted means, and the PMT-Expenditure group serves as the reference category. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

Figure 14: Worriedness across targeting arms



Note: The worriedness index is based on Hidrobo et al. (2023) and constructed from Likert-scale questions capturing household concerns across ten domains: food, health care, education, debt, employment, family, extremism, conflict, natural disasters, and displacement. Responses range from 1 (Not worried at all) to 4 (Very worried). Estimates represent weighted means, and the PMT-Expenditure group serves as the reference category. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

Figure 15: PTSD score across targeting arms



Note: The PTSD module was adapted from the International Trauma Questionnaire and asks about three categories of trauma-related behaviour: re-experiencing in the here and now; avoidance; and a sense of current threat. Each item is rated on a 5-point Likert scale from 0 (“Not at all”) to 5 (“Extremely”). The responses are summed to produce a dimensional PTSD score which ranges from 0 to 24. Estimates represent weighted means, and the PMT-Expenditure group serves as the reference category. Vertical lines indicate 90% confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

104. Overall, the analysis on psychological well-being confirms what emerged by looking at food security: despite selecting different households, on average, the different targeting approaches do not lead to substantially different outcomes across the entire target population in either the short or medium run.

Community, socioeconomic, and human capital development outcomes

105. The lack of substantial differences across targeting methods is confirmed when considering socioeconomic, human capital, and community outcomes (see Annex A7). No differences are observed in social cohesion (Figure A7.6) and trust (Figure A7.7). For household finances (Table A7.2) and employment (Table A7.3), differences detected are small and inconsistent.

106. For human capital development and agency indicators (Table A7.4), there were notable differences on the indicators related to women’s decision-making. During round 2, respondents under PMT-CARI and Categorical targeting both reported higher levels of women’s agency, which were significant at the 5 percent level for food-related decisions and the 1 percent level for major purchases. Given that these results appear by round 2, it might be related to the emergency cash injection for discontinued households, which is not the subject of this evaluation.

107. The results show that no targeting method consistently outperforms the others. These findings are in line with the limited evidence from the targeting literature, which generally show that, while different approaches might lead to the selection of somewhat different

groups, such variation does not translate into differences in programme effectiveness as measured by the average well-being across the entire target population. Still, it is possible that the observed average masks substantial heterogeneity across subgroups, which is the focus of the next section.

6.4 Which types of households benefit the most from each targeting approach?

Summary of findings: When comparing targeting methods across sociodemographic groups/criteria, the study finds no systematic differences in outcomes. The evaluation shows that the three PMT methods select vulnerable households that are worse-off when compared to non-selected (discontinued). When the Categorical approach is used, continued and discontinued households are on similar levels (for rCSI and worriedness), and results in fewer highly vulnerable households being selected – indicating poorer targeting performance. This result helps explain why households excluded under the Categorical method were found to be more likely to appeal the exclusion decision.

108. Although the previous section showed no differences in *average welfare* across targeting methods, it is still possible that the methods lead to significant differences for specific subgroups of the population. To test for this possibility, this section disaggregates the previous analysis, splitting the sample according to specific characteristics. The selected characteristics can be categorized into two groups:

- common sociodemographic markers of vulnerability (e.g., male vs. female head of household); and
- continuation/discontinuation status.

By sociodemographic markers of vulnerability

109. The study identified six sociodemographic criteria that were of interest to the LBCO to include in this analysis: household head was female; household was a minor/elderly; household head had a low education level; household had a member with a disability; household had below-median wealth; and household had a high dependency ratio. In the interest of space and length, the analysis here focuses on three key outcomes – FCS, food expenditure, and life satisfaction.

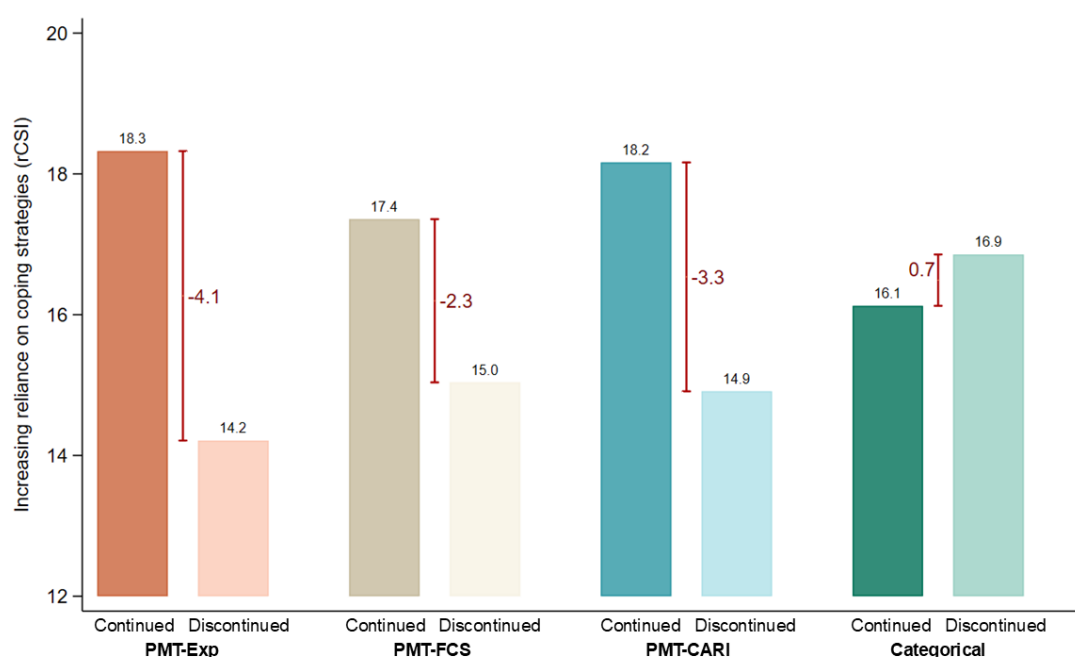
110. Mirroring the results from the previous section, this analysis also shows that no targeting method consistently advantaged a specific subgroup of the population across these outcomes and over time. While households headed by women are generally worse-off across the board, households headed by women that were assigned to the Categorical method show larger FCS scores by round 1 (+3.1, $p < 0.10$), when compared against those households assigned to PMT-Exp. This is likely due to the fact that the Categorical method selects households headed by women at a higher rate when compared to the PMT methods.

111. Full regression results, disaggregated by survey round, are reported in Tables A7.5 to A7.10 in Appendix A7.

By discontinuation status

112. The analysis shows the differences in outcomes between the 43 percent continued and 57 percent discontinued households for each targeting method at data collection round 1. It is important to note that this analysis does not provide an estimate of the causal impact of discontinuation itself, which is addressed in section 6.5.
113. Figure 16 shows the gaps in rCSI by targeting method during round 1 of data collection. Households selected for continued assistance by all three PMT methods were worse off at the time of selection when compared to those selected for discontinuation – indicated by higher rCSI scores, or increased reliance on negative coping strategies.

Figure 16: Descriptive differences in rCSI between discontinued and continued households by targeting method during survey round 1 (May 2024)



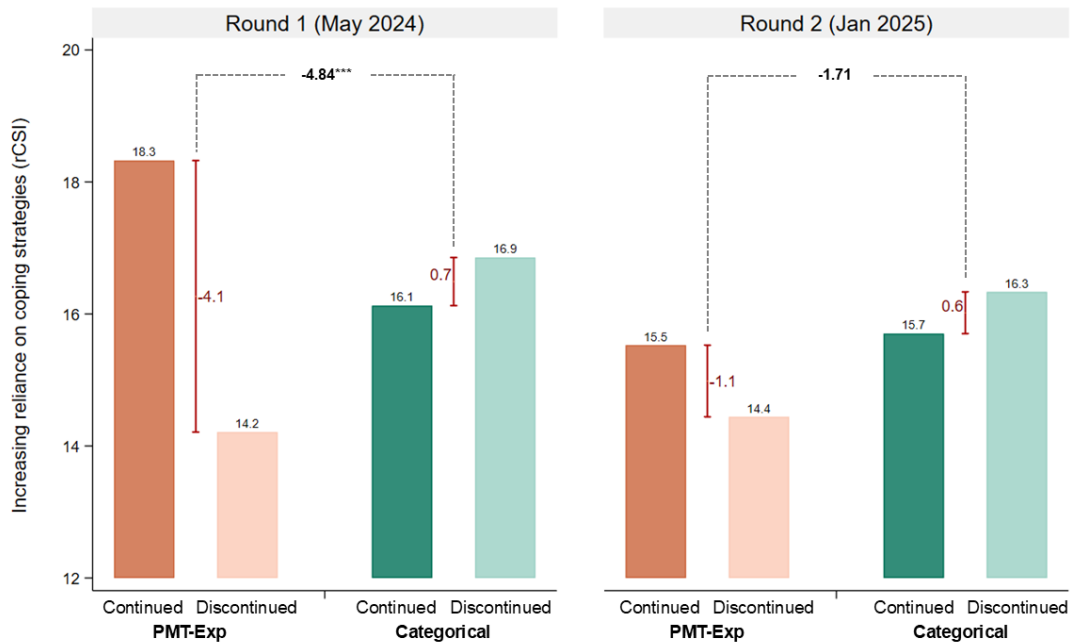
Note: The figure presents weighted means of the reduced Coping Strategies Index. The red bars indicate the gap/difference between continued (“Cont”) and discontinued (“Disc”) households. However, these gaps are descriptive and do not provide the true estimate of the impact of discontinuation.

114. A possible explanation for this result is that the PMT methods are better at selecting more vulnerable households – people who still rely on coping strategies even when provided with food assistance. By contrast, the Categorical method may lack precision by prioritizing households that are relatively better-off – people who have less need for coping strategies when given support.⁸⁹ This is also supported by the findings from section 6.2. The findings there showed that the Categorical method excluded poorer households at a much higher rate than the PMT-based methods.
115. To further validate this observation, the analysis compares the Categorical method with PMT-Exp in depth: Figure 17 shows the unbiased difference in gaps between the two

⁸⁹ A counterargument could also be made that the Categorical method is able to select households that are more responsive to assistance as opposed to the households selected by the PMT methods.

targeting methods, with a large, significant estimate for round 1 (-4.84, $p < 0.01$), though this difference shrinks in round 2. The result for round 1 indicates that when Categorical targeting is implemented instead of PMT-Exp, there is a significant difference in who benefits between discontinued and continued households.

Figure 17: Comparison of the differential impacts of targeting on rCSI between PMT-Exp and Categorical



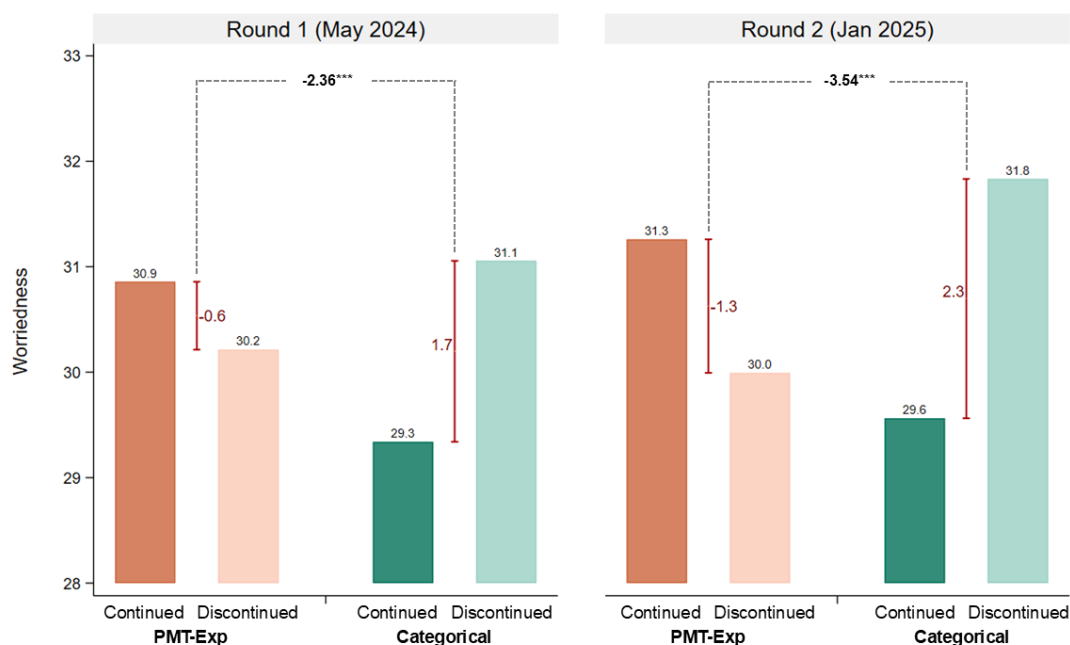
Note: The figure presents weighted means of the reduced Coping Strategies Index. The red bars indicate differences between continued and discontinued households in the same targeting arm, while the dotted bars provide the combined effect, which is a sum of the differences, or “difference in differences”. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

116. A similar pattern is observed when considering the worriedness outcome (see Figure 18). In this case worriedness may be seen as a proxy for vulnerability: households in need of assistance would likely express higher worriedness. Under PMT-Exp, despite no longer receiving assistance, discontinued households were less worried than their continued counterparts. However, the opposite is true for Categorical households as the discontinued under this method expressed much higher worriedness. The unbiased differences show that the gaps between PMT-Exp and Categorical for both survey rounds were significant (-2.36, $p < 0.01$ for round 1; -3.54, $p < 0.01$ for round 2). This supports the explanation that Categorical is less precise than the PMT methods and excludes more vulnerable households in need of assistance.

117. These results indicate that, although the Categorical method performs just as well as the PMT methods when comparing averages across all households (as discussed in section 6.3), considering discontinuation status provides a different picture. Categorical targeting appears to select fewer vulnerable households and reverses the trend in which group has better outcomes. This might, to some extent, explain why households excluded under the Categorical method were more likely to appeal, as shown in section 6.2.

118. Full results for FCS, rCSI, food expenditure, and worriedness are reported in Table A7.11 in Appendix A7.

Figure 18: Comparison of the differential impacts of targeting on worriedness between PMT-Exp and Categorical



Note: The figure presents weighted means of the worriedness score. The red bars indicate differences between continued and discontinued households in the same targeting arm, while the dotted bars provide the combined effect, which is a sum of the differences, or “difference in differences”. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Part 2: The consequences for households excluded from assistance

119. The analysis now studies the consequences of ending assistance on the well-being of discontinued households (sections 6.4 to 6.5). As explained earlier (in section 4.3), this analysis takes advantage of variation in the discontinuation of assistance for the “marginal” households that would be eligible for assistance under some methods but not others.⁹⁰ The focus of this and the following sections is to estimate the causal effect of the discontinuation of assistance and to shed light on its heterogeneous effects across different subgroups. Continued assistance households and smaller eligibility cells are over-represented in the study sample so that a clean comparison with discontinued households can be made. Population-wide estimates can be retrieved by applying sampling weights.

6.5 What is the impact of ending humanitarian food assistance on households' well-being?

Summary of findings: Discontinuing food assistance led to an immediate decline in household food consumption, particularly for items included in the WFP basket. While these effects faded over time, a reduction in consumption for some food groups persisted 12

⁹⁰ The sampling strategy maximizes the power for the comparison discussed in this section, between continued and discontinued assistance households, but does not directly provide a representative sample across the full population.

months after discontinuation. Households partially compensated by increasing food spending, but not enough to fully replace the lost assistance. Discontinued households increased reliance on negative coping strategies, including borrowing food, reducing meals, and selling assets. Psychological well-being also declined in the short term, with lower life satisfaction lasting through the medium term. Trust in aid agencies fell sharply and remained low throughout the study period. While there were no significant differences in household finances and other welfare outcomes, discontinued households had a significantly lower rate for male employment and a significantly higher one for female employment by round 2. Overall, while households showed some ability to adapt, the end of assistance had clear negative effects on well-being.

120. This section is structured according to the different outcomes of interest and it distinguishes between results in the short- (survey round 1) and medium- (survey round 2) run. One important caveat to keep in mind is that, when round 2 data were collected, most discontinued households had been receiving three months of emergency cash assistance, provided by WFP in response to the conflict with Israel. The comparisons presented in this section should, therefore, be interpreted as providing a lower bound of the true medium-run effect of ending food assistance, as the emergency cash likely diminished the contrast between continued and discontinued households.⁹¹ (See [section 3.1](#) for information on the emergency cash transfer size.)

Food Security

Food Consumption Score (FCS)

121. Figure 19 presents the adjusted mean FCS for continued and discontinued households across the two rounds.⁹² At round 1, discontinued households recorded an average FCS of 54,⁹³ which implies a 3-percent drop compared to the average FCS of 55.5 among continued assistance households ($p < 0.01$). This indicates a decline in food security immediately following the discontinuation of support. However, by round 2, the gap narrowed, and the difference was no longer statistically significant, likely due to the fact the discontinued households received additional emergency cash by that point.

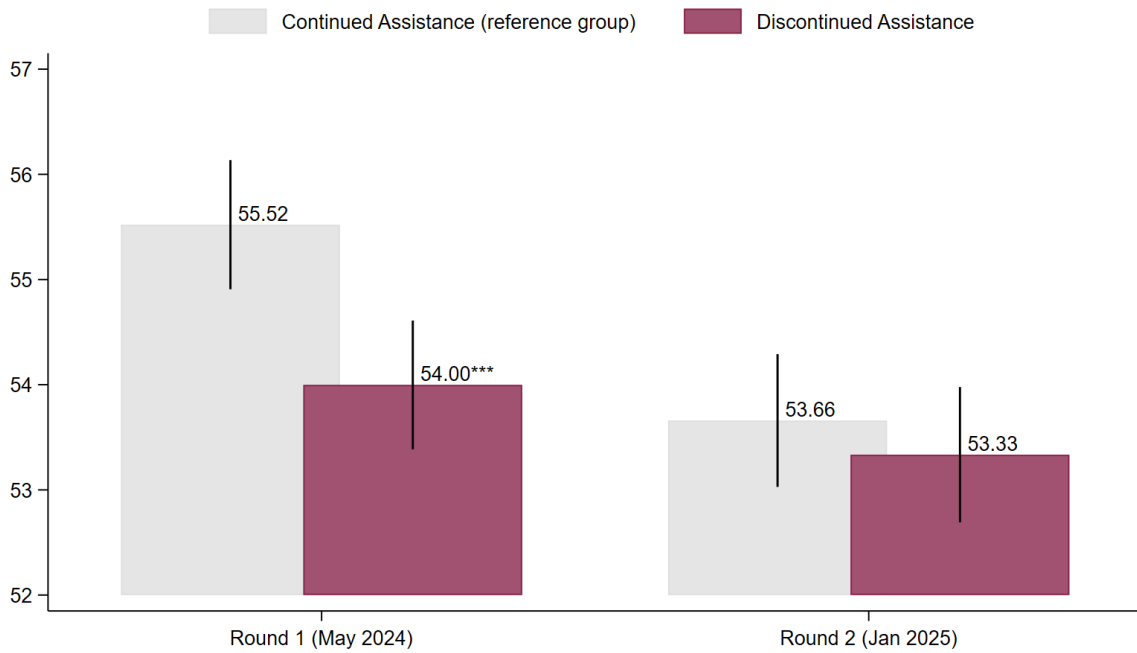
122. This result can reflect resilience, with excluded households finding ways to restore their food consumption levels over time, but it might also partly reflect the benefit of the emergency cash assistance that most discontinued households received by round 2.

⁹¹ It is possible to estimate the impact of discontinuation for households whose status was unaffected by the emergency cash intervention. However, this is outside the scope of this report.

⁹² Adjusted means refer to estimated outcome values obtained from regression models that control for additional variables (see Annex A3). Unlike raw means, which reflect simple averages, adjusted means account for differences in observable characteristics across groups, providing an accurate comparison of the discontinuation effects. For the discontinuation analysis, it is essential to control for the “eligibility” spaces defined in section 4.3.

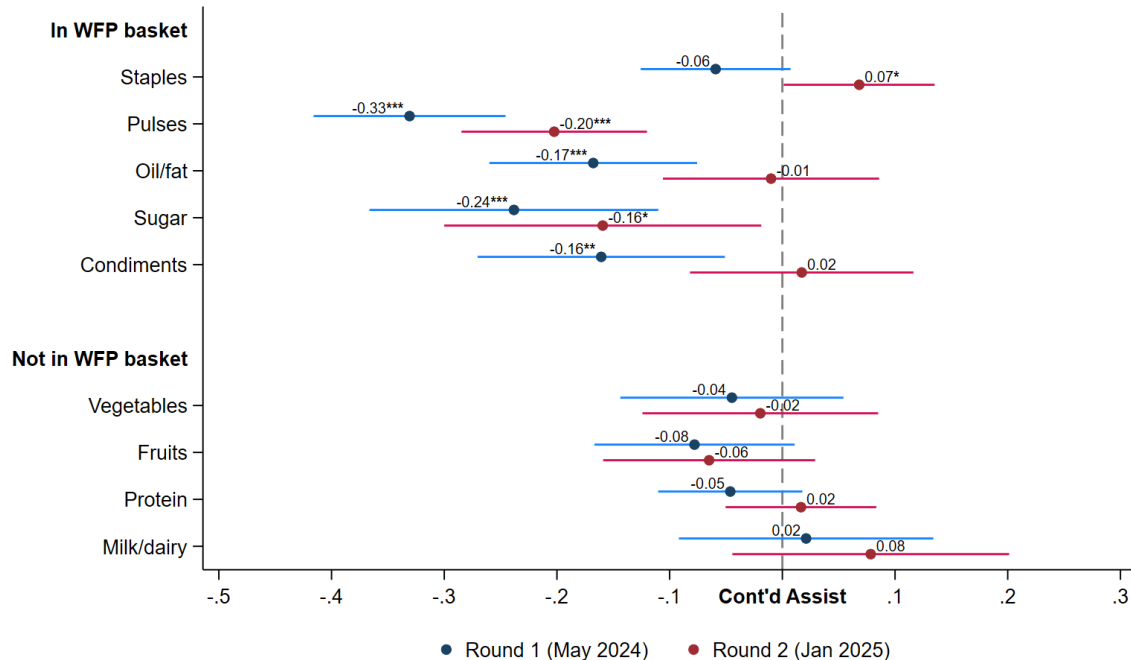
⁹³ The threshold used is the one adjusted for high sugar consumption: food security is classified as poor if households have a score below 28, borderline if they have a score between 28 and 42, and acceptable if they have a score above 42.

Figure 19: Impact of the discontinuation of assistance on FCS



Note: The figure presents the adjusted mean of the FCS, which ranges from 0 to 112. Adjusted means refer to estimated outcome values obtained from regressions that control for additional variables. Control variables were selected using post-double selection Lasso and always include eligibility-space fixed effects and a control for long notice households. For round 2, controls also include the random assignment of survey modules and survey quality. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 20: Impact of the discontinuation of assistance on FCS components



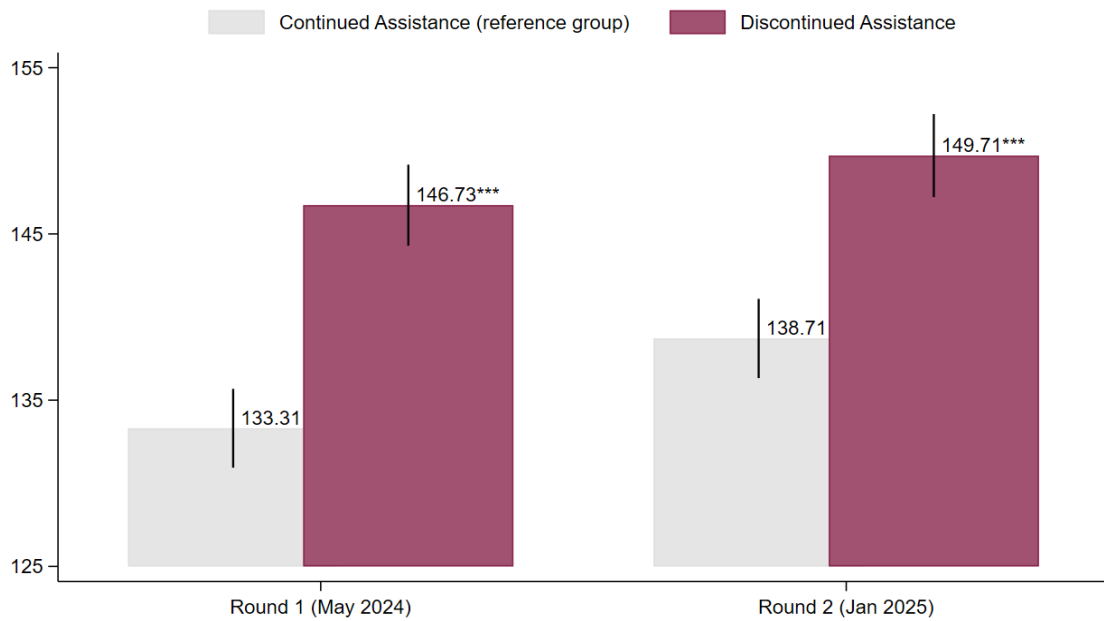
Note: The figure displays point estimates of the impact of discontinuation on the number of days each food group was consumed; outcome values range from 0 to 7 days. The y axis lists the groups, while the x axis shows differences in days consumed. Negative values indicate lower consumption by discontinued households than those who continued receiving assistance. Control variables were selected using post-double selection Lasso and always include eligibility-space fixed effects and a control for long notice households. For round 2, controls also include the random assignment of survey modules and survey quality. Horizontal lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

123. To better understand the consequences of the discontinuation of assistance on food security, the analysis disaggregates the overall FCS into its different food groups. These variables capture the number of days each food group was consumed in the households in the previous seven days. The results summarized in Figure 20 show reductions in food consumption in both rounds 1 and 2 for items included in the WFP in-kind assistance basket.
124. In the short run, the end of assistance caused reductions in the consumption of pulses by 8.4 percent (-0.33 days, $p < 0.01$), oil and fat by 2.9 percent (-0.17 days, $p < 0.01$), sugar by 5.9 percent (-0.24 days, $p < 0.01$), and condiments by 3.1 percent (-0.16 days, $p < 0.10$), while there was no clear impact on staple consumption. By round 2, these drops narrowed, but were still present, especially for pulses (-0.20 days, $p < 0.01$) and sugar (-0.16 days, $p < 0.10$), while there was a small increase in staple consumption, equivalent to 1.1 percent (+0.07 days, $p < 0.10$). No significant short- or medium-term changes were observed in the consumption of food groups not included in the WFP basket, such as vegetables, fruits, proteins or dairy.

Food expenditure

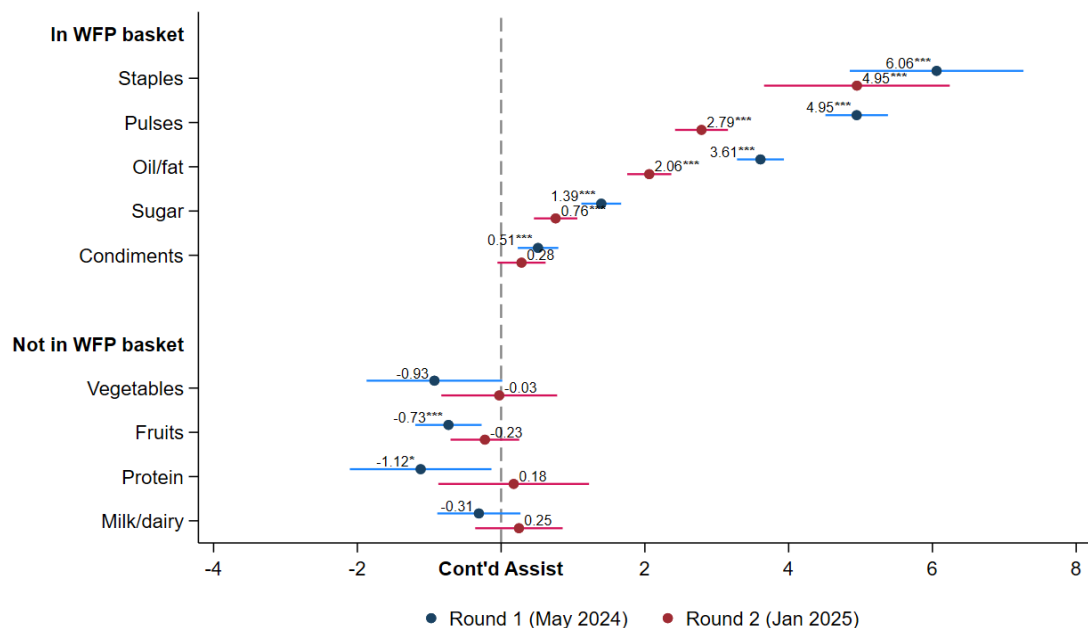
125. The consequences of the end of assistance are evident when looking at household food expenditure. Figure 21 compares monthly expenditure (in USD) on food between households that continued receiving assistance and those whose assistance was discontinued. In both rounds, discontinued households reported significantly higher food expenditure: in round 1, they spent on average USD 13.42 (10.1 percent) more per month, and in round 2, USD 11.00 (7.9 percent) more, compared to households that continued receiving in-kind food assistance.
126. When disaggregating food expenditure across categories, the higher spending is driven by items previously included in the WFP food basket. As shown in Figure 22, discontinued households spent an average of USD 6.06 (17.1 percent) more on staples in round 1 and USD 4.95 (13.2 percent) more in round 2 than households that continued receiving assistance. For pulses, the differences are USD 4.95 (94.1 percent) in round 1 and USD 2.79 (51.6 percent) in round 2. For oil and fat, they are USD 3.61 (64.9 percent) in round 1 and USD 2.06 (28.9 percent) in round 2. For sugar, they are USD 1.39 (30.7 percent) in round 1 and USD 0.76 (13.6 percent) in round 2. Spending on condiments is also higher by USD 0.51 (9.1 percent) in round 1.
127. In the short run there is some substitution across food items, with a drop in expenditure in items not included in the WFP basket: expenditure on fruits drops by USD 0.73 (7.3 percent) and expenditure on proteins by USD 1.12 (4.2 percent). All these effects are precisely estimated and statistically significant, with most of them being significant at the 1-percent level.

Figure 21: Impact of the discontinuation of assistance on food expenditure (USD)



Note: The figure presents the impact of discontinuation on household food expenditure, in US dollars. This includes food acquired through cash purchases, credit, and own production. The outcome reflects total spending over the previous 30 days. Expenditure was winsorized at the 95th percentile to reduce the influence of outliers. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects and a control for long notice households. For round 2, additional controls include the random assignment of survey modules and survey quality. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 22: Impact of the discontinuation of assistance on food expenditure by components of the food basket (USD)



Note: The figure shows point estimates of changes in household food expenditure due to discontinuation in USD. The y-axis lists food groups; the x-axis shows differences in spending. Values to the right indicate higher spending by discontinued households. Expenditure was winsorized at the 95th percentile to reduce the influence of outliers. Control variables were selected using post-double selection Lasso and always include eligibility-space fixed effects and a control for long notice households. For round 2,

additional controls include the random assignment of survey modules and survey quality. Horizontal lines indicate 90 percent confidence intervals. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

128. These are important results as they indicate that, on average, households have room to compensate for the end of assistance through their own resources. However, the estimated overall increase in expenditure corresponds to roughly 40 percent of the market value of a WFP parcel. This implies that, while some compensation takes place, on average, households end up with fewer overall food resources.⁹⁴ This explains the drop in food consumption in some of the food categories included in the assistance basket highlighted above.
129. Figure 22 also shows that expenditure on items outside the WFP basket – such as vegetables, fruits, protein and dairy – decreased in the short-term. Although the effects are not large, they indicate that, in the aftermath of losing assistance, households partly compensate for the increased expenditure on basic staples by lowering their expenditure on more diverse or nutrient-dense food groups. Over time, they manage to readjust their expenditure on these additional food items, closing the gap with continued assistance households.
130. Following discontinuation of support, consumption and expenditure patterns show that staples, such as cereals, continued to be consumed by households at relatively stable levels. This is likely due to their role as a primary calorie source, while nutrient-dense components of the basket saw sharper declines. Such food consumption patterns underscore a common coping strategy under resource constraints, where energy needs are prioritized over balanced and healthy diets.⁹⁵ The decline in the consumption of pulses, in particular, raises concerns about the potential for worsening nutritional outcomes in the long run, including increased risk of protein and micronutrient deficiencies.⁹⁶
131. Figure 23 also summarizes the impacts of the discontinuation of assistance on total and non-food expenditure. These are disaggregated by payment modality, either cash or credit, to illustrate the potential trade-offs that households may have implemented to compensate for the loss of assistance while maintaining food expenditure.
132. The increase in food expenditure for discontinued households is driven by purchases on credit, which is also reflected in higher total expenditure on credit. In round 1, total credit expenditure is USD 6.27 higher ($p < 0.01$) and, in round 2, USD 7.47 higher ($p < 0.05$) compared to households that continued receiving assistance, with no meaningful change in cash spending. For non-food expenditure, there is no statistically significant impact in round 1, but in round 2 there is a small yet significant increase in credit purchases (USD 4.74, $p < 0.10$). Total expenditure mirrors the rise in food expenditure, increasing by USD 15.80 ($p < 0.01$) in round 1 and by USD 10.77 in round 2, although the latter is not statistically significant.

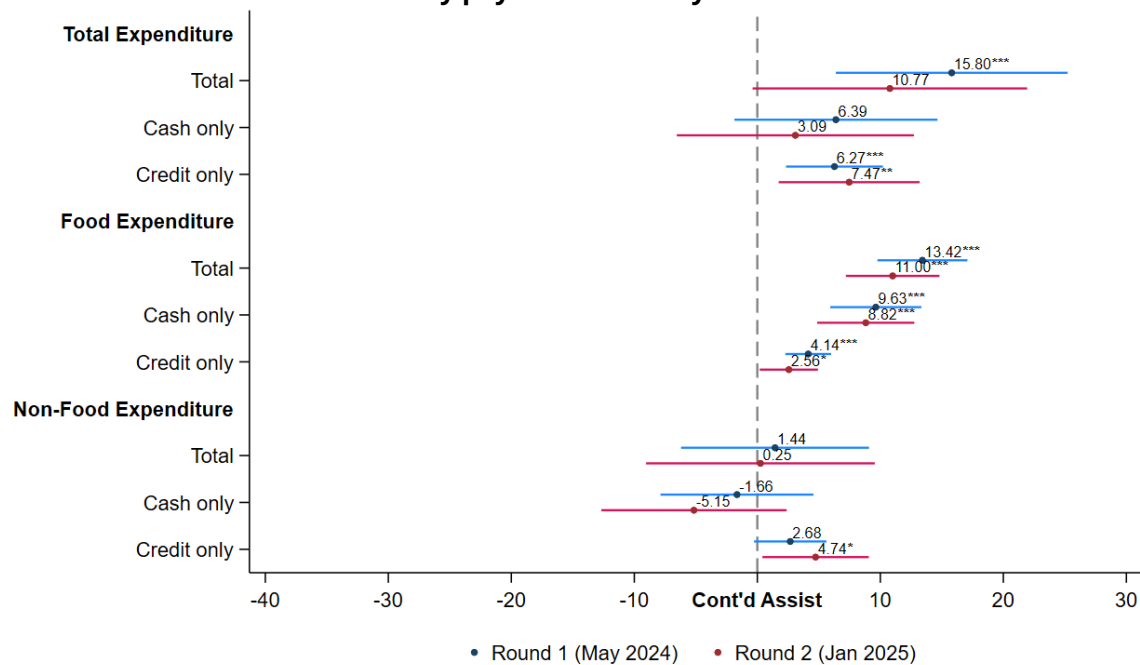
⁹⁴ In rounds 1 and 2, households receiving assistance (continued households) spent USD 159.75 and USD 156.99, respectively, which is higher than the expenditure of discontinued households.

⁹⁵ Food and Agriculture Organization (Ed.) 2008. *The State of Food Insecurity in the World 2008: High food prices and food security – threats and opportunities*.

⁹⁶ Tumen, S. 2016. The Economic Impact of Syrian Refugees on Host Countries: Quasi-experimental evidence from Turkey. *American Economic Review*, 106(5), 456-460.

133. As with food expenditure, this increase is entirely driven by the credit component. These findings show that households whose assistance was discontinued may have relied more heavily on credit to maintain overall spending levels after the loss of support.

Figure 23: Impact of the discontinuation of assistance on total expenditure (USD) by payment modality



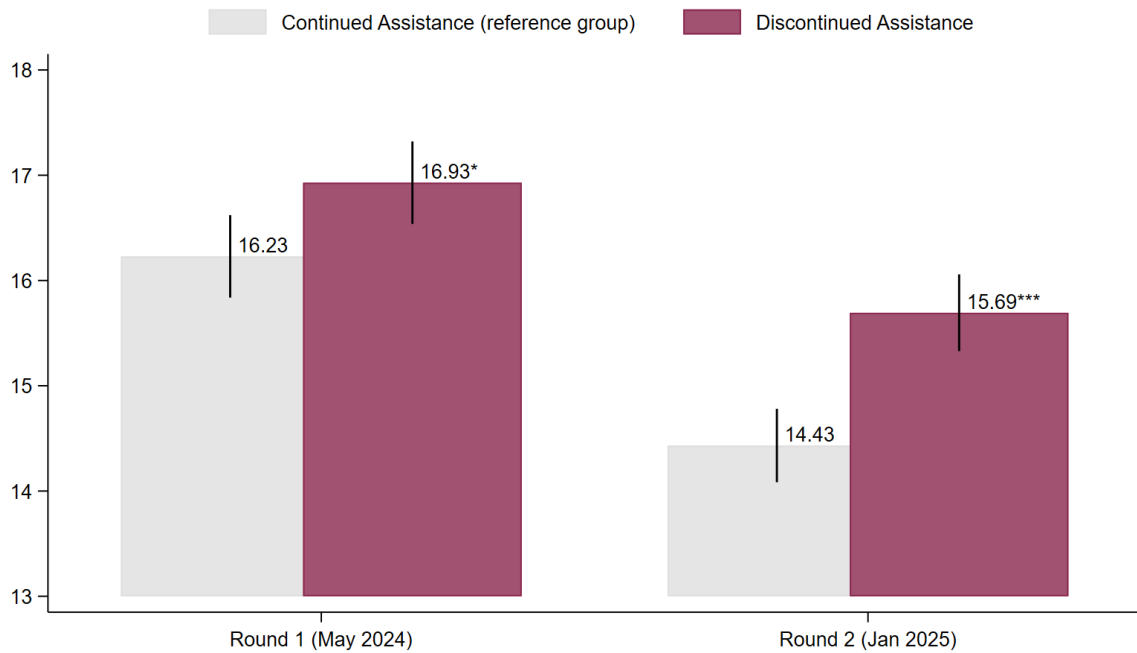
Note: The figure shows point estimates of changes in household expenditure (USD) due to discontinuation. The y-axis lists expenditure groups; the x-axis shows differences in spending. Values to the right indicate higher spending by discontinued households. Each estimate reflects the change in spending on a specific category. These categories correspond to food expenditure, one-month non-food expenditure, including more frequent and recurring items such as rent, electricity, heating, water, communication, and similar routine costs, and six-month household non-food expenditure, including education, clothing, celebrations, construction, agriculture, and other occasional expenses. Expenditure was winsorized at the 95th percentile to reduce the influence of outliers. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects and a control for long notice households. For round 2, additional controls include the random assignment of survey modules and survey quality. Horizontal lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Coping strategies

134. Given that the increase in food expenditure is, on average, not sufficient to fully compensate for the loss of in-kind assistance, households likely rely on coping strategies to address their needs.

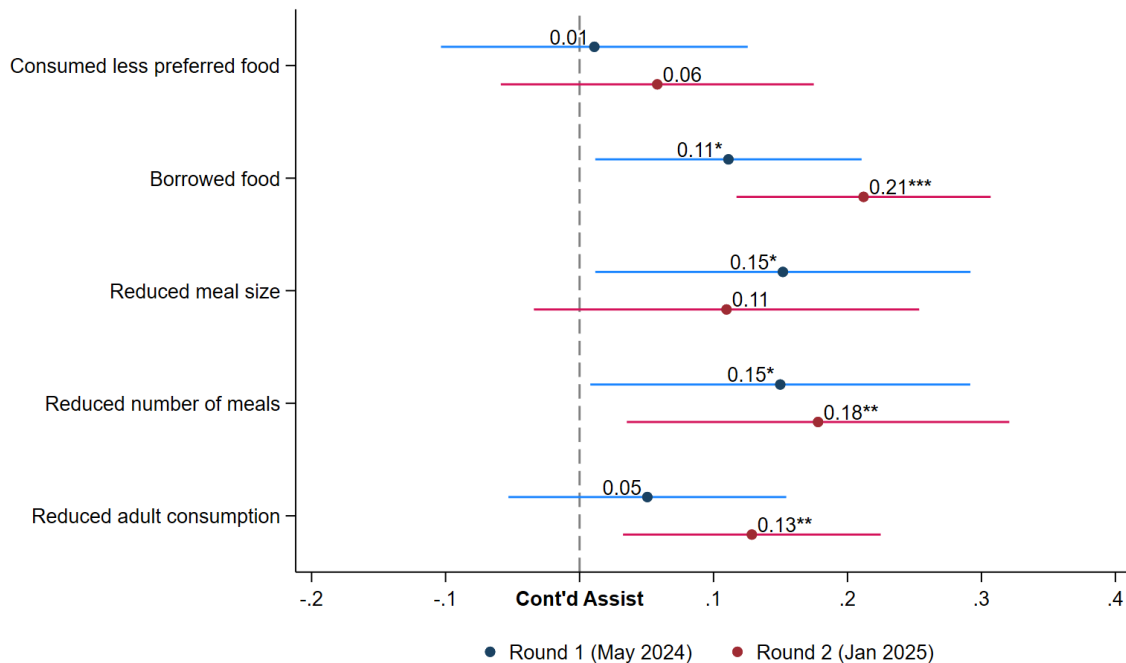
135. Figure 24 shows the impact of the discontinuation of assistance on the rCSI, where a higher value of the index indicates higher reliance on food-related harmful coping strategies. The results show a widening gap between continued and discontinued households over time. At round 1, the rCSI of discontinued households was on average 0.70 points higher than those who continued receiving assistance, equivalent to a 4.3-percent increase ($p < 0.1$). By round 2, this gap widened to 1.26 points or 8.7 percent ($p < 0.01$), though rCSI scores were lower overall in this round.

Figure 24: Impact of the discontinuation of assistance on rCSI



Note: The figure presents adjusted mean values of the reduced Coping Strategies Index, which ranges from 0 to 56. Higher scores indicate greater reliance on negative coping strategies. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects and a control for long notice households. For round 2, additional controls include the random assignment of survey modules and survey quality. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

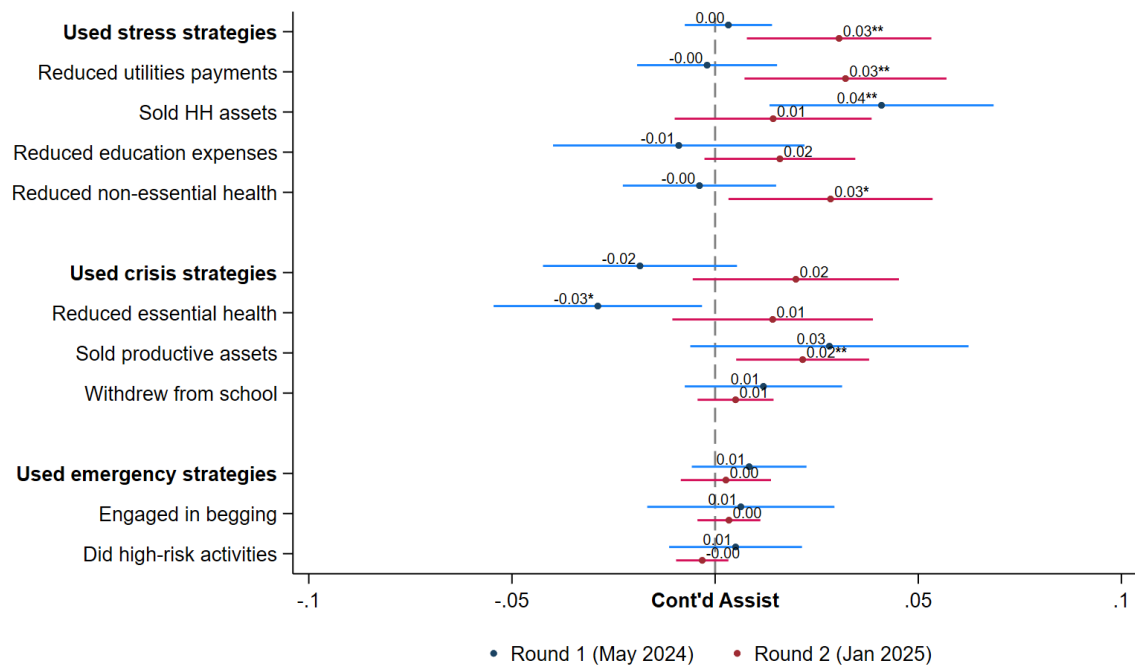
Figure 25: Impact of the discontinuation of assistance on rCSI components



Note: The figure displays the estimated impact of discontinuation on the frequency of consumption-related coping strategies. Each point represents the change in the number of days a specific strategy was adopted by households that discontinued assistance during the past seven days. Values to the right indicate more frequent use compared to households that continued receiving assistance. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects and a control for long notice households. For round 2, additional controls include the random assignment of survey modules and survey quality. Horizontal lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

136. Figure 25 illustrates the impact of discontinuation on the different components of the rCSI index. At round 1, discontinued households reported significantly higher reliance on borrowing of food (+0.11 days or 9.6 percent higher, $p < 0.1$), reducing meal size (+0.15 days or 5.4 percent higher ($p < 0.10$), and reducing the number of meals (+0.15 days or 5.5 percent higher, $p < 0.10$).
137. By round 2, the differences in reliance on consumption coping strategies became more pronounced: borrowing food was higher among the discontinued by 23.3 percent (+0.21 days, $p < 0.01$), reducing the number of meals increased by 7.3 percent for (+0.18 days, $p < 0.05$), and reducing adult food consumption increased by 15.3 percent for (+0.13 days, $p < 0.05$). These effects indicate a concerning widening gap over time following the removal of assistance.
138. While rCSI specifically captures food-related coping strategies, Figure 26 shows the consequences of the discontinuation of assistance on other livelihood coping strategies with a recall period of six months (LCSI questions).

Figure 26: Impact of the discontinuation of assistance on the probability of adopting negative livelihood-related coping strategies in the past six months



Note: The figure shows the estimated impact of discontinuation on the likelihood of adopting livelihood-related coping strategies. Each point reflects the change in probability for discontinued households, with values to the right indicating increased use relative to those who continued receiving assistance. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects and a control for long notice households. For round 2, additional controls include the random assignment of survey modules and survey quality. Horizontal lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

139. At round 1, discontinued households were 4 percentage points more likely to sell household assets compared to those who continued receiving assistance ($p < 0.05$), but otherwise showed no differences. By round 2, reliance on stress-related coping strategies persisted, with discontinued households being 3 percentage points more likely to adopt at

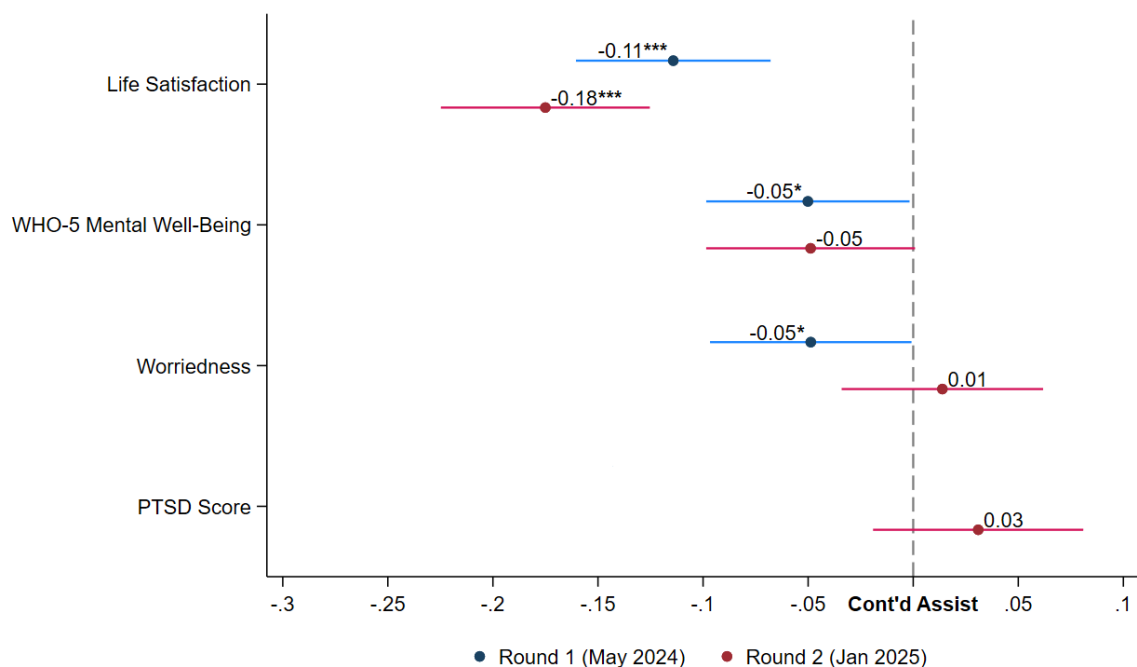
least one such strategy ($p < 0.05$). They were also 3 percentage points more likely to reduce utility payments ($p < 0.05$) and to cut non-essential health expenditures ($p < 0.1$). They were also 2 percentage points more likely to sell productive assets ($p < 0.05$).

140. These patterns indicate that, in the absence of food assistance, over time households increasingly resort to harmful coping mechanisms (stress strategies), by liquidating productive assets and by reducing spending on other (potentially essential) services. The emergency cash discontinued households received did not fully buffer these negative impacts (also, because part of the six-month recall period, households still went without the additional cash). On the positive side, there is no indication that households are pushed to adopt crisis or emergency coping strategies, such as engaging in begging or taking up high-risk activities.

Psychological well-being

141. By lowering the availability of food resources, the discontinuation of assistance might also negatively affect the psychological well-being of household members. Figure 27 shows the impact of the discontinuation of assistance on the four measures of psychological well-being included in this study.

Figure 27: Impact of the discontinuation of assistance on psychological well-being



Note: The figure displays estimated effects of discontinuation on standardized indices of psychological well-being. Each point shows the change in standard deviations due to discontinued assistance. Estimates to the left indicate lower scores. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects and a control for long notice households. For round 2, additional controls include the random assignment of survey modules and survey quality. Horizontal lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

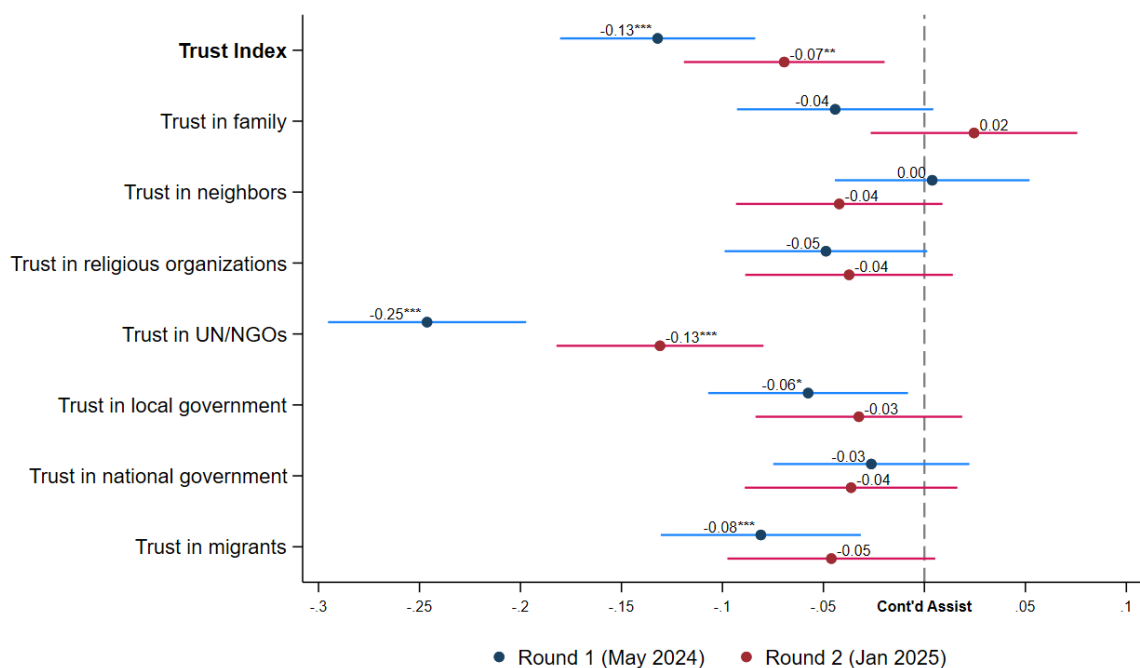
142. The results show that life satisfaction of discontinued households declined significantly in both rounds: by 0.11 standard deviations in round 1 ($p < 0.01$) and by 0.18 standard deviations in round 2 ($p < 0.01$). Similarly, mental well-being (as measured by the WHO-5

index) and the worriedness index both declined in the short term due to the discontinuation of assistance, with reductions of 0.05 standard deviations each in round 1 ($p < 0.1$). However, neither of these two effects persists by round 2, indicating a potential short-lived deterioration in these dimensions of psychological well-being.⁹⁷ For the PTSD score, which was only collected in round 2, there are no differences between discontinued and continued assistance households.

Trust and social cohesion

143. The discontinuation of assistance may have also affected households' levels of trust and social cohesion. Figure 28 presents the impact of the discontinuation of assistance on the composite trust index, which declined significantly by 0.13 standard deviations in round 1 ($p < 0.01$) and was still lower by 0.07 standard deviations ($p < 0.05$) in round 2. This overall decline is mainly driven by a very sharp drop in trust toward United Nations agencies and non-governmental organizations (NGOs), falling by 0.25 standard deviations in round 1 and 0.13 in round 2.⁹⁸

Figure 28: Impact of the discontinuation of assistance on trust



*Note: The figure shows estimated effects of discontinuation on standardized trust measures. All outcomes except the composite Trust Index are based on Likert-scale items. Points reflect changes in standard deviations, with values to the left indicating lower trust among discontinued households. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects and a control for long notice households. For round 2, additional controls include the random assignment of survey modules and survey quality. Horizontal lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.*

⁹⁷ When disaggregating the indices by their components, worriedness about food security remains significantly worse in round 2 for discontinued households.

⁹⁸ United Nations agencies and NGOs are grouped under the same umbrella (also in the original survey question), as extensive piloting suggested that respondents often struggled in distinguishing between the two groups.

144. While disappointment at the decision to interrupt assistance is expected, these results indicate that such disappointment translates into a mistrust towards the institutions that provided assistance in the first place. By round 2, one year had passed since the end of food assistance, and most discontinued households were receiving the new WFP emergency cash assistance. This result indicates that, once it is negatively impacted, trust is difficult to recover.
145. Figure 28 also shows that, in round 1, there was a significant reduction in trust toward the local government by 0.06 standard deviations ($p < 0.10$) and toward migrants by 0.08 ($p < 0.01$). The latter finding might be the result of a perception of unfairness, as assistance to Syrian refugees (provided by WFP through a different programme) continued throughout the study period. Although this effect appears short-lived and dissipated by round 2 (potentially due to the provision of the emergency cash support), it provides a warning signal, as such resentment can worsen some community relations, particularly in the context of Lebanon's ongoing economic and political crisis.⁹⁹ For both rounds, there is no significant impact of funding discontinuation on trust in family, neighbours, religious organizations, or the central government.
146. Effects on social cohesion outcomes are reported in Figure A7.8 in Appendix A7. In round 1, households that were discontinued from assistance were 6 percentage points less likely to report that a household member volunteered, compared to those who continued receiving support ($p < 0.05$). However, these effects were short-lived and discontinuation did not lead to significant impacts on other aspects of social cohesion, including households' perceived closeness to their community, their participation in community associations, or the presence and intensity of local tensions.

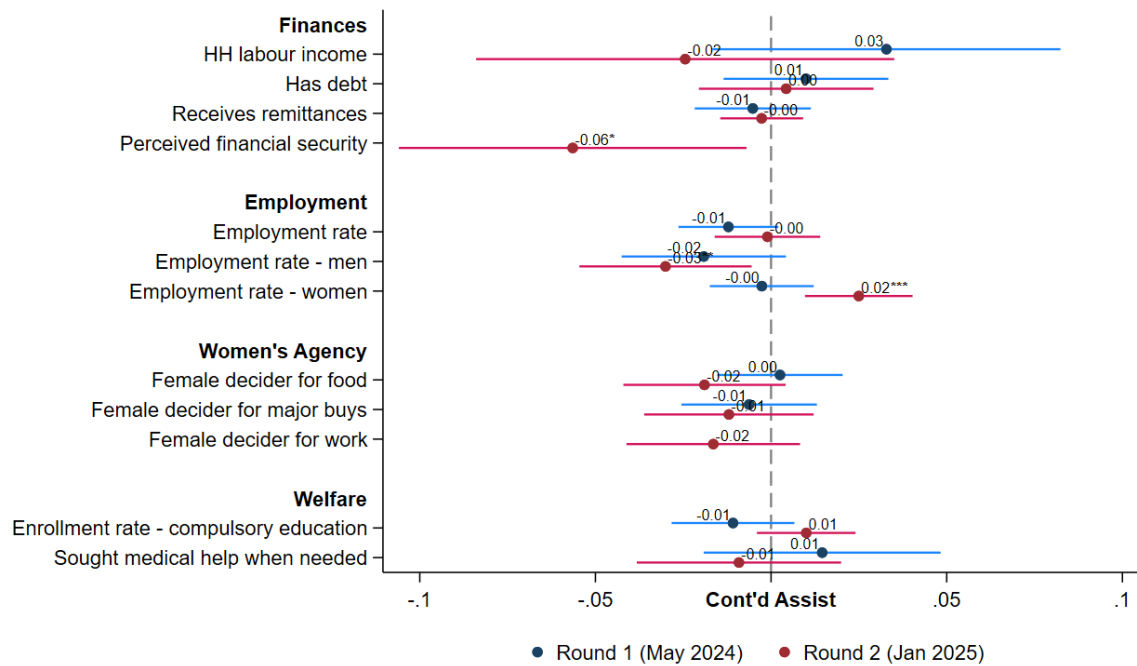
Socioeconomic and human development outcomes

147. Figure 29 shows the impact of the discontinuation on household finances, employment, women's agency, and general household welfare. For household finances, although there are no differences detected related to labour income, whether a household has debt, and whether a household receives remittances, discontinued households' perception of their financial security is worse than that of continued assistance households in round 2 (-0.06 sd, $p < 0.10$).¹⁰⁰ Similarly, there are no significant differences in any of the indicators related to women's agency, health, or schooling.
148. While employment rate did not differ between continued and discontinued households in either round, the composition of the workforce changed as a consequence of discontinuation: by round 2, the employment rate for men is 3 percentage points lower among the discontinued households ($p < 0.05$), while the employment rate for women is 2 percentage points higher ($p < 0.01$). What is motivating this change in household gender roles among the discontinued households is unclear and is left for further investigation with the LBCO.

⁹⁹ Tumen, S. 2016. The Economic Impact of Syrian Refugees on Host Countries: Quasi-experimental Evidence from Turkey. *American Economic Review*, 106(5), 456-460. <https://doi.org/10.1257/aer.p20161065>

¹⁰⁰ The question on perceived financial security was only asked during round 2.

Figure 29: Impact of the discontinuation of assistance on finances, employment, women’s agency, and general welfare



Note: The figure shows standardized point estimates of the impact of discontinuation, where values left of zero indicate worse outcomes for discontinued households. Except household labour income and perceived financial security which are both standardized, all other variables on the y-axis are based on proportions whose estimates indicate percentage point differences. Income was also winsorized at the 99th percentile to reduce the influence of outliers. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects and a control for long notice households. For round 2, additional controls include the random assignment of survey modules and survey quality. Horizontal lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.6 Which types of households are most affected by the discontinuation of assistance?

Summary of findings: Households that were deemed eligible for assistance under relatively more targeting methods were more severely affected by its discontinuation. They are less able to compensate for the end of assistance by increasing expenditure, more likely to rely on harmful coping strategies, and have worse food security outcomes.

When considering heterogeneity based on common sociodemographic criteria of vulnerability, the results generally show no significant differences.

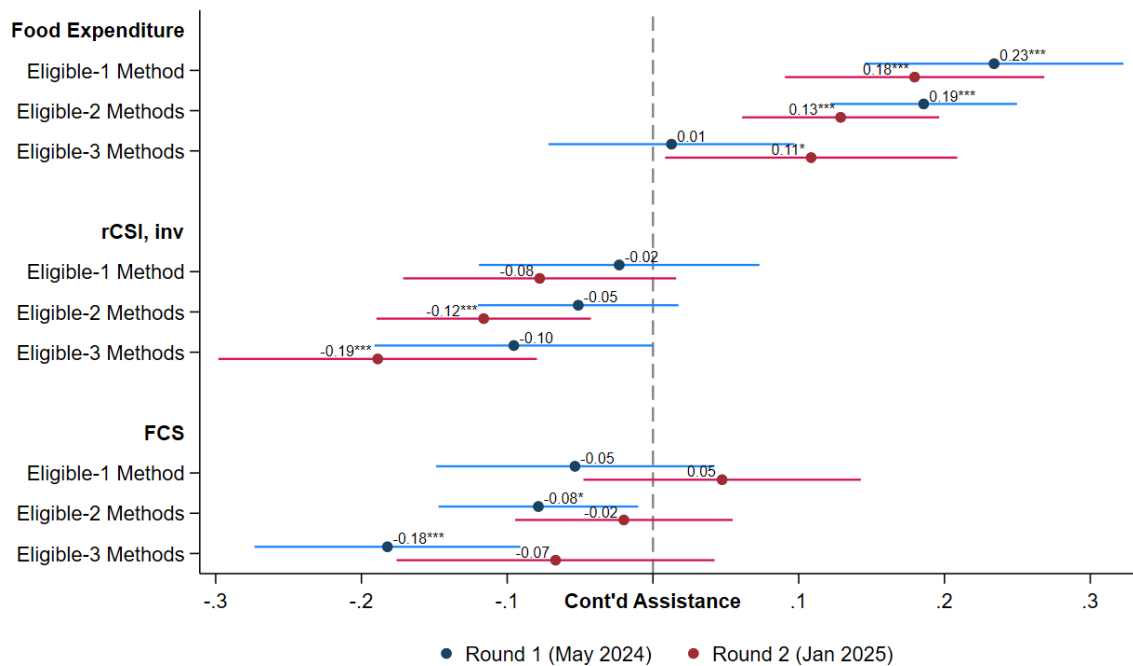
Heterogeneity in impact by eligibility

149. Figures 30 and 31 disaggregate the impact of discontinuation by the number of targeting methods under which a household is eligible.

150. A clear trend emerges in relation to food expenditure (see Figure 30): households eligible under only one method compensate for the discontinuation of in-kind assistance by increasing food expenditure by 0.2 standard deviations ($p < 0.01$). Households eligible under two methods also increase expenditure, though to a lesser extent (0.2 in round 1 and 0.1 in round 2, $p < 0.01$), while those eligible under three methods show no significant change in food expenditure in round 1 and increase food expenditure by 0.1 standard

deviations in round 2, indicating lower capacity to compensate for the assistance discontinuation with their own purchases.

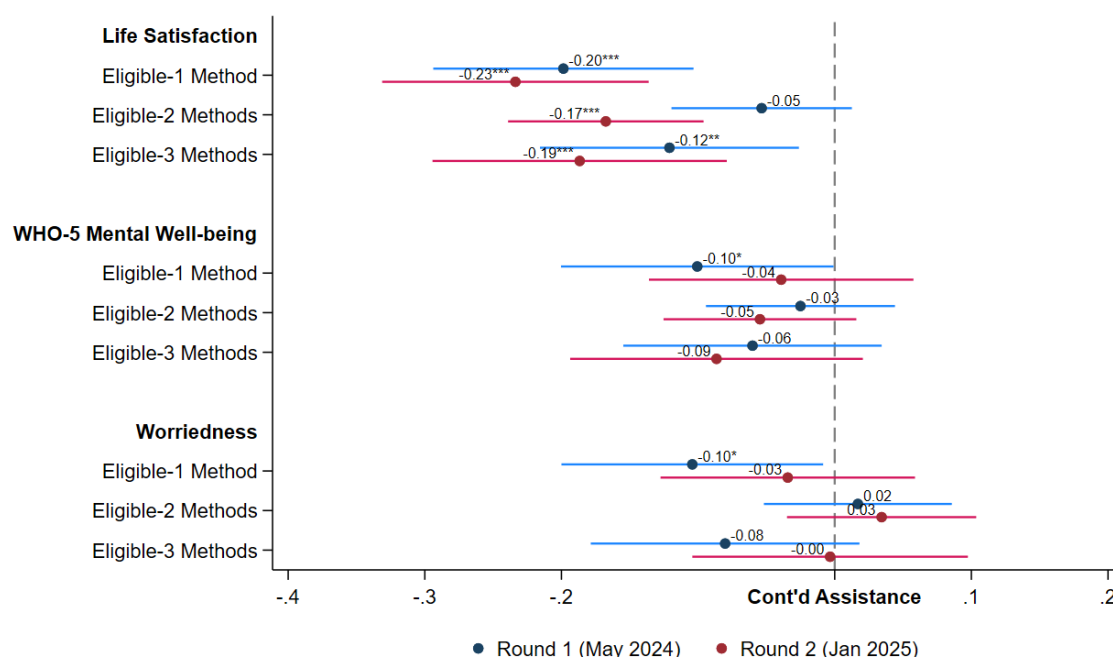
Figure 30: Impacts of discontinuation on food security by degree of eligibility



Note: The figure displays estimated effects of discontinuation on standardized indices of food security indicators from separate regressions, based on separate regressions by the number of targeting methods a household is eligible under. Each point shows the change in standard deviations due to discontinued assistance. Estimates to the left indicate lower outcomes. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects and a control for long notice households. For round 2, additional controls include the random assignment of survey modules and survey quality. Horizontal lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

151. A similar, though less pronounced, gradient is observed for rCSI and FCS, where households eligible under more methods show greater increases in food-related coping strategies and larger declines in food consumption. These findings show that households identified as eligible for assistance by multiple targeting methods are relatively more vulnerable and less able to compensate for the withdrawal of assistance. When these households are excluded from assistance, they suffer more severe negative consequences.
152. Turning to the well-being indicators in Figure 31, the relationship is less clear, with no consistent or striking gradient in the impact of discontinuation by eligibility. This shows that, unlike food-related outcomes, changes in psychological well-being indicators may be less sensitive to the degree of vulnerability.

Figure 31: Impacts of discontinuation on psychological well-being by degree of eligibility



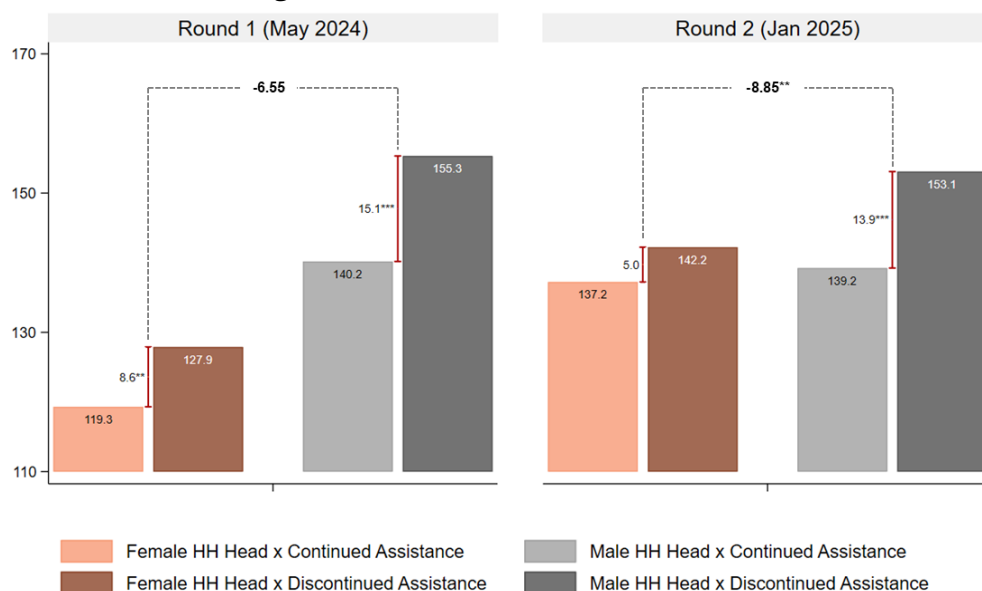
Note: The figure displays estimated effects of discontinuation on standardized indices of well-being indicators from separate regressions, based on separate regressions by the number of targeting methods a household is eligible under. Estimates to the left indicate lower outcomes. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects and a control for long notice households. For round 2, additional controls include the random assignment of survey modules and survey quality. Horizontal lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Heterogeneity by socioeconomic characteristics

153. In this section, the analysis focuses on whether the discontinuation of assistance had differential effects on three demographic criteria: gender of the household head; whether a household is below median wealth; and whether a household has a member with a disability. For simplicity, focus is on three key outcomes: food expenditure, FCS, and life satisfaction.
154. When considering households headed by women, although results for the three outcomes were somewhat lower compared to households headed by men, the differences were generally not statistically significant (see Figure 32 below and Figures A7.9 and A7.10 in Appendix A7). The exception was food expenditure in round 2.
155. As shown in the right panel of Figure 32, by round 2, among households headed by men, those who were discontinued from food aid increased their food expenditure by almost USD 14 ($p < 0.01$), but among households headed by women, while there is also a gap, no significant difference in expenditure was detected. This resulted in a difference in food expenditure of almost USD 9 ($p < 0.05$) between the different households by round 2. However, when observing these results it is important to keep in mind that most discontinued households in round 2 received emergency cash assistance and this finding is therefore difficult to compare with the round 1 finding. (Selection for cash assistance was based on geography and gender-related factors were not considered.)

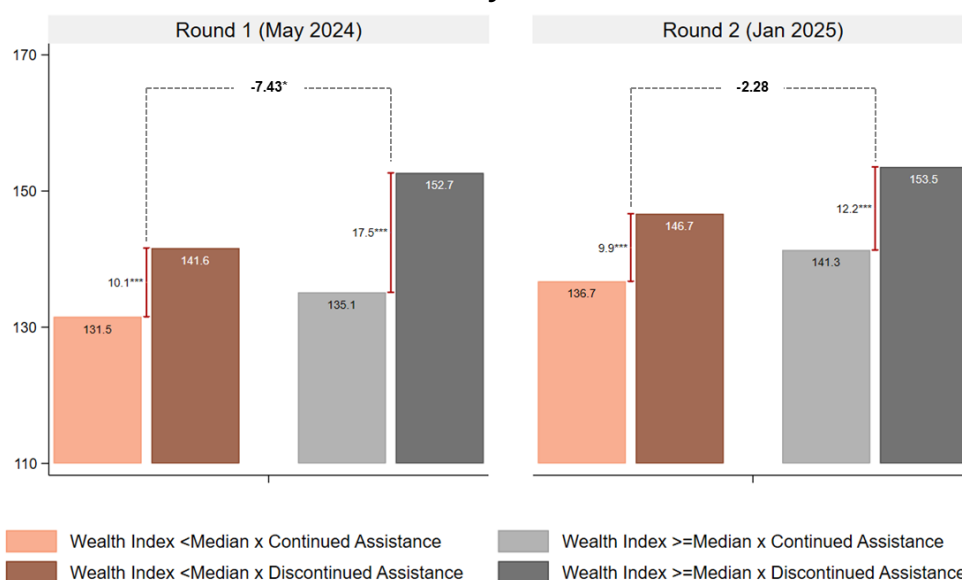
156. These findings may reflect more limited budgetary flexibility among households headed by women to compensate for the loss of food assistance. Although, given the lack of differences along the other outcomes, this result should be interpreted with caution.

Figure 32: Differential impacts of discontinuation on food expenditure in USD by gender of the household head



Note: The figures show mean values for combinations of household head gender (female or male) and programme status (continued or discontinued) for each round. The point estimates within each gender group indicate the group-specific average treatment effect. Differences between gender groups indicate heterogeneous treatment effects. Statistical significance is denoted by asterisks: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Figure 33: Differential impacts of discontinuation on food expenditure in USD by wealth

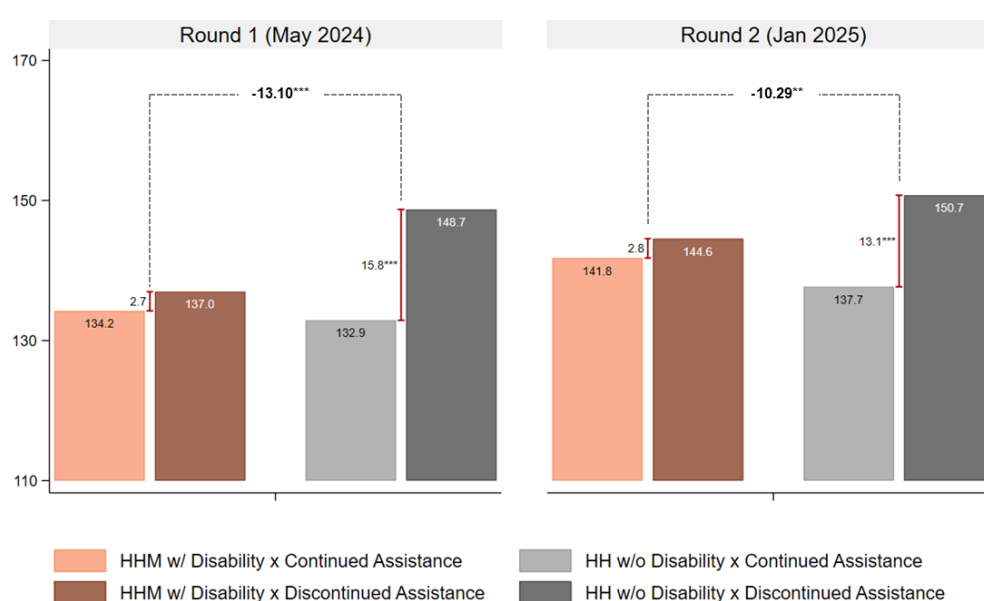


Note: The figures show mean values for combinations of household wealth (below or at/above median) and programme status (continued or discontinued) for each round. The point estimates within each wealth group indicate the group-specific average treatment effect. Differences between wealth groups indicate heterogeneous treatment effects. Statistical significance is denoted by asterisks: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

157. When examining the differential effects of discontinuation based on wealth – as measured through the baseline wealth index – no systematic differences emerged (see Figure 33 and Figures A7.11 and A7.12). Only food expenditure in round 1 was significant ($p < 0.10$), as shown in Figure 32.

158. Along the same lines, Figures 34, A7.13, and A7.14 present heterogeneous effects of discontinuation depending on whether a household member has a disability or chronic illness. No differences emerge in any of the survey rounds.

Figure 34: Differential impacts of discontinuation on food expenditure in USD by presence of a household member with a disability



Note: The figures show mean values for combinations of the presence of disability (having or not having a household member with a disability/chronic illness) and programme status (continued or discontinued) for each round. The point estimates within each disability group indicate the group-specific average treatment effect. Differences between disability groups indicate heterogeneous treatment effects. Statistical significance is denoted by asterisks: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Part 3: Whether and how households prepare for programme exit

159. Given the adverse consequences of the discontinuation of assistance, it is important to understand whether and how households that are given six months advance notice might prepare for it, and whether they manage to smooth their consumption. Given that they are informed that support will be disrupted, do they reduce present consumption to save resources for the near future? This section focuses on households that were randomly assigned to receive an extended six-month window of assistance (until June 2024). It compares their food security, expenditure patterns, coping behaviours, and psychological well-being to: (i) continued assistance households that kept receiving assistance throughout the study period (beyond June 2024); and (ii) discontinued assistance households that stopped receiving assistance in January 2024 (following traditional protocols that notified them the month assistance ended).

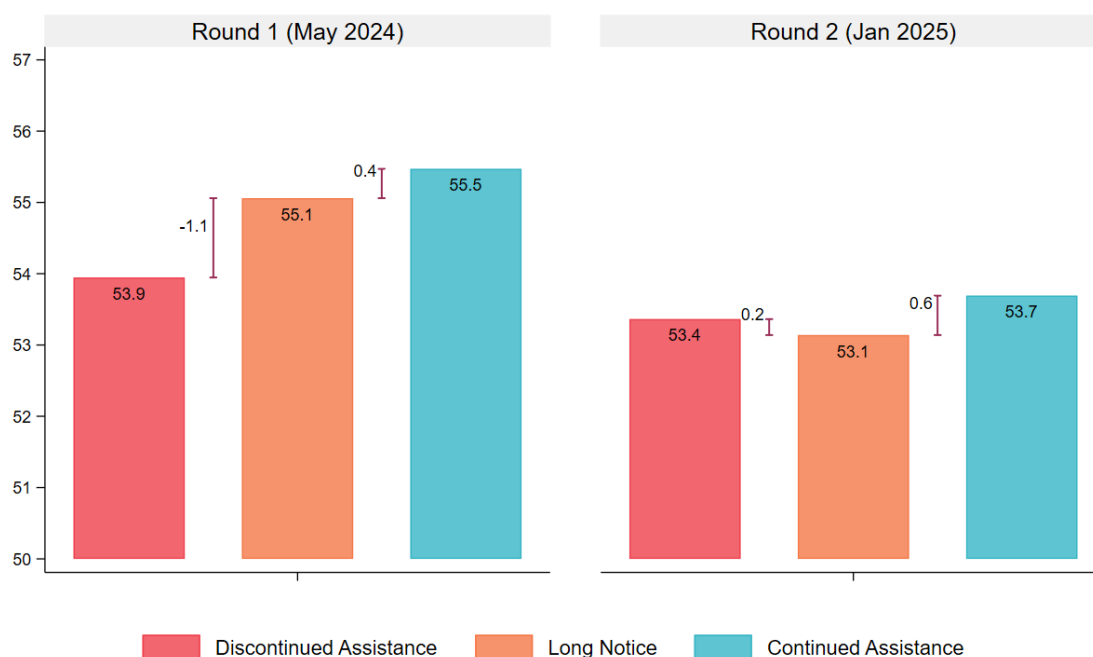
6.7 Do households prepare for the end of assistance?

Summary of findings: Comparing “long notice” households to continued and immediately discontinued recipients, the analysis finds no evidence that early warning led to early preparation and improved later outcomes. While long notice households resembled continued recipients in the short run, they converged with discontinued households by the end of the study period. By round 2, indeed, long notice households exhibited food security, expenditure, coping behaviour, and psychological well-being outcomes very similar to those of households who lost assistance without advance notice. This underscores the need for more effective strategies to support households transitioning off humanitarian assistance in fragile settings.

160. Figure 35 compares the FCS for the three groups of households: discontinued assistance households (in red); long notice households that received assistance until June 2024 (in orange); and continued assistance households (in blue). No significant differences in the FCS are observed across groups or survey rounds. This is not too surprising, given the overall small impact of discontinuation of assistance on FCS documented above.

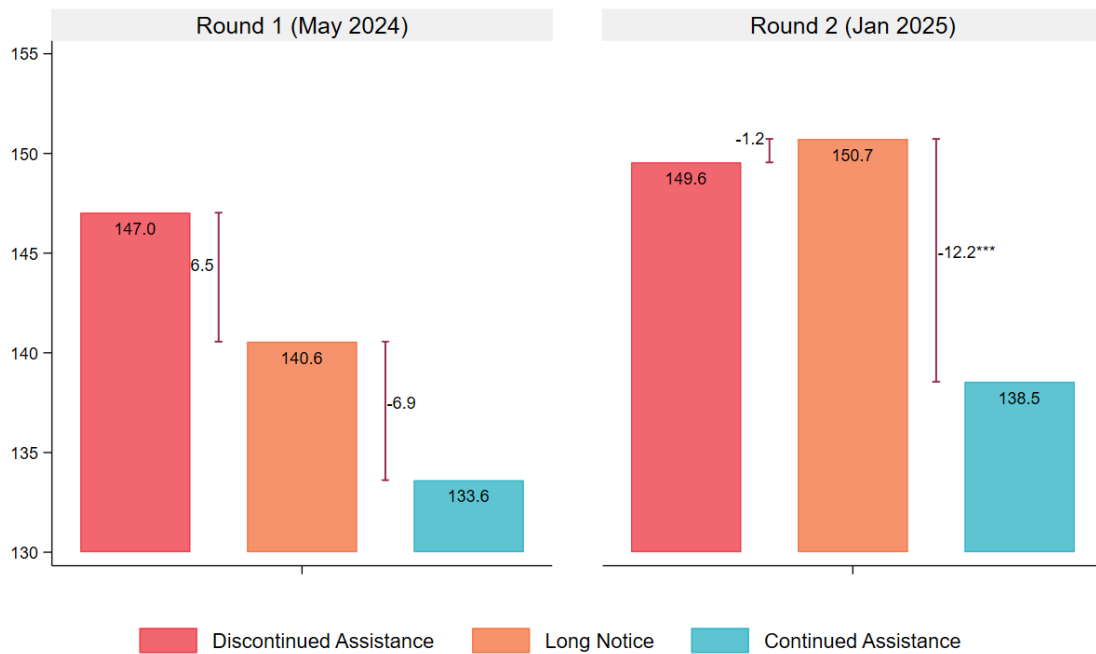
161. Figure 36 focuses on food expenditure. In round 1, long notice households showed no significant difference in spending compared to either discontinued or continued assistance households. By round 2, their expenditure closely matched that of discontinued households, exceeding that of continued assistance households by USD 12.2 ($p < 0.01$). This indicates that by round 2, long notice households were spending at comparable levels to discontinued households to offset the loss of assistance.

Figure 35: FCS for long notice, continued and discontinued households



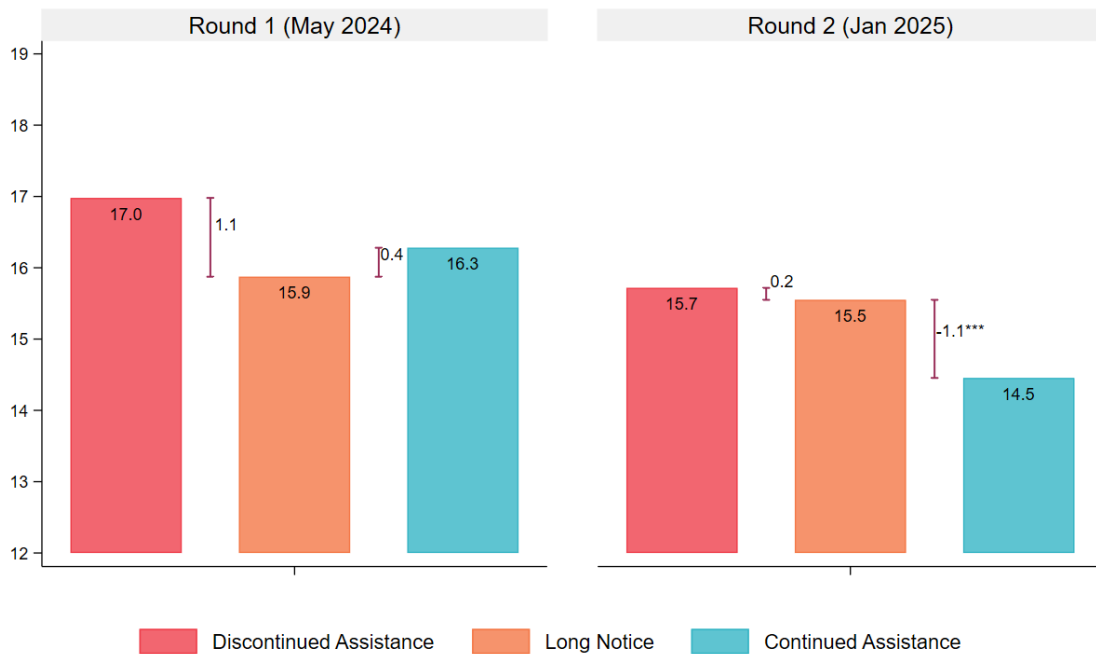
Note: The figure presents the adjusted mean of the FCS, which ranges from 0 to 112. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects. For round 2, additional controls include the random assignment of survey modules and survey quality. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 36: Food expenditure (USD) for long notice, continued and discontinued households



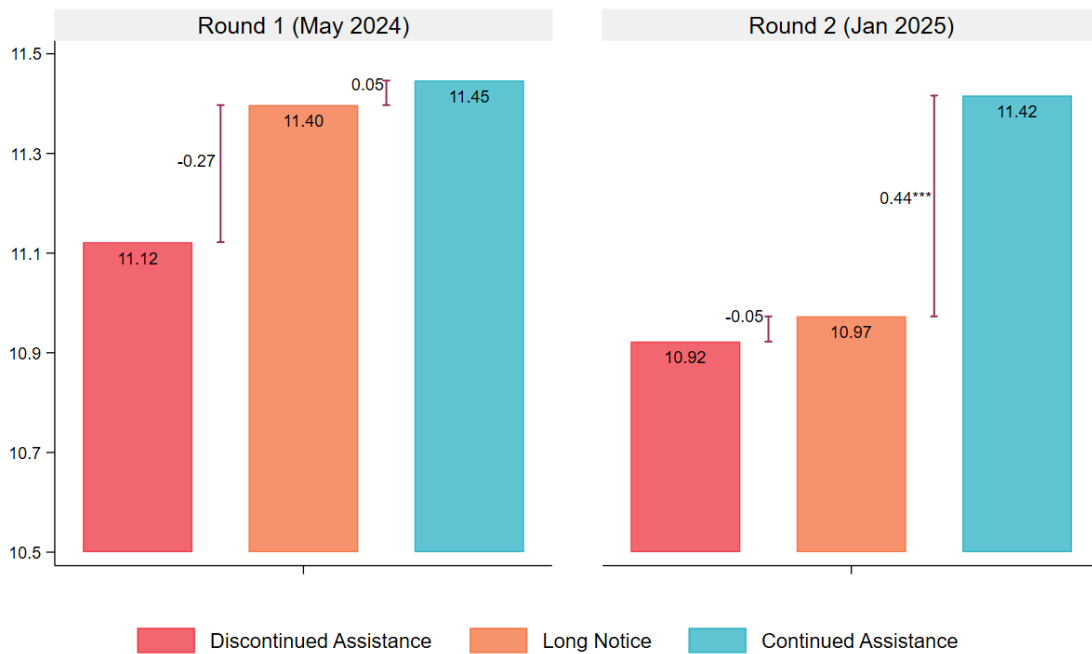
Note: The figure presents the adjusted mean of food expenditure in USD. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects. For round 2, additional controls include the random assignment of survey modules and survey quality. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 37: rCSI for long notice, continued and discontinued households



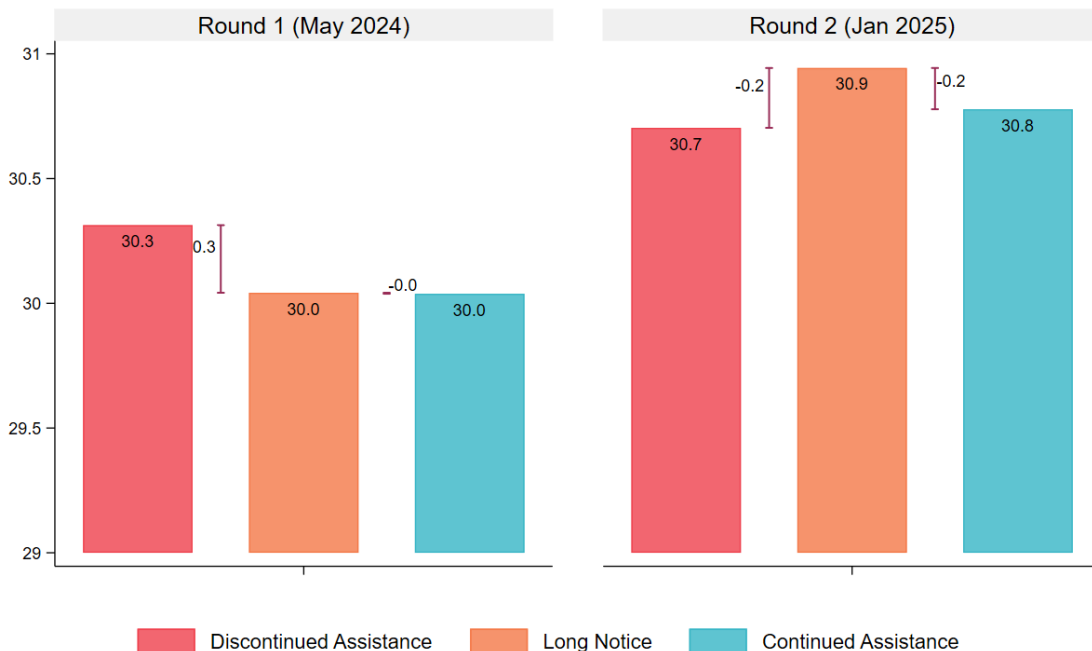
Note: The figure presents adjusted mean values of the reduced Coping Strategies Index, which ranges from 0 to 56. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects. For round 2, additional controls include the random assignment of survey modules and survey quality. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 38: Life satisfaction long notice, continued and discontinued households



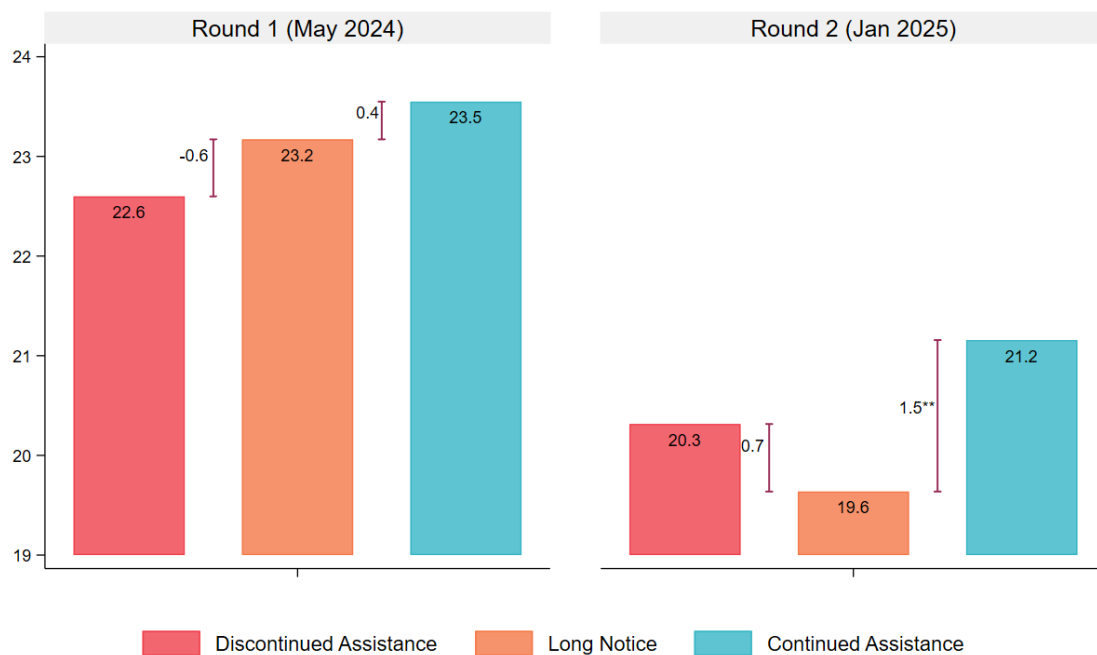
Note: The figure presents the life satisfaction index, which ranges from 5 to 20. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects. For round 2, additional controls include the random assignment of survey modules and survey quality. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 39: Mental well-being (WHO-5) for long notice, continued and discontinued households



Note: The figure presents the WHO-5 index, which ranges from 0 to 100. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects. For round 2, additional controls include the random assignment of survey modules and survey quality. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 40: Worriedness for long notice, continued and discontinued households



Note: The figure presents the worriedness index, which ranges from 10 to 40. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects. For round 2, additional controls include the random assignment of survey modules and survey quality. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

162. Figure 37 focuses on household coping strategies related to food consumption, as captured through rCSI. While at round 1 there is no clear difference across groups, by round 2, households in the long notice group appear indistinguishable from households in the (short notice) discontinued group. Both groups report higher use of coping strategies compared to households whose assistance continued. Taken together, the results on expenditure and rCSI indicate that long notice households did not change behaviour in anticipation of the discontinuation of assistance, and by round 2 exhibited the same conditions as households that had stopped receiving assistance six months earlier.
163. It is still possible that the longer notice period allowed households to prepare psychologically. To explore this, Figure 38 presents the results on life satisfaction. These confirm the previous results: by round 1, long notice households were not different from other groups, but by round 2, they exhibit a drop in life satisfaction compared to continued assistance households of a similar magnitude as other discontinued households.
164. The evolution of mental well-being, reported in Figure 39, shows a similar pattern, although in this case, by round 2 the long notice group displays even worse mental well-being than the discontinued group (although the difference is not statistically significant). Finally, Figure 40, shows no clear pattern in terms of worriedness, as differences across groups are small and not statistically significant in either round.
165. These findings indicate that, during round 1, households in the long notice window (which, by then, were still receiving assistance) behave similarly to households assigned to

continue receiving assistance. By round 2, they look generally indistinguishable from households that had stopped receiving assistance six months earlier.

166. These results appear in line with some existing evidence from fragile settings and point to households' inability to save and smooth the adverse consequences of negative shocks, even when they are anticipated.^{101,102,103} This indicates that alternative strategies and tools are needed to support households' transition out of assistance.

¹⁰¹ Goldin, J., Homonoff, T., & Meckel, K. 2022. Issuance and Incidence: SNAP Benefit Cycles and Grocery Prices. *American Economic Journal: Economic Policy*, 14(1), 152–178. <https://doi.org/10.1257/pol.20190777>

¹⁰² Hastings, J., & Washington, E. 2010. The first of the Month Effect: Consumer Behaviour and Store Responses. *American Economic Journal: Economic Policy*, 2(2), 142–162. <https://doi.org/10.1257/pol.2.2.142>

¹⁰³ Shapiro, J. M. 2005. Is there a Daily Discount Rate? Evidence from the Food Stamp Nutrition Cycle. *Journal of Public Economics*, 89(2–3), 303–325.

7. Conclusions and considerations for future programming

7.1. Conclusions

Part 1: Prioritizing among households already being supported when scaling down

167. **The four targeting approaches do not lead to substantial differences in terms of average food security, psychological well-being, social cohesion, and socioeconomic indicators among the target population based on their overall averages.** These findings are consistent with the limited evidence from experimental assessments on targeting, which generally show that, while different approaches lead to the *selection* of different groups, such variation does not necessarily translate into differences in programme effectiveness as measured by the average well-being of the target population.
168. **Categorical targeting underperforms with respect to the three PMT-based approaches in terms of higher appeal rates and their selection of more vulnerable households.** There were substantially more appeals among discontinued households under Categorical targeting. This indicates that reliance on coarse sociodemographic criteria, rather than the more granular information embedded in PMT methods, can generate greater targeting errors, and reduce the perceived legitimacy of eligibility decisions. Comparisons of baseline wealth indicate that households discontinued under the Categorical approach are more likely to be asset-poor than those excluded under PMT-based methods. Consistent with this evidence, when comparing continued versus discontinued households, the Categorical method behaved differently from all three PMT approaches. The PMT methods appeared more likely to retain households that were worse-off – according to rCSI, food expenditure, and worriedness indicators – whereas the Categorical approach did not show this pattern.
169. **The probability of appealing the discontinuation decision increases with the number of methods under which a household qualifies.** Discontinued households that were deemed eligible by multiple other methods were more likely to appeal their exclusion. In the reference group – discontinued households not identified as eligible by any method – exactly 50 percent submitted an appeal. This rate rises to 62 percent among discontinued households deemed eligible under three methods.

Part 2: The consequences for households cut from assistance

170. **In the short term (six months), the discontinuation of food assistance leads to lower food consumption, greater reliance on harmful coping strategies (including borrowing or purchasing food on credit), declines in life satisfaction and mental well-being, and reduced trust in aid agencies.** While some of these effects attenuate in the medium term (12 months), reductions persist in the consumption of specific food groups included in the WFP basket, such as pulses. Discontinued households partially compensate for the loss of assistance by increasing food expenditures, particularly on items previously

provided by WFP, but these increases are generally insufficient to fully offset the lost transfers.

171. Over time, discontinued households increasingly resort to negative coping strategies. Psychological well-being also deteriorates, with lower life satisfaction persisting a year after discontinuation. Finally, trust in aid agencies fell sharply as a result of discontinuation and remained low 12 months after the end of assistance. Overall, while households exhibit some capacity to adjust, the withdrawal of assistance generates substantial and lasting adverse effects on both socioeconomic conditions and mental well-being.
172. **More vulnerable households suffer the most from the end of assistance.** When comparing impacts on different segments of the population, the results show that more vulnerable households are also more severely affected by the discontinuation of assistance. For example, the impacts on the economic well-being of discontinued households that were deemed vulnerable under three targeting methods were relatively stronger. They are less able to increase spending on food to compensate for the end of assistance; and more likely to rely on harmful coping strategies and have lower food security. The findings highlight how socioeconomic constraints limit the ability of households to adapt to the end of programmes.

Part 3: Whether and how households prepare for programme exit

173. **Early notice does not mitigate the negative impact of the end of assistance.** The “long notice” analysis shows no clear evidence that a longer warning window helped households prepare for the end of assistance and improve later outcomes. While long notice households behaved similarly to the households that continued to receive assistance in the short run, they converged with discontinued households by the end of the study period. When their assistance ended, they exhibit food security, expenditure, coping behaviour, and psychological well-being outcomes very similar to households who lost assistance without advance notice. These results indicate that households did not manage to smooth consumption in anticipation of the shock. This is likely due to limited ability to save or store resources. This underscores the need for more effective strategies to support households transitioning off humanitarian assistance in fragile settings, such as in Lebanon.

7.2. Key considerations

174. Although grounded in a specific context, the findings from this study have broader relevance for humanitarian operations globally. Based on the evidence, four key considerations for future WFP programming are outlined:

Consideration 1: Combining targeting methods to improve accuracy and fairness

175. The analysis shows that greater accuracy can be achieved by combining multiple targeting methods and prioritizing households that are consistently identified as vulnerable across approaches. Institutionalizing this approach in WFP operations, as already piloted in Lebanon with new programmes based on these results, could enhance fairness and effectiveness in future scale-up and scale-down efforts.

Consideration 2: Assessing trade-offs between precision and practicality

176. In the current global scale-down in humanitarian funding, it is critical to ensure that targeting accurately selects the most vulnerable. Although the four methods yield similar average well-being in the overall target population, Categorical targeting produces larger well-being gaps between continued and discontinued households and is associated with higher rates of appeals. These findings indicate lower targeting precision. At the same time, Categorical targeting is generally easier to apply in contexts where extensive survey data is lacking or data collection is infeasible. Policymakers thus need to balance the lower precision with the potentially higher cost of conducting PMT-based targeting. When more detailed household data is available, combining a prediction model approach with the simpler categorical scoring is expected to improve selection precision.

Consideration 3: Developing more deliberate and evidence-based exit strategies to maintain household well-being

177. When assistance is scaled back after the peak of an emergency, efforts are made to retain the most vulnerable households. Those who have assistance discontinued experience significant adverse effects, both in the short and medium term. This result points to a broader structural issue in the design of humanitarian (and development) programmes: most programmes lack clear exit strategies and their continuation or termination is typically dictated by funding cycles rather than rigorous assessments of ongoing need. Developing more deliberate and evidence-based exit strategies could reduce the risk of harm when assistance ends and help ensure that the gains made during the programme are not lost.

Consideration 4: More careful support to transition households towards the end of assistance

178. This study represents a first step toward understanding how to implement more effective exit strategies. It finds that simply providing advance notice of discontinuation does little to help households smooth consumption or adjust effectively. This indicates the need for alternative approaches that minimize the disruption caused by the end of the programme. Potential interventions could include behavioural tools that help households plan for the transition, psychological support, or temporary bridging assistance to facilitate the transition. Such strategies may also help sustain trust in humanitarian actors, which appears to deteriorate sharply when assistance ends abruptly.

8. Bibliography

- Afridi, F., Abhiroop M., & Soham Sahoo. 2016. Female Labor Force Participation and Child Education in India: Evidence from the National Rural Employment Guarantee Scheme. *IZA Journal of Labor & Development*, Vol 5, no. 1: 7. <https://doi.org/10.1186/s40175-016-0053-y>.
- Aiken, E. et al. 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., Ohlenburg, T., & Blumenstock, J. 2023. *Moving targets: When does a poverty prediction model need to be updated?* Proceedings of the 6th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies, 117–117. <https://doi.org/10.1145/3588001.3609369>
- Alatas, V. et al. 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. et al. 2012. Targeting the Poor: Evidence from a Field Experiment in Indonesia. *American Economic Review*, 102(4), 1206–1240. <https://doi.org/10.1257/aer.102.4.1206>
- Altındağ, O. et al. 2021. Targeting humanitarian aid using administrative data: Model design and validation. *Journal of Development Economics*, 148, 102564. <https://doi.org/10.1016/j.jdeveco.2020.102564>
- Anderson, M. B. 1999. *Do No Harm: How Aid Can Support Peace--Or War*. Lynne Rienner Publishers.
- 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>
- Araar, A., & Tiberti, L. 2020. *Optimal Targeting and Poverty Reduction with fixed budget and under imperfect information*. [https://dasp.ecn.ulaval.ca/dasp3/refs/Optimal Targeting and Poverty Reduction april2020.pdf](https://dasp.ecn.ulaval.ca/dasp3/refs/Optimal%20Targeting%20and%20Poverty%20Reduction%20april2020.pdf)
- Banerjee, A. V. et al. 2017. Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programmes. *The World Bank Research Observer*, 32(2), 155–184. <https://doi.org/10.1093/wbro/lkx002>
- Beaman, L. et al. 2021. Can Network Theory-Based Targeting Increase Technology Adoption? *American Economic Review*, 111(6), 1918–1943. <https://doi.org/10.1257/aer.20200295>
- Bossuroy, T. et al. 2022. Tackling psychosocial and capital constraints to alleviate poverty. *Nature*, 605(7909), 291–297. <https://doi.org/10.1038/s41586-022-04647-8>
- Brück, T. et al. 2019. Assets for Alimentation? The Nutritional Impact of Assets-based Programming in Niger. *The Journal of Development Studies*, 55(sup1), 55–74. <https://doi.org/10.1080/00220388.2019.1687876>
- Bruni, V., & Sterck, O. 2025) The Welfare and Market of Effects of Delays in Humanitarian Assistance (SSRN Scholarly Paper No. 5190878). *Social Science Research Network*. <https://doi.org/10.2139/ssrn.5190878>

- Cameron, L., & Shah, M. 2014. Can Mistargeting Destroy Social Capital and Stimulate Crime? Evidence from a Cash Transfer Programme in Indonesia. *Economic Development and Cultural Change*, 62(2), 381–415. <https://doi.org/10.1086/674102>
- Caria, S. et al. 2020. An Adaptive Targeted Field Experiment: Job Search Assistance for Refugees in Jordan. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3689456>
- Cloitre, M. et al. 2018. The International Trauma Questionnaire: Development of a self-report measure of ICD-11 PTSD and Complex PTSD. *Acta Psychiatrica Scandinavica*. DOI: 10.1111/acps.12956
- Crosta, T. et al. 2024. *Unconditional Cash Transfers: A Bayesian Meta-Analysis of Randomized Evaluations in Low and Middle Income Countries* (Working Paper No. 32779). National Bureau of Economic Research. <https://doi.org/10.3386/w32779>
- Deaton, A., & Zaidi, S. 2002. *Guidelines for Constructing Consumption Aggregates for Welfare Analysis* (No. No. 135; LSMS Working Paper). World Bank. <http://hdl.handle.net/10986/14101>
- Delhey, J. & Welzel, C. 2012. Generalizing Trust: How Outgroup-Trust Grows Beyond Ingroup-Trust (October 18, 2012). *World Values Research*, WVR Volume 5, Number 3. <http://dx.doi.org/10.2139/ssrn.2390636>
- Devereux, S. et al. 2017. The targeting effectiveness of social transfers. *Journal of Development Effectiveness*, 9(2), 162–211. <https://doi.org/10.1080/19439342.2017.1305981>
- Dustmann, C., Landersø, R., & Andersen, L. H. 2024a. Refugee Benefit Cuts. *American Economic Journal: Economic Policy*, 16(2), 406–441. <https://doi.org/10.1257/pol.20220062>
- Dustmann, C., Landersø, R., & Andersen, L. H. 2024b. Unintended Consequences of Welfare Cuts on Children and Adolescents. *American Economic Journal: Applied Economics*, 16(4), 161–185. <https://doi.org/10.1257/app.20230519>
- Follett, L., & Henderson, H. 2023. A hybrid approach to targeting social assistance. *Journal of Development Economics*, 160, 103002. <https://doi.org/10.1016/j.jdeveco.2022.103002>
- Food and Agriculture Organization (Ed.). 2008. *The state of food insecurity in the world 2008: High food prices and food security - threats and opportunities*. FAO.
- Friedman, M. 1957. *The Permanent Income Hypothesis*. NBER Chapters, 20–37.
- Gelman, M. et al. 2014. Harnessing naturally occurring data to measure the response of spending to income. *Science*, 345(6193), 212–215. <https://doi.org/10.1126/science.1247727>
- Goldin, J., Homonoff, T., & Meckel, K. 2022. Issuance and Incidence: SNAP Benefit Cycles and Grocery Prices. *American Economic Journal: Economic Policy*, 14(1), 152–178. <https://doi.org/10.1257/pol.20190777>
- Gram-Hansen, B. J. et al. 2019. *Mapping Informal Settlements in Developing Countries using Machine Learning and Low Resolution Multi-spectral Data*. Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, 361–368. <https://doi.org/10.1145/3306618.3314253>
- Grosh, M. et al. 2022. Revisiting Targeting in Social Assistance: A New Look at Old Dilemmas. The World Bank. <https://doi.org/10.1596/978-1-4648-1814-1>

- Hall, O. et al. 2023. A review of machine learning and satellite imagery for poverty prediction: Implications for development research and applications. *Journal of International Development*, 35(7), 1753–1768. <https://doi.org/10.1002/jid.3751>
- Hastings, J., & Washington, E. 2010. The First of the Month Effect: Consumer Behaviour and Store Responses. *American Economic Journal: Economic Policy*, 2(2), 142–162. <https://doi.org/10.1257/pol.2.2.142>
- Haushofer, J. et al. 2025. Targeting Impact versus Deprivation. *American Economic Review*, 115(6), 1936–1974. <https://doi.org/10.1257/aer.20221650>
- Hidrobo, M., Karachiwalla, N., & Roy, S. 2023. *The Impacts of Cash Transfers on Mental Health and Investments: Experimental evidence from Mali*. <https://doi.org/10.1016/j.jebo.2023.10.016>
- Hillebrecht, M., Klonner, S., & Pacere, N. A. 2023. The dynamics of poverty targeting. *Journal of Development Economics*, 161, 103033. <https://doi.org/10.1016/j.jdeveco.2022.103033>
- Integrated Food Security Phase Classification. 2025. IPC Acute Food Insecurity Analysis October 2024–March 2025: Lebanon. https://www.ipcinfo.org/fileadmin/user_upload/ipcinfo/docs/IPC_Lebanon_Acute_Food_Insecurity_Oct2024_Mar2025_Report.pdf
- International Organization for Migration. 2024. Mobility Snapshot—Round 67—12-12-2024. Global Data Institute Displacement Matrix. <https://dtm.iom.int/datasets/mobility-snapshot-round-67-12-12-2024>
- Jeong, D., & Trako, I. 2022. *Cash and In-Kind Transfers in Humanitarian Settings: A Review of Evidence and Knowledge Gaps* (No. 10026; Policy Research Working Paper). World Bank Group. <https://openknowledge.worldbank.org/entities/publication/2152a365-c73f-5bfd-9070-6f1fa3a81f03>
- Kagin, J. et al. 2024. *The Cost of Inaction: Impacts of WFP Assistance Shortfalls on Food Security Outcomes in Somalia*. World Food Programme. <https://hdl.handle.net/10919/121145>
- Kagin, J. et al. 2024. *The Cost of Inaction: Impacts of WFP Refugee Assistance Shortfalls on Food Security Outcomes in Uganda*. World Food Programme. <https://hdl.handle.net/10919/121144>
- Khan, S., & Knotek, E. S. 2011. How do households respond to uncertainty shocks? *Economic Review*, 96(Q II). <https://ideas.repec.org//a/fip/fedker/y2011iqiinv.96no.2.html>
- Kondylis, F., & Loeser, J. 2024. *Cash transfer size and persistence*. <http://johnloeser.github.io/assets/kl.pdf>
- Manea, R. E. et al. 2025. *Impact Evaluation of Farmer Support Activity in Ghana*. World Food Programme, Rome. <https://www.wfp.org/publications/ghana-farmer-support-activity-impact-evaluation>
- 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.

- 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>
- Raleigh, C. et al. 2010. Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research*, 47(5), 651–660. <https://doi.org/10.1177/0022343310378914>
- Schnitzer, P., & Stoeffler, Q. 2024. Targeting Social Safety Nets: Evidence from Nine Programmes in the Sahel. *The Journal of Development Studies*, 60(4), 574–595. <https://doi.org/10.1080/00220388.2023.2291325>
- Shapiro, J. M. 2005. Is there a daily discount rate? Evidence from the food stamp nutrition cycle. *Journal of Public Economics*, 89(2), 303–325. <https://doi.org/10.1016/j.jpubeco.2004.05.003>
- Sterck, O., & Bruni, V. 2025. The Welfare and General-Equilibrium Impacts of Aid Cuts (SSRN Scholarly Paper No. 5270846). *Social Science Research Network*. <https://doi.org/10.2139/ssrn.5270846>
- Stoeffler, Q., Mills, B., & Del Ninno, C. 2016. Reaching the Poor: Cash Transfer Programme Targeting in Cameroon. *World Development*, 83, 244–263. <https://doi.org/10.1016/j.worlddev.2016.01.012>
- Tumen, S. 2016. The Economic Impact of Syrian Refugees on Host Countries: Quasi-experimental Evidence from Turkey. *American Economic Review*, 106(5), 456–460. <https://doi.org/10.1257/aer.p20161065>
- UNOCHA. 2024. Global Humanitarian Overview 2024: October Update. <https://www.unocha.org/attachments/b9488007-b44e-4a35-9827-6a9f8488194e/10%20GHO%202024%20October%20Update.pdf>
- Vaitla, B., Coates, J., & Maxwell, D. 2015. *Comparing Household Food Consumption Indicators to Inform Acute Food Insecurity Phase Classification*. FHI 360/Food and Nutrition Technical Assistance III Project (FANTA).
- World Bank. 2024. *Lebanon Poverty and Equity Assessment 2024: Weathering a Protracted Crisis* (English). World Bank Group. <http://documents.worldbank.org/curated/en/099052224104516741/P1766511325da10a71ab6b1ae97816dd20c>
- World Food Programme. 2021. *Lebanon Annual Country Report 2020*.
- World Food Programme. 2023a. *Evaluation of Nigeria WFP Country Strategic Plan 2019–2022* [Centralized Evaluation Report]. <https://docs.wfp.org/api/documents/WFP-0000146137/download/>
- World Food Programme. 2023b. *Impacts of the Cost of Inaction on WFP Food Assistance in Sudan (2021 & 2022)*. <https://reliefweb.int/report/sudan/impacts-cost-inaction-wfp-food-assistance-sudan-2021-2022-april-2023>
- World Food Programme. 2023c. *Lebanon Annual Country Report 2022*.

World Food Programme. 2023d. *The Cost of Inaction: The consequences of not being able to reach those most in need in Eastern DRC*. <https://reliefweb.int/report/sudan/impacts-cost-inaction-wfp-food-assistance-sudan-2021-2022-april-2023>

World Food Programme. 2025. *Summary of Evaluation Evidence: Prioritization (2021–2023)* [WFP Evidence Summary]. <https://docs.wfp.org/api/documents/WFP-0000165455/download/>

World Health Organization. 2024. The World Health Organization-Five Well-Being Index (WHO-5). <https://www.who.int/publications/m/item/WHO-UCN-MSD-MHE-2024.01>

9. Acronyms

ACLED	Armed Conflict Location & Event Data
CARI	Consolidated Approach for Reporting Indicators of Food Security
CASN	Central Arab Statistical Network
CH	Cadre Harmonisé
FCS	Food Consumption Score
GRF	Generalized Random Forests
HH	Household
HHH	Household Head
HHM	Household Member
IPC	Integrated Food Security Phase Classification
ISDC	International Security and Development Center
IV	Instrumental variable
LASSO	Least Absolute Shrinkage and Selection Operator
LBCO	Lebanon Country Office
LBP	Lebanese Pound
LCSI	Livelihood Coping Strategy Index
LHS	Lebanon Household Survey
LSMS	Living Standards Measurement Study
MDDI	Multidimensional Deprivation Index
MEB	Minimum Expenditure Basket
NGO	Non-governmental organization
ODK	Open Data Kit
OLS	Ordinary Least Squares
PDM	Post-Distribution Monitoring
PMT	Proxy-Means Test
PTSD	Post-Traumatic Stress Disorder
rCSI	Reduced Coping Strategies Index
RCT	randomized controlled trial
UNHCR	United Nations High Commissioner for Refugees
USD	United States DollarWFP World Food Programme
WHO	World Health Organization

10. Annexes

A1. Programme implementation additional details

Targeting threshold

179. To determine the number of households to continue receiving assistance in 2024, the LBCO established a targeting threshold of 43 percent. This threshold corresponds to 25,160 households and was derived using the MEB approach, based on data from the 2023 Lebanon Household Survey (LHS) with a sample size of 3,597 households. The MEB was set at LBP 3,238,315 per capita per month – or approximately USD 71 at January 2023 prices – representing the minimum level of monthly consumption required to meet basic needs.
180. However, the LHS dataset is not fully representative of the programme’s target population, particularly with respect to key indicators of vulnerability. Given that econometric models used to predict household vulnerability are sensitive to the quality and representativeness of the training data, some adjustments were necessary. To better approximate the programme’s target population – which is expected to consist of generally poorer and more food-insecure households – the LBCO applied a trimming procedure, excluding the top 25 percent of LHS observations based on per capita expenditure.¹⁰⁴ The analysis led to the identification of 43 percent as the targeting threshold.
181. In parallel, the LBCO routinely conducts Post-Distribution Monitoring (PDM) surveys to track the food security and consumption levels of assisted households. The most recent PDM survey was conducted in July 2023 (before the baseline census) on a representative sample of 1,000 households in the in-kind food assistance programme. The survey found that 44 percent had consumption levels below the MEB threshold, which is closely aligned with the adopted targeting threshold.
182. Based on this analysis, the LBCO set the final targeting threshold at 43 percent. The resulting 25,160 households serve as the target caseload across all targeting approaches tested in the study. Table A1.1 shows the geographical distribution of the households in the in-kind food assistance programme.

¹⁰⁴ The 25 percent threshold was chosen following simulations that tested the impact of alternative trimming levels (1 percent, 5 percent, 20 percent, 25 percent, and 50 percent) on targeting errors.

Table A10.1: Distribution of in-kind assistance households across governorates and districts

Governorate	District	Households	% of households from pool selected per governorate
Akkar	Akkar	5,540	78%
Baalbek-Hermel	Baalbek	2,317	62%
	El Hermel	399	78%
Beirut	Beirut	1,076	37%
Bekaa	Rachaya	148	28%
	West Bekaa	330	30%
	Zahlé	827	33%
Mount Lebanon	Aley	839	30%
	Baabda	1,745	37%
	Chouf	678	24%
	El Meten	912	29%
	Jbeil	215	21%
	Keserwane	515	21%
Nabatieh	Bint Jbeil	447	29%
	El Nabatieh	1,020	36%
	Hasbaya	240	37%
	Marjaayoun	428	38%
North Lebanon	Bcharreh	213	42%
	El Batroun	229	23%
	El Koura	532	36%
	El Minieh-Dennie	1,492	58%

	Tripoli	1,979	72%
	Zgharta	641	49%
South Lebanon	Jezzine	133	25%
	Saida	1,067	31%
	Sour	1,198	37%
Total in-kind programme		73,520	

Food parcel details

183. The parcels include ten items and are meant to provide at least 1,300 kcal per person per day. A detailed breakdown of the parcel's composition is provided in Table A1.2.

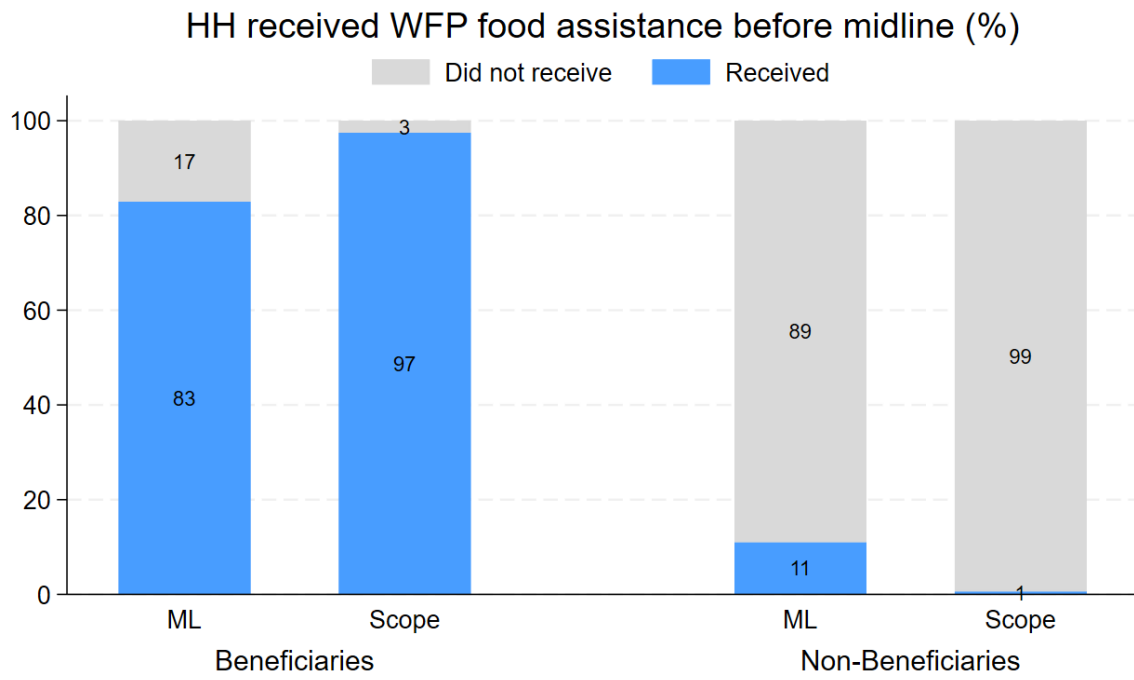
Table A10.2: Composition of the in-kind food parcel distributed by WFP in Lebanon

Commodity	Quantity	Food group	kcal contribution ^{1/}	Contribution to the basket
Egyptian rice	3 kg	Cereals	10,800	47%
Brown bulgur	3 kg	Cereals	10,260	
Pasta	4 kg	Cereals	14,840	
Lentils	4 kg	Pulses	13,720	27%
White beans	1 kg	Pulses	3,330	
Chickpeas	1 kg	Pulses	3,640	
Sunflower oil	1.84 l	Oils & Fat	15,000	20%
Salt	500 g	Condiments	0	0%
Tomato paste	400 g	Condiments	328	
Sugar	1 kg	Sugar	3,870	

Note: The contribution represents kcal equivalencies of the defined quantity of a basket for two people.

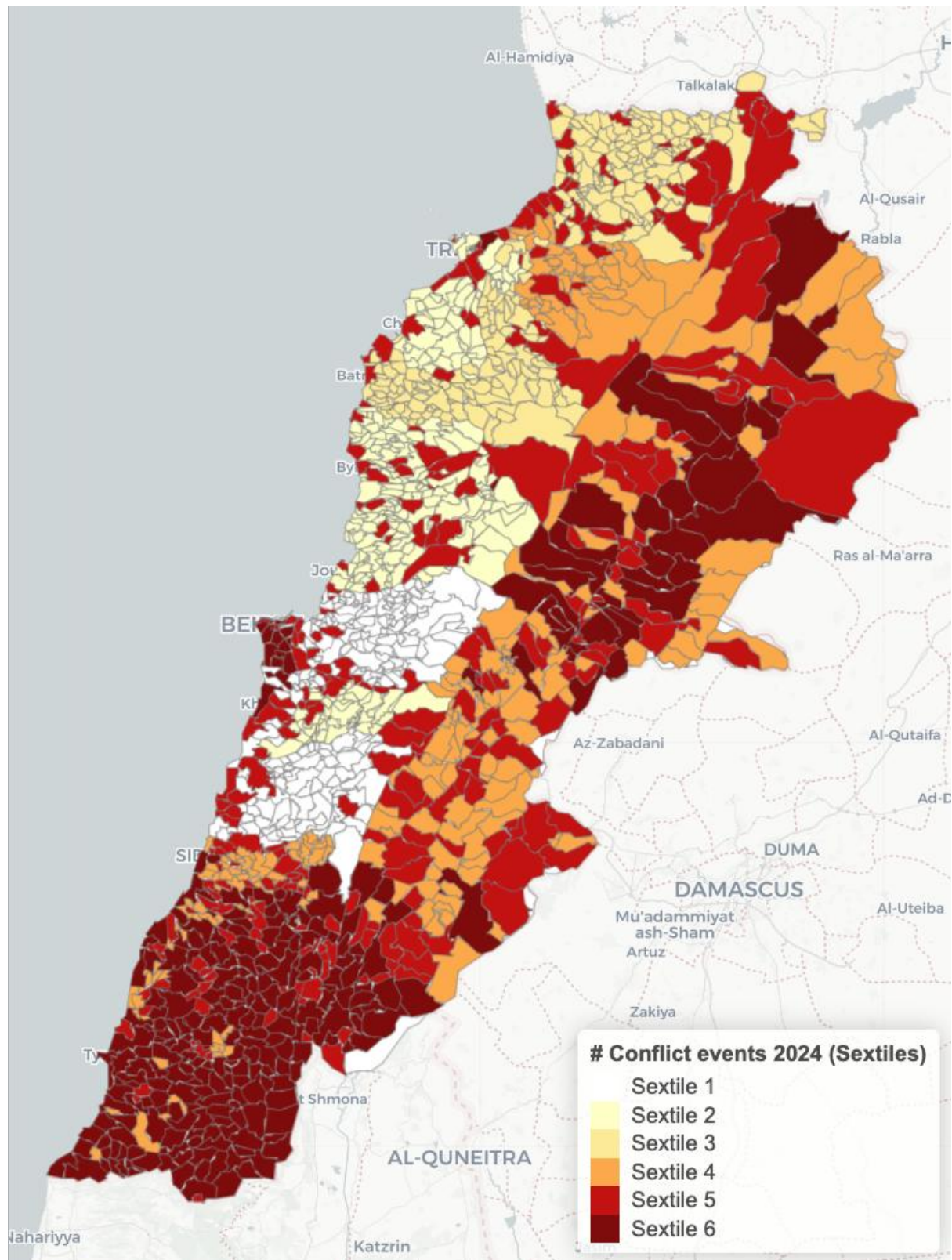
184. Each month, households receive a message with an individual code to collect their parcel from one of the many collection points disseminated across the country. Administrative data from the LBCO show that households regularly collect their parcels: in 2023, 89 percent of households collected their parcels every month and virtually everyone (99 percent) collected them at least for half of the months.

Figure A10.1: Adherence to implementation/study protocol



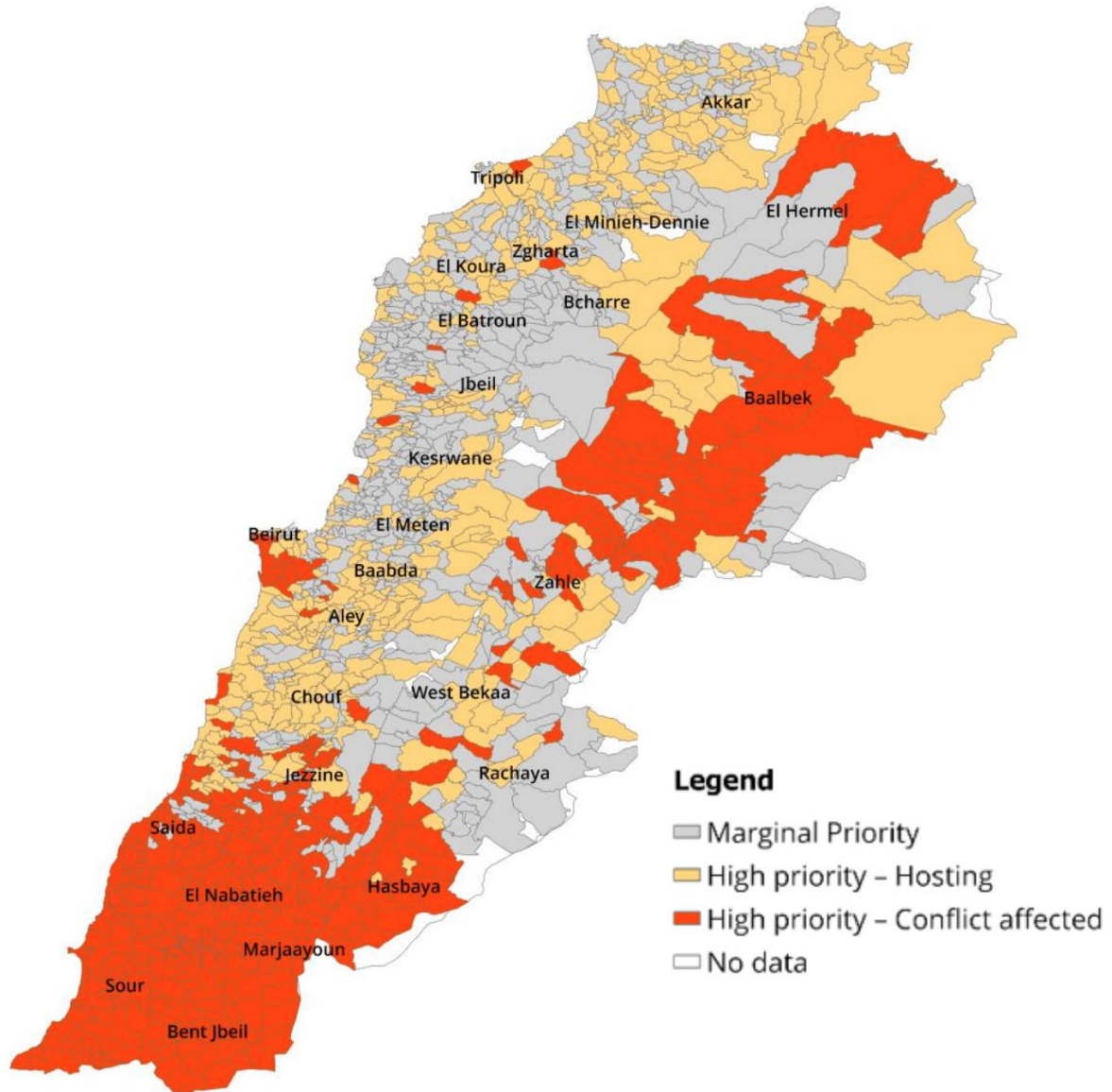
Note: The figure displays the validation of receipt of in-kind food assistance with WFP Lebanon's scope data (Scope) and round 1 of survey data collection (ML).

Figure A10.2: Conflict events registered July to December 2024 at sub-district level



Note: Own visualization. Data source: Armed Conflict Location & Event Data (ACLED)

Figure A10.3: Cadastre-level prioritization for cash distribution



Note: The figure is obtained from LBCO, and shows the three priorities for cash eligibility at the cadaster levels (marginal, high hosting, and high conflict affected).

A2. Targeting additional details

Proxy means test (PMT) models

185. PMT models are among the most widely used approaches for targeting development and humanitarian assistance in contexts where administrative capacity is limited and direct measures of need are not readily observable. In this study, three separate PMT models are developed to capture distinct dimensions of vulnerability. The first model uses total household expenditure as a proxy for poverty. The second relies on the Food Consumption Score (FCS), a commonly used indicator of food security based on dietary diversity and food frequency. The third model is based on the Consolidated Approach to Reporting Indicators of Food Security (CARI), which offers a broader, multidimensional perspective by combining information on food expenditure, food security, and adoption of harmful coping strategies to address food shortages.
186. The selected models and outcomes correspond to the specific proxies used to target the population of interest. In this case, the outcomes of interest are total expenditure, the FCS, and CARI. These indicators are not typically available for the entire population and collecting them at scale may neither be reliable¹⁰⁵ nor cost-efficient. Households may misreport income or expenditure data, either intentionally or due to recall difficulties. WFP does not recommend using self-reported outcome measures directly for targeting purposes.¹⁰⁶ To overcome this limitation, predictive modelling is used to estimate these outcomes based on observable characteristics. Indicators are collected through short interviews that are faster to conduct. They are also more and reduce the risk of misreporting as the questions asked are mostly based on observable variables such as demographics or assets. .
187. In practice, once the outcomes have been defined, they must be available in the selected training dataset. This dataset should also include a broad set of potential predictors that are similarly available in the data covering the population of interest. The choice of outcome and the selection of training data are therefore mutually dependent, as it is not possible to estimate a model for an outcome that is not observed in the training data.
188. Implementing PMT involves a structured sequence of steps that rely on two key data sources. The first is a training dataset that contains the outcome of interest, such as poverty or food insecurity, along with a broad set of potential predictors. The second is a population-level dataset or registry that includes observable characteristics for the population to which the model will be applied. The outcome selected must be available in the training dataset, and there must be strong overlap in the predictors across both datasets to allow for accurate prediction.

¹⁰⁵ See WFP. 2021. *Targeting and Prioritization: Operational Guidance Note*. <http://docs.wfp.org/api/documents/WFP-0000122035/download/>

¹⁰⁶ For instance, collecting detailed household consumption or expenditure data can take between 1.5 and 3 hours per household on average, depending on the length of the recall period and the number of items included in the questionnaire. This makes such data collection costly and time-consuming at scale. See: Deaton, A. & Zaidi, S. 2002. *Guidelines for Constructing Consumption Aggregates for Welfare Analysis*. LSMS Working Paper No. 135. World Bank. <http://hdl.handle.net/10986/14101>.

189. All three PMT models in this study follow the same core steps. The first step is to define the outcome variable, which reflects the specific dimension of vulnerability targeted by the intervention. The second step involves selecting an appropriate training dataset and identifying the set of common predictors. These predictors are used to estimate a range of candidate models. The final model is selected based on its performance using a validation or out-of-sample dataset, with careful testing and calibration to improve predictive accuracy. The third step consists of applying the model to the population dataset to generate predicted scores. Households are ranked according to their scores, and the eligible population is selected based on their position in the distribution.

190. Below is more detailed information on each of these three steps:

Step 1: defining analytical framework (targeting outcome)

191. Vulnerabilities in humanitarian contexts are inherently multidimensional. However, each agency tends to focus on a particular dimension that aligns with its mandate. For example, the World Bank primarily targets poverty reduction, the World Food Programme (WFP) focuses on food insecurity, and the United Nations High Commissioner for Refugees (UNHCR) prioritizes protection. While there is some overlap among these objectives, there is limited evidence on how targeting different outcomes may affect the overall effectiveness of a programme. This creates a knowledge gap in understanding whether collaboration across agencies or alignment with government programmes that pursue different targeting objectives can help achieve broader development or humanitarian goals.

Table A10.3: Indicators tested for PMT models

Indicator	Type	Econometric	Threshold for decision	Target population	Nature
Total expenditure per capita	Continuous	OLS		Energy deficient population	Poverty
Food expenditure per capita	Continuous	OLS		Energy deficient population	Poverty
Poverty	Binary	Logit	2,100 kcal per person per month	Energy deficient population	Poverty
Multidimensional Deprivation Index (MDDI)	Composite index	OLS	Deprivations as in MPI methodology	Chronically poor/deprived	Poverty
Food Consumption Score (FCS)	Continuous	OLS		Food insecure	Food insecurity
Food insecurity	Binary	Logit	Thresholds depending on country's sugar consumption	Food insecure	Food insecurity
CARI	Composite index	Poisson	CARI thresholds	Food insecure and vulnerable	Food Insecurity and poverty

192. In Lebanon, in-kind assistance is integrated into the national social protection system, which employs targeting to identify households in need. Reflecting this, several alternative

outcome variables were tested to guide targeting and assess their suitability. These included food expenditure, total household expenditure, and the MDDI developed by WFP. The FCS was also considered, along with different specifications of CARI. For CARI, both a binary indicator of food insecurity and a Poisson specification using the four-category food security classification were explored. In the case of the FCS, models were tested using both a binary indicator and the continuous score as the outcome variable. This range of outcomes captures the multidimensional nature of vulnerability and enables a comparison of how alternative proxies perform in identifying households most in need of assistance. Table A2.1 presents the indicators tested for prediction.

Step 2: training data and model validation

193. The LHS is the most comprehensive dataset available in Lebanon that includes all outcomes of interest for this study. It was selected as the training dataset to identify potential outcome variables and relevant covariates for modelling and predicting vulnerability in the target population. It is a collaborative effort led by WFP, the World Bank, and UNHCR, and includes data on residents and non-residents. This combination of breadth, depth, and relevance makes the LHS a suitable and reliable source for training the targeting models used in this evaluation. The LHS was collected in 2023 and includes data from 3,597 households. To enhance the relevance of the training data for this evaluation, the sample was trimmed to include only the 75 percent of households with the lowest total expenditure. This restriction ensures greater alignment between the training data and the characteristics of the population most likely to be selected for assistance. It also ensures that the training sample more accurately represents the poorest segments of the Lebanese population. Once the training dataset is identified, the outcome indicators are calculated according to the procedures described below.

194. Expenditure:

- The total expenditure aggregate includes food and non-food components. Food expenditure is collected using a seven-day recall period and then multiplied by four to estimate monthly values. Non-food expenditure is reported over two different recall periods: 30 days for regular expenses such as electricity and transport; and six months for less frequent expenditures such as education or asset purchases. The total also accounts for meals consumed outside the home. Food expenditure is derived from detailed consumption data. Households are asked whether they consumed various food groups, subgroups, or specific items in the past seven days. For each item consumed, households report the quantity using locally relevant units of measurement, followed by the amount paid for that quantity. Both food and non-food expenditure aggregates are then converted to a monthly value. For food expenditure, this is done by multiplying the seven-day recall amount by four. For non-food items, monthly expenditure is derived from two recall periods: 30 days for regular expenses; and six months for infrequent expenses, with the latter divided by six to obtain a monthly estimate. To adjust for household composition, total expenditure is converted to per capita terms using adult equivalence scales. Under this approach, the first adult in the household is assigned a weight of 1.0, the second adult and each additional

household member aged 14 or older receive a weight of 0.5, and each child under 14 receives a weight of 0.3.

- The final outcome used for modelling is monthly per capita expenditure adjusted by adult equivalences. This variable can be used in two ways: as a continuous measure or as a binary indicator based on a predefined threshold. One commonly used threshold is the World Bank poverty line, which corresponds to the estimated cost of meeting basic caloric needs (2,100 kilocalories per person per day) plus essential non-food items. In this case, WFP's MEB was used instead of the traditional poverty line. The MEB reflects the cost of a set of food and non-food goods and services considered essential by vulnerable households. The MEB is typically developed using expenditure data from vulnerable populations and is validated by national coordination platforms such as the food security cluster or cash working group. When used as a threshold, the expenditure variable becomes a binary indicator, suitable for modelling using logistic regression. When used as a continuous variable, standard linear regression is typically applied.
- For the continuous specification, the final per capita monthly expenditure variable is log-transformed to improve model performance. This transformation enhances prediction accuracy without affecting the selection process, as targeting is ultimately based on the relative ranking of predicted values rather than their absolute levels.

195. Food security:

- The FCS is used as an indicator of household food security. It is calculated based on the frequency with which different food groups are consumed over the seven days preceding the survey. Households report how many days in the past week they consumed items from each group, using a standardized list of aggregated food categories. The FCS includes nine food groups: staples, pulses, vegetables, fruits, animal protein, dairy, sugar, oils, and condiments. Each food group is assigned a weight that reflects its relative nutritional value. The final score is obtained by multiplying the number of days each group was consumed by its corresponding weight, and then summing across all groups. This results in a continuous variable ranging from 0 to a theoretical maximum of 112. Table A2.2 presents the food groups and the standard weights used in the FCS calculation, which follow internationally recognized guidelines.

Table A10.4: Construction of the Food Consumption Score (FCS)¹⁰⁷

Food group	Examples	Weight	Max. weighted score
Main staples	Cereals, tubers, roots	2	14
Pulses	Beans, lentils, peas	3	21
Vegetables	Spinach, tomatoes, carrots, etc.	1	7

¹⁰⁷ WFP. 2008. *Food Consumption Analysis*. Vulnerability and Mapping Unit, WFP Rome. World Food Programme. https://documents.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp197216.pdf

Fruits	Oranges, mangoes, bananas, etc.	1	7
Meat & fish	Beef, poultry, fish, eggs	4	28
Dairy	Milk, cheese, yogurt	4	28
Sugar	Sugar, honey	0.5	3.5
Oils & fats	Oil, butter	0.5	3.5

196. **CARI:**

- CARI is a composite indicator that combines three key dimensions of household vulnerability: the FCS; household expenditure; and livelihood-based coping strategies. Based on data collected through household surveys, CARI classifies households into four food security categories: Food Secure, Marginally Food Secure, Moderately Food Insecure, and Severely Food Insecure. This classification system enables the generation of representative food security estimates at different geographic levels, including national, district, regional, or livelihood zones. It also supports vulnerability profiling and targeting in the design and implementation of WFP programmes. CARI converts continuous indicators into a categorical scale ranging from 1 to 4, in line with the classification frameworks used by the Integrated Food Security Phase Classification (IPC) and the Cadre Harmonisé, which are widely applied in food security assessments.
- In this evaluation, the CARI indicator was constructed following the methodology described in Table A2.3. A binary version of the indicator was also created, identifying households in categories 3 and 4 as food insecure. This binary variable was used as an alternative specification for model estimation.

Table A10.5: Construction and classification of CARI¹⁰⁸

Domain	Indicator	CARI classification			
		Food Secure (1)	Marginally Food Secure (2)	Moderately Food Insecure (3)	Severely Food Insecure (4)
Food Consumption	Food Consumption Score (FCS)	Acceptable (FCS>42)		Borderline (42<FCS>28)	Poor (FCS<28)
Economic Vulnerability	Per capita expenditure	Above the Minimum Expenditure Basket (MEB) - includes food and non-food expenditure.		Below MEB but above Survival MEB (food only)	Below SMEB - minimum food expenditure

¹⁰⁸ WFP. 2015. *Consolidated Approach to Reporting Indicators of Food Security (CARI)* (2nd ed.). Vulnerability and Mapping Assessments Team. <https://www.indikit.net/document/421-consolidated-approach-to-reporting-indicators-of-food-security-cari>

Coping capacity	Coping Strategies Index related to livelihoods	No coping	Stress coping (e.g. use savings, borrow money)	Crises coping (e.g. take the CH out of school)	Emergency coping (e.g. migrate, child labour)
-----------------	--	-----------	--	---	--

197. Once the outcome variables were defined, the next step involved identifying the set of covariates that are common to both the training dataset (LHS) and the dataset covering the population of interest. This alignment ensures that models can be trained on one dataset and consistently applied to another using the same set of predictors. To facilitate the use of both traditional econometric models and non-traditional machine learning approaches, these covariates were transformed into a machine-readable format. This process included converting categorical variables into binary indicators, standardizing continuous variables, generating interaction and polynomial terms, and addressing missing values. The final set of potential predictors comprised 293 variables, all available in both datasets. These included the original variables as well as their machine-readable transformations.
198. To determine the most appropriate modelling approach for each outcome, several functional forms were tested, including Ordinary Least Squares (OLS), logistic regression, and Poisson models, depending on the nature of the dependent variable. Various methods for covariate selection and coefficient estimation were also explored. The first was stepwise regression, where covariates are introduced sequentially based on improvements in model fit. While widely used, stepwise methods have been criticized for their reliance on researcher discretion and their susceptibility to overfitting.
199. To address these concerns, automated regularization techniques were considered. These included Ridge regression, Least Absolute Shrinkage and Selection Operator (LASSO), and Elastic Net. LASSO, which penalizes less informative coefficients and shrinks them towards zero, provided the best performance across all outcomes based on out-of-sample mean squared error following Altındağ et al.¹⁰⁹ A the third non-parametric and more advance approach using a Generalized Random Forest (GRF) was also tested, as implemented in Haushofer et al,¹¹⁰ a non-parametric machine learning technique that captures complex interactions and non-linearities in the data. Table A2.4 presents the out-of-sample precision of the selected models.

¹⁰⁹ Altındağ, O. et al. 2021. Targeting humanitarian aid using administrative data: Model design and validation. *Journal of Development Economics*, 150, 102564. <https://doi.org/10.1016/j.jdeveco.2020.102564>

¹¹⁰ Haushofer, J. et al. 2025. Targeting Impact versus Deprivation. *American Economic Review*, 115(6), 1936-1974. <https://doi.org/10.1257/aer.20221650>

Table A10.6: Performance of the PMT models

Outcome	Model	Performance indicator	OLS (SW)		LASSO		GRF	
			Training	# covs selected	Training	# covs selected	Training	# covs selected
Consumption total (per capita)	Linear	MSE	0.46	51.00	0.21	93.00	0.23	0.23
		R2 (OOS)	0.55		0.56		0.49	0.49
Food Insecure (CARI binary)	Logit	MSE	780.59	45.00	0.56	31.00	0.09	0.09
		D2 (OOS)	0.22		0.24			
FCS	Linear	MSE	17.28	47.00	282.85	95.00	308.84	308.84
		R2 (OOS)	0.27		0.13		0.23	0.23

200. Consistent with the findings of Altındağ et al., LASSO regression outperformed the other methods tested in this evaluation. It delivered the highest out-of-sample accuracy across all outcomes and was therefore selected as the preferred specification for the PMT models.

Step 3: model selection and application

201. The third step involves applying the selected model to the population of interest. This process begins by estimating and selecting the best-performing model using the training data, and then applying the resulting coefficients to the baseline census to generate predicted scores. For this step to be valid, the covariates used during model estimation must be available and coded in the same way in the population dataset. Once the outcome values are predicted, households are ranked according to their predicted scores, and households are selected based on their relative position in the distribution, using the cut-off defined by the LBCO, specifically the bottom 43 percent most vulnerable households for each method.

Categorical approach

202. The fourth targeting approach applied was categorical targeting. This method is commonly used in humanitarian contexts due to its simplicity, transparency, and ease of implementation. Unlike statistical models, it does not rely on complex estimation procedures, making it especially useful in settings where data on the target population is limited. The core idea is to define specific dimensions or categories that identify vulnerable households and target those who meet the selected criteria. The key design challenge lies in choosing the categories. These may vary based on the context and data availability. Categories can be used directly to determine eligibility or combined into a score to rank households by vulnerability.

203. In this study, a score-based approach was used. The selection of categories and their weights was guided by expert opinion. An online survey was conducted in November 2023 with 28 stakeholders who have operational or research experience in food security and humanitarian assistance in Lebanon. The survey included covariates closely aligned with those used in the PMT approaches. However, in contrast to the proxy means-test

approaches, all variables were transformed into dummy variables for human classification, resulting in a set of 55 variables. This comprehensive set encompasses all facets of the original covariate list, excluding highly correlated variables (e.g. owning a TV and owning a flatscreen TV). It also excludes transformed variables such as squared variables, interactions, and normalizations. The full list of variables is provided in Table A2.5. This variable list aligns with 33 inclusion criteria (“include [in humanitarian food assistance] if”) and 22 exclusion criteria.

Table A10.7: Categorical score weighting

Variable	Weight
Household owns a car, truck or bus	-27
Household has a dryer	-16
Household has a washing machine	-14
Exclude if the main household income is from a business with more than 5 workers	-14
Household has a microwave	-11
Household has air conditioner	-8
At least one child attends private school	-7
Household has a dishwasher	-7
Household has heating system	-6
Exclude if the main household income is from unskilled regular labour	-5
Household has satellite cable	-4
Head of household completed tertiary education	-3
Household has a separate freezer	-3
Household owns a smartphone	-2
Exclude if the main household income is from skilled regular labour	-2
Household has a flat screen TV	-1
Household owns a laptop or computer	-1
Household has WiFi	-1
Household uses a private grid	0
Household uses gas for cooking	0
Household has electric burner for cooking	0
Household has a water heating system	0
Head of household doesn't know how to read and write	1.2
More than 1.5 persons per room, e.g., 2 rooms for 3 people	1.2
Head of household only completed primary	1.8
At least one household member has a psychological disability	2.4
The dwelling is not an independent building	2.4
Household is not connected to the public grid, and uses sources such as battery	2.4
Household source of water is an unprotected borehole, well or spring, or rainwater collection	3
Household hosts at least one Syrian refugee	3.2
At least one child under 5yo	3.4
At least one child under 2yo	4.4

household head has a psychological disability	4.6
Household floor is soil	4.6
The main household income is from a small informal agricultural or fishing business (< 5 workers)	5
Household head has a sensorial disability	5.2
More than 6 household members	5.6
Household shares the toilet with neighbours	5.8
Household doesn't own the dwelling	5.8
At least one household member has a physical disability	6
Household does not use a built-in toilet	6
At least one elderly (65yo or more)	6.6
At least one school-age child not enrolled	7
The main household income is from occasional labour	7
At least one member of the household with a chronic disease, such as diabetes	7.4
At least one household member has a sensorial disability	8.4
Head of household suffers any disability	10.6
Household head is an elderly	15
Household headed by women	15
At least one household member with disability	15
No household member in working age (15–64yo)	15.2
Household head has a physical disability	15.8
Household head is a minor	17
Widowed/divorced/single household head (monoparental household)	23.4
The household does not have any income	27

Note: The full list of variables provided to the stakeholder for ranking and ordered by final weights.

204. Experts were asked to select and rank up to ten variables they considered important for identifying Lebanese households in need of food assistance. Their responses were used to generate a categorical score for each household, taking into account how often a given variable was selected and the ranking it received.¹¹¹ The survey was provided in English and Arabic and sent out via email to stakeholders in humanitarian affairs related to Lebanon. In total, 28 experts participated in the study: 71 percent of survey respondents were female; 25 percent were aged under 30 years; 43 percent were aged between 30 and 39 years; 61 percent of the respondents worked for a non-governmental organization (NGO); and 82 percent were based in Lebanon.

205. Based on the survey responses, a categorical score was constructed using a weighted ranking system. Each respondent's vote contributed points to a variable. If a variable secured the top spot, it received 10 points, followed by 9 points for the second rank, and so forth, with the tenth rank earning one point. The scores for each variable were then summed up. Notably, the variable ranked highest was "Include if the household does not have any income," a unanimous choice among all participants. There was a pronounced

¹¹¹ In practice, every time a variable secured the top spot, it received 10 points, followed by 9 points for the second rank, and so forth, with the tenth rank earning one point. The scores for each variable were summed up. Lastly, the scores for the exclusion criteria were converted to negative values.

bias towards inclusion criteria. Consequently, the range for inclusion criteria was adjusted to align with the range used for exclusion criteria (0 to 27 points). Moreover, the scores were converted for the exclusion criteria to take negative values. The weights of variables are provided in Table A2.5.

206. Following the assessment of characteristics in the baseline census, households were assigned a score. If households surpassed the 43 percent threshold, then they were selected to continue receiving assistance.

A3. Empirical strategy

207. This appendix provides a more detailed description of the specific analysis used to address each research question, under the original design, which was pre-registered in the American Economic Association's registry for randomized controlled trials (AEARCTR-0011831).

Q1: *What are the differences across the four methods in terms of who gets selected to continue receiving assistance?*

208. In order to answer this question, the analysis compares the profiles of households selected for assistance by each targeting approach. It begins with a simple comparison of means across households selected under each method. This is followed by a series of multivariate regressions each on 142 household characteristics to identify which traits are systematically prioritized by each approach. Joint significance is assessed using F-tests. Through this approach, the analysis reveals how different targeting methods implicitly favour distinct household profiles.

Q2: *Which targeting approach leads to higher acceptance (as proxied by the lowest number of formal appeals)?*

209. The analysis makes use of administrative data collected by WFP on appeals submitted by households from the discontinued population to compare outcomes across the different targeting methods. More specifically, the analysis estimates a regression equation of the form:

$$Y_{hi} = \beta_1 * Targ^2_h + \beta_2 * Targ^3_h + \beta_3 * Targ^4_h + \varepsilon_{hi} \quad (1)$$

210. Y_{he} denotes whether the household h belonging to targeting method i submitted a complaint. $Targ^i_h$ is an indicator variable for households assigned to targeting method i . Targeting method 1 (PMT model targeting per-capita expenditure) is considered the 'status quo' reference approach. ε_{he} is the error term. The analysis always utilizes robust standard errors to account for heteroskedasticity.

Q3: *Which targeting approach leads to overall higher well-being among the population?*

211. The analysis makes use of survey data to compare outcomes across households randomly assigned to the different targeting methods. More specifically, the analysis estimates a regression equation of the form:

$$Y_{he} = \beta_1 * Targ^2_{he} + \beta_2 * Targ^3_{he} + \beta_3 * Targ^4_{he} + \mathcal{Q} * X_{he} + \varepsilon_{he} \quad (2)$$

212. Where Y_{he} is an outcome of interest (e.g. household's FCS, expenditure, etc.) for household h belonging to eligibility space e . $Targ^i_h$ is an indicator variable for households assigned to targeting method i . In the specification above, as well as in the empirical analysis, targeting method 1 (PMT model targeting per-capita expenditure) is considered the 'status quo' reference approach. X_{he} indicates a vector of baseline control for 'long notice' in rounds 1 and 2. In round 2, the vector additionally includes a survey quality dummy and a dummy for the randomization of outcome modules across surveys, implemented to reduce the time required for data collection. Finally, ε_{he} is the error term.

213. The coefficients β_i capture the average difference in Y between households assigned to targeting method i and the reference targeting method. By comparing the coefficients, it is also possible to assess the relative performance of the targeting methods against each other.
214. The analysis always uses robust standard errors to account for heteroskedasticity. It also applies weights to account for the sampling design, which selects equal numbers of households from each eligibility space regardless of their actual population size. This adjustment corrects for the overrepresentation of smaller spaces and ensures that the estimates are representative of the entire population.

Q4: Which types of households benefit the most from each targeting approach?

215. To investigate the presence of heterogeneous effects across the population, equation (1) is augmented with the set of interactions between the targeting arm indicators $Targ^j_h$ and the baseline covariate of interest Cov_{he} . As shown in equation 2, the interaction terms $Targ^j_h * Cov_{he}$ are the coefficients of interest, capturing how the effect of each targeting method, relative to the reference method (PMT Exp), varies with the baseline covariate Cov_{he} . X_{he} indicates a vector of baseline control for 'long notice' in rounds 1 and 2. In round 2, the vector additionally includes a survey quality dummy and a dummy for the randomization of outcome modules across surveys, implemented to reduce the time required for data collection. ϵ_{he} is the error term.

$$Y_{he} = \beta_1 * Targ^2_{he} + \beta_2 * Targ^3_{he} + \beta_3 * Targ^4_{he} + \beta_4 * Targ^2_{he} * Cov_{he} + \beta_5 * Targ^3_{he} * Cov_{he} + \beta_6 * Targ^4_{he} * Cov_{he} + \theta Cov_{he} + \Omega * X_{he} + \epsilon_{he} \quad (2)$$

216. The study examines the following set of potential sources of heterogeneous effects: household head was female, household was a minor/elderly, household head had a low education level, household had a member with a disability, household had below-median wealth, household had a high dependency ratio, and discontinuation status.

Q5: What is the impact of ending humanitarian food assistance on households' well-being?

217. In order to assess the consequences of ending humanitarian assistance, the analysis estimates the following empirical model:

$$Y_{he} = \alpha * Endofassistance_{he} + \Omega * X_{he} + \gamma_e + \epsilon_{he} \quad (3)$$

218. Where $Endofassistance_{he}$ is an indicator variable for households assigned to end assistance. The coefficient α thus captures the average impact of ending assistance on the outcome Y . Importantly, the inclusion of eligibility space fixed effects ensures that the estimation exploits variation in assistance status *within* eligibility spaces. This also implies that the estimated effects holds for households in the sample that belong to eligibility spaces where there is variation in assistance status – that is, all spaces except for E0 and E1E2E3E4. In equation 3, γ_e represents eligibility space fixed effects, as the randomization was stratified by eligibility. X_{he} indicates a vector of baseline controls selected using post-double selection Lasso, and always including eligibility space fixed effects, a control for long notice households, a survey quality dummy, and a dummy for the randomization of outcome modules across surveys, implemented to reduce the time required for data collection. ϵ_{he} is the error term.

Q6: Which types of households are most affected by the discontinuation of assistance?

219. Mirroring the approach described earlier for Q3, the analysis tests for the presence of heterogeneous effects by augmenting equation (3) with the baseline covariate of interest Cov_{he} and its interaction with $Endofassistance_{he}$. Also, in this case, the analysis considers the pre-specified potential sources of heterogeneous effect examined in Q2, including whether the household head was female, household was a minor/elderly, household head had a low education level, household had a member with a disability, household had below-median wealth, and whether a household had a high dependency ratio.

$$Y_{he} = \alpha * Endofassistance_{he} + \Omega * X_{he} + \eta * X_{he} * Cov_{he} + \theta Cov_{he} + \gamma_e + \varepsilon_{he} \quad (4)$$

220. Where $Endofassistance_{he}$ is an indicator variable for households assigned to end assistance. The coefficient α thus captures the average impact of ending assistance on the outcome Y . Cov_{he} represents the baseline covariate of interest. $X_{he} * Cov_{he}$ is the coefficient of interest, capturing how the effect of each discontinuation varies with the baseline covariate Cov_{he} . γ_e represents eligibility space fixed effects, as the randomization was stratified by eligibility. X_{he} indicates a vector of baseline controls selected using post-double selection Lasso, and always including eligibility space fixed effects, a control for long notice households, a survey quality dummy, and a dummy for the randomization of outcome modules across surveys, implemented to reduce the time required for data collection. ε_{he} is the error term.

Q7: Do households prepare for the end of assistance? Does a longer notice period help households attenuate any negative consequence?

221. The analysis investigates the behaviour of households that received the long notice by estimating the following equation:

$$Y_{he} = \Phi_1 * longnotice_{he} + \Omega * X_i + \gamma_j + \varepsilon_{ij} \quad (5)$$

222. Where $longnotice_{he}$ is an indicator variable for households randomly assigned to the long notice, γ_e represents eligibility space fixed effects, as the randomization was stratified by eligibility. X_{he} indicates a vector of baseline controls selected using post-double selection Lasso, and always including eligibility space fixed effects, a control for long notice households, a survey quality dummy, and a dummy for the randomization of outcome modules across surveys, implemented to reduce the time required for data collection. ε_{he} is the error term.

223. The long notice sample can be compared to two alternative groups of households: the continued assistance households; and the discontinued (short-notice) households. In order to understand whether long notice households are taking any preventive action to prepare for the expected end of assistance, the cleanest comparison is the one against the continued assistance households in survey round 1. By that time, both continued assistance and long notice households were still receiving food assistance and, thus, any difference in their outcomes would have been driven by the fact that long notice households were expecting the end of their support soon (the inclusion of eligibility space fixed effects once again ensures that the two sets of continued assistance and long notice households would be perfectly comparable). By survey round 2, even long notice households stopped receiving assistance. In this case, the comparison between long- and

short- notice households would reveal the combined (potential) benefit of providing additional six months of food assistance and a longer notice.

A4. Randomization, sampling, balance, and attrition

Randomization of households into targeting methods

Table A4.1 shows the relatively equal allocation of the study population into the four targeting arms.

Table A10.1: Allocation of households across eligibility spaces and targeting methods

Eligibility space	Total (%)	Random allocation to four targeting methods			
		Applied PMT-Exp	Applied PMT-FCS	Applied PMT-CARI	Applied Categorical
Total	58,501	14 624	14 626	14 625	14 626
Selected for assistance	25 160 (43%)	6 289	6 289	6 290	6 292
Not selected	33 341 (57%)	8 335	8 337	8 335	8 334
E0	15 295 (26.14%)	3 824	3 824	3 824	3 823
E1E2E3E4	7 727 (13.21%)	1 931	1 932	1 932	1 932
E1	2 724 (4.66%)	681	681	681	681
E1E2	706 (1.21%)	176	177	176	177
E1E2E3	7 753 (13.25%)	1 939	1 938	1 938	1 938
E1E2E4	458 (0.78%)	114	115	114	115
E1E3	2 825 (4.83%)	707	706	707	705
E1E3E4	1 810 (3.09%)	453	452	452	453
E1E4	1 153 (1.97%)	288	289	287	289
E2	1 945 (3.32%)	485	486	487	487
E2E3	1 390 (2.38%)	347	347	348	348
E2E3E4	2 332 (3.99%)	583	583	583	583
E2E4	2 845 (4.86%)	712	711	711	711
E3	703 (1.20%)	176	176	176	175
E3E4	616 (1.05%)	154	154	154	154
E4	8 219 (14.05%)	2 054	2 055	2 055	2 055

Note: Number of households eligible under combinations of targeting methods and random allocation to method T (horizontal shares in parentheses). Values coloured in green refer to retargeted households, while values coloured in red refer to discontinued households.

Survey sampling details

224. The main analysis of the evaluation draws on two rounds of household survey data. The questionnaire administered during both survey rounds was developed by the evaluation team and covered a broad range of dimensions related to household characteristics and well-being. The first round of data collection from the study sample took place between May to June 2024, just before the end of assistance for the long notice households. The second (and last) round of data collection took place between January to February 2025.
225. For the survey, households were stratified by assigned targeting method and eligibility space. This resulted in 64 groups: 4 assigned targeting methods x 16 eligibility spaces. Under each one of the 64 groups, 100 households were then randomly selected, leading to a total sample of 6,400 households.
226. This sampling strategy results in a sample that is not only equally divided across the four targeting arms (1,600 households per arm) but also across households selected and excluded from assistance (3,200 households per group). The set of excluded households also includes 960 'long notice' households and 2,240 discontinued ('short notice') households.¹¹²

Survey roll-out and attrition

227. The planned sample composition – as well as the actual roll-out of household data collection for both survey rounds – can be found in Table A4.2. In cases where a household could not be surveyed, it was replaced with another household from the same eligibility-targeting stratum.
228. Among the households assigned to continued assistance and short-notice groups, 99 percent of the target sample was successfully surveyed during round 1 of household interviews and 94 percent during round 2, with no evidence of differential attrition across targeting arms or discontinuation status (see Table A4.3).
229. For the long notice households, however, only 31 percent of the target sample was surveyed during round 1 due to an error in the list of households used by the survey firm. During round 2, this issue was resolved, and 93 percent of the original long notice sample was successfully surveyed. This means that the round 1 analysis of long notice households will have lower statistical power, which should be taken into account when interpreting results.

¹¹² By giving equal weight to each stratum, the sample maximizes the power for the comparison between continued and discontinued assistance. While this does not mirror the population distribution across eligibility spaces, by applying population weights, it is possible to retrieve population-wide estimates.

Table A1.2: Planned sample composition and roll-out by eligibility space at rounds 1 and 2

Eligibility Cells	Targeting Method 1			Targeting Method 2			Targeting Method 3			Targeting Method 4			Continued Assistance			Discontinued Assistance		
	Plan	R1	R2	Plan	R1	R2	Plan	R1	R2	Plan	R1	R2	Plan	R1	R2	Plan	R1	R2
e0	100	98	101	100	97	101	100	97	100	100	96	100	0	0	0	400	388	402
e1	100	100	101	100	99	101	100	101	100	100	102	103	100	100	101	300	302	304
e1e2	100	101	101	100	97	101	100	98	102	100	97	102	200	198	202	200	195	204
e1e2e3	100	100	101	100	100	101	100	99	101	100	101	100	300	299	303	100	101	100
e1e2e3e4	100	100	102	100	100	102	100	101	102	100	103	103	400	404	409	0	0	0
e1e2e4	100	100	101	100	97	97	100	101	96	100	99	100	300	296	298	100	101	96
e1e3	100	101	101	100	98	101	100	96	100	100	100	102	200	197	201	200	198	203
e1e3e4	100	100	101	100	98	101	100	98	100	100	100	101	300	298	302	100	98	101
e1e4	100	98	101	100	99	101	100	98	100	100	100	100	200	198	201	200	197	201
e2	100	101	104	100	103	102	100	100	100	100	102	101	300	103	102	100	303	305
e2e3	100	100	102	100	99	101	100	102	103	100	100	101	200	201	204	200	200	203
e2e3e4	100	102	105	100	100	101	100	100	100	100	103	105	100	303	306	300	102	105
e2e4	100	99	101	100	104	102	100	99	102	100	103	100	200	207	202	200	198	203
e3	100	99	101	100	99	101	100	98	100	100	100	100	300	98	100	100	298	302
e3e4	100	97	101	100	94	101	100	98	101	100	103	100	200	201	201	200	191	202
e4	100	95	102	100	95	100	100	98	100	100	98	100	300	98	100	100	288	302
Total	1 600	1 591	1 626	1 600	1 579	1 614	1 600	1 584	1 607	1 600	1 607	1 618	3 600	3 201	3 232	2 800	3 160	3 233

Note: The table displays the planned number of households for the study sample, the ones reached in round 1 and the ones reached in round 2 disaggregated by targeting approach and eligibility.

Table A1.3: Survey sample and attrition

Treatment	Planned sample	Response at round 1 (%)	Response at round 2 (%)
Panel A: Targeting			
PMT-Expenditure	1 600	89.3	93.5
PMT-FCS	1 600	88.8	94.0
PMT-CARI	1 600	89.9	93.5
Categorical	1 600	89.0	94.1
F-test for balance across all groups		0.415	0.297
Panel B: Discontinuation			
Continuation	3 200	79.2	93.7
Discontinuation	3 200	99.3	93.8
F-test for balance across all groups		634.126***	0.008
Panel C: Long Notice vs Short Notice vs Households with Continued Assistance			
Long Notice	946	31.0	93.0
Short Window	2 254	99.5	94.0
Continued Households	3 200	99.3	93.8
F-test for balance across all groups		1149.701***	0.414
Panel D: Long notice across targeting methods			
PMT-Expenditure	245	27.3	87.8
PMT-FCS	245	28.6	92.7
PMT-CARI	245	34.3	86.5
Categorical	245	29.8	93.1
F-test for balance across all groups		1.767	0.2557

Sampling balance

230. Table A4.4 compares baseline characteristics from the census and the study sample, and it shows minor differences between the two that reflect the sampling strategy illustrated above.¹¹³

Table A1.4: Balance in baseline characteristics

	Population (N=58,501)	Endline Sample (N=6,461)
Household headed by a woman	0.30 (0.46)	0.33 (0.47)
Age of the household head	55.40 (13.92)	55.99 (13.93)
Household head divorced/separated or widowed	0.21 (0.41)	0.22 (0.42)
Household head is married	0.68 (0.47)	0.65 (0.48)
Household head never attended school	0.21 (0.41)	0.23 (0.42)
Household head with tertiary education only	0.07 (0.26)	0.05 (0.22)
Number of household members	3.86 (1.82)	3.81 (1.80)
Dependency ratio (non-productive/household members)	0.35 (0.30)	0.35 (0.30)
At least one household member has a disability	0.33 (0.47)	0.36 (0.48)
At least one child attends private school	0.17 (0.37)	0.15 (0.36)
Skilled regular job private	0.03 (0.16)	0.02 (0.15)
Business owner informal	0.13 (0.34)	0.12 (0.32)
Unemployed	0.31 (0.46)	0.33 (0.47)
No Syrians in household	0.05 (0.36)	0.06 (0.40)
Household does not own house	0.51 (0.50)	0.53 (0.50)
Dwelling area per household member	36.88 (28.09)	35.52 (26.29)
Flush toilet	0.86 (0.35)	0.85 (0.36)
Piped water access	0.22 (0.42)	0.24 (0.43)
Use gas heating	0.16 (0.36)	0.14 (0.35)
Use gas for cooking	0.73 (0.44)	0.73 (0.44)
Use electricity from own solar panel	0.09 (0.29)	0.09 (0.29)
WiFi internet (e.g. Cable / ADSL)	0.52 (0.50)	0.48 (0.50)
Computer / Laptop	0.08 (0.26)	0.04 (0.20)
Mobile internet	0.49 (0.50)	0.45 (0.50)
Flat screen TV	0.63 (0.48)	0.60 (0.49)

¹¹³ Given that the sample was randomly selected from each eligibility cell, and the composition of each cell is known, it is always possible to apply population weights to retrieve population-wide statistics from the study sample.

Vacuum cleaner	0.31 (0.46)	0.24 (0.43)
Microwave	0.16 (0.37)	0.10 (0.30)
Air conditioner	0.18 (0.38)	0.13 (0.34)
Scooter / Moped / Motorcycle	0.09 (0.29)	0.10 (0.29)
Car	0.37 (0.48)	0.30 (0.46)

Note: Mean values of baseline characteristics for baseline population and endline sample with standard deviations in parentheses.

231. Tables A4.5 to A4.7 compare these characteristics across targeting arms, discontinuation of assistance, and long notice status. The tables show that, as expected, the randomizations and study design led to observationally equivalent groups.¹¹⁴ Table A4.8 displays the randomization balance across all districts, confirming that there are no systematic differences across groups.

Table A1.5: Balance across targeting methods

	Coef of selection (robust SE)			
	PMT-Exp	PMT-FCS	PMT-CARI	Categorical
Household headed by a woman	0.001 (0.004)	0 (0.004)	0.002 (0.004)	-0.003 (0.004)
Age of the household head	-0.052 (0.134)	-0.085 (0.133)	0.163 (0.133)	-0.027 (0.133)
Household head divorced/separated or widowed	-0.003 (0.004)	0.003 (0.004)	0.004 (0.004)	-0.004 (0.004)
Household head is married	0 (0.004)	-0.001 (0.004)	0.001 (0.004)	0.001 (0.004)
Household head never attended school	0.004 (0.004)	0 (0.004)	-0.006 (0.004)	0.002 (0.004)
Household head with tertiary education	0.004 (0.002)	0 (0.002)	-0.001 (0.002)	-0.003 (0.002)
Number of household members	0.009 (0.017)	-0.003 (0.017)	0.01 (0.017)	-0.016 (0.017)
Dependency ratio (non-productive/household members)	-0.001 (0.003)	-0.001 (0.003)	0.002 (0.003)	-0.001 (0.003)
At least one household member has a disability	0.005 (0.004)	-0.008* (0.004)	0.001 (0.004)	0.002 (0.004)
At least one child attends private school	-0.006* (0.004)	0.001 (0.004)	0.005 (0.004)	0 (0.004)
Skilled regular job private	0 (0.002)	0.002 (0.002)	0.002 (0.002)	-0.004*** (0.001)
Business owner informal	-0.001 (0.003)	-0.004 (0.003)	0.001 (0.003)	0.003 (0.003)
Unemployed	-0.005 (0.004)	0.005 (0.004)	-0.003 (0.004)	0.003 (0.004)
No Syrians in household	-0.002 (0.003)	-0.001 (0.003)	0.006 (0.004)	-0.003 (0.003)
Household does not own house	-0.005 (0.005)	0.003 (0.005)	0.001 (0.005)	0.002 (0.005)
Dwelling area per household member	-0.168 (0.269)	0.163 (0.269)	-0.044 (0.267)	0.049 (0.268)
Flush toilet	-0.001 (0.003)	0.004 (0.003)	0.002 (0.003)	-0.005 (0.003)

¹¹⁴ The tables only include a subset of the variables available from the baseline census. When considering the entire set of variables, there are a total of 142 possible comparisons. Among these, randomization across the four targeting methods shows nine instances (6.3 percent) where an F-test indicates joint significance of the targeting methods on characteristics at the 10 percent level ($p < 0.1$), three instances (2.1 percent) at 5 percent level ($p\text{-value} < 0.05$), and zero instances at 1 percent level ($p\text{-value} < 0.01$). When comparing continued and discontinued assistance, there are 17 instances (12 percent) where the difference is significant at 10 percent level ($p\text{-value} < 0.1$), ten instances (7 percent) where it is significant at 5 percent level ($p\text{-value} < 0.05$), and one instance (0.7 percent) where it is significant at 1 percent level ($p\text{-value} < 0.01$). Finally, when comparing long notice and control, there are 15 instances (10.5 percent) where the difference is significant at the 10 percent level ($p\text{-value} < 0.1$), eight instances (5.6 percent) where it is significant at 5 percent level ($p\text{-value} < 0.05$), and two instances (1.4 percent) where it is significant at 1 percent level ($p\text{-value} < 0.01$).

Piped water access	0.002 (0.004)	0 (0.004)	0.002 (0.004)	-0.004 (0.004)
Use gas heating	-0.001 (0.003)	0.002 (0.003)	0.003 (0.003)	-0.004 (0.003)
Use gas for cooking	0.002 (0.004)	-0.002 (0.004)	-0.002 (0.004)	0.002 (0.004)
Use electricity from own solar panel	-0.004 (0.003)	0.006** (0.003)	-0.001 (0.003)	-0.001 (0.003)
WiFi internet (e.g. Cable / ADSL)	0.004 (0.005)	0.004 (0.005)	0.001 (0.005)	-0.008* (0.005)
Computer / Laptop	0 (0.003)	0 (0.003)	0.001 (0.003)	-0.001 (0.003)
Mobile internet	0.005 (0.005)	-0.003 (0.005)	-0.003 (0.005)	0.001 (0.005)
Flat screen TV	-0.007 (0.005)	0.002 (0.005)	0.001 (0.005)	0.004 (0.005)
Vacuum cleaner	0 (0.004)	0.004 (0.004)	-0.005 (0.004)	0.001 (0.004)
Microwave	0.002 (0.004)	0.005 (0.004)	-0.001 (0.004)	-0.005 (0.003)
Air conditioner	0 (0.004)	0.006 (0.004)	0.003 (0.004)	-0.008** (0.004)
Scooter / Moped / Motorcycle	0.001 (0.003)	0.006** (0.003)	-0.003 (0.003)	-0.004 (0.003)
Car	0 (0.005)	-0.003 (0.005)	0.001 (0.005)	0.003 (0.005)

*Note: Separate OLS regressions of study arm randomization on covariate with robust standard error in parentheses. N=58,501. Statistical significance is denoted by: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Table A1.6: Balance across assistance discontinuation status

	Mean (SD)		Coef of targeting (robust SE)	
	Discontinued	Continued	without elig FE	with elig FE
Household headed by a woman	0.26 (0.44)	0.35 (0.48)	0.083*** (0.004)	-0.003 (0.005)
Age of the household head	55.38 (13.36)	55.42 (14.62)	0.033 (0.118)	0.165 (0.139)
Monoparental household head	0.19 (0.39)	0.24 (0.42)	0.045*** (0.003)	-0.004 (0.004)
Household head is married	0.71 (0.46)	0.63 (0.48)	-0.072*** (0.004)	0.003 (0.005)
Household head never attended school	0.16 (0.36)	0.28 (0.45)	0.125*** (0.003)	0 (0.005)
Household head with tertiary education	0.10 (0.30)	0.03 (0.18)	-0.064*** (0.002)	0.002 (0.003)
Number of household members	3.66 (1.62)	4.12 (2.04)	0.459*** (0.016)	0.001 (0.017)
Dependency ratio (non-productive/ household members)	0.32 (0.30)	0.38 (0.30)	0.059*** (0.003)	0 (0.004)
At least one household member has a disability	0.25 (0.43)	0.43 (0.49)	0.172*** (0.004)	-0.013*** (0.005)
At least one child attends private school	0.22 (0.41)	0.10 (0.30)	-0.116*** (0.003)	-0.008** (0.004)
Skilled regular job private	0.04 (0.20)	0.01 (0.09)	-0.031*** (0.001)	-0.001 (0.001)
Business owner informal	0.15 (0.35)	0.11 (0.31)	-0.037*** (0.003)	0.002 (0.004)
Unemployed	0.24 (0.43)	0.39 (0.49)	0.141*** (0.004)	0.005 (0.004)
No Syrians in household	0.03 (0.29)	0.07 (0.44)	0.035*** (0.003)	0.005 (0.004)
Household does not own house	0.46 (0.50)	0.57 (0.49)	0.118*** (0.004)	-0.007 (0.006)
Dwelling area per household member	39.70 (28.65)	33.13 (26.87)	-6.576*** (0.231)	0.21 (0.312)
Flush toilet	0.93 (0.26)	0.76 (0.43)	-0.167*** (0.003)	0.004 (0.004)
Piped water access	0.21 (0.41)	0.24 (0.43)	0.033*** (0.004)	0.001 (0.005)
Use gas heating	0.21 (0.40)	0.09 (0.29)	-0.112*** (0.003)	-0.007* (0.004)
Use gas for cooking	0.78 (0.41)	0.67 (0.47)	-0.11*** (0.004)	-0.001 (0.005)
Use electricity from own solar panel	0.12 (0.32)	0.06 (0.25)	-0.052*** (0.002)	-0.004 (0.003)
WiFi internet (e.g. Cable / ADSL)	0.65 (0.48)	0.35 (0.48)	-0.297*** (0.004)	-0.004 (0.006)
Computer / Laptop	0.12 (0.32)	0.02 (0.15)	-0.093*** (0.002)	0.004* (0.002)
Mobile internet	0.53 (0.50)	0.43 (0.50)	-0.102*** (0.004)	0.002 (0.006)
Flat screen TV	0.74 (0.44)	0.49 (0.50)	-0.256*** (0.004)	0.001 (0.006)
Vacuum cleaner	0.43 (0.49)	0.14 (0.35)	-0.285*** (0.004)	0.005 (0.005)
Microwave	0.24 (0.43)	0.05 (0.23)	-0.191*** (0.003)	-0.001 (0.003)
Air conditioner	0.26 (0.44)	0.07 (0.26)	-0.185*** (0.003)	-0.005 (0.004)
Scooter / Moped / Motorcycle	0.10 (0.31)	0.08 (0.27)	-0.024*** (0.002)	0.002 (0.003)
Car	0.51 (0.50)	0.18 (0.39)	-0.327*** (0.004)	-0.006 (0.005)

Note: Separate OLS regressions of targeting selection on covariate with robust standard error in parentheses. N=58,501.

Statistical significance is denoted by: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A1.7: Balance across long notice and control

	Mean (SD)		Coef of long notice (robust SE)	
	Control	Long notice	without elig FE	with elig FE
Household headed by a woman	0.26 (0.44)	0.30 (0.46)	0.038*** (0.015)	-0.001 (0.014)
Age of the household head	55.37 (13.35)	55.98 (13.83)	0.611 (0.448)	0.378 (0.395)
Household head divorced/separated or widowed	0.19 (0.39)	0.22 (0.41)	0.028** (0.013)	0.005 (0.013)
Household head is married	0.71 (0.45)	0.67 (0.47)	-0.037** (0.015)	-0.001 (0.014)
Household head never attended school	0.16 (0.36)	0.23 (0.42)	0.075*** (0.014)	0.024* (0.014)
Household head with tertiary education	0.10 (0.30)	0.05 (0.23)	-0.046*** (0.007)	-0.003 (0.008)
Number of household members	3.66 (1.61)	3.81 (1.76)	0.157*** (0.057)	0.031 (0.048)
Dependency ratio (non-productive/household members)	0.32 (0.30)	0.34 (0.31)	0.012 (0.01)	-0.017* (0.01)
At least one household member has a disability	0.25 (0.43)	0.33 (0.47)	0.084*** (0.015)	0.003 (0.014)
At least one child attends private school	0.22 (0.41)	0.19 (0.39)	-0.031** (0.013)	0.015 (0.012)
Skilled regular job private	0.04 (0.20)	0.03 (0.16)	-0.015*** (0.005)	0.007 (0.005)
Business owner informal	0.15 (0.36)	0.13 (0.33)	-0.023** (0.011)	0 (0.011)
Unemployed	0.24 (0.43)	0.30 (0.46)	0.061*** (0.015)	-0.003 (0.013)
No Syrians in household	0.03 (0.29)	0.06 (0.36)	0.023** (0.011)	0.007 (0.012)
Household does not own house	0.45 (0.50)	0.52 (0.50)	0.066*** (0.016)	-0.006 (0.016)
Dwelling area per household member	39.81 (28.68)	36.09 (27.34)	-3.727*** (0.888)	-0.102 (0.805)
Flush toilet	0.93 (0.25)	0.87 (0.33)	-0.058*** (0.011)	-0.013 (0.011)
Piped water access	0.21 (0.41)	0.23 (0.42)	0.02 (0.014)	-0.006 (0.014)
Use gas heating	0.21 (0.41)	0.16 (0.37)	-0.049*** (0.012)	0.004 (0.012)
Use gas for cooking	0.78 (0.41)	0.74 (0.44)	-0.04*** (0.014)	-0.003 (0.015)
Use electricity from own solar panel	0.12 (0.32)	0.10 (0.30)	-0.015 (0.01)	0.011 (0.01)
WiFi internet (e.g. Cable / ADSL)	0.65 (0.48)	0.52 (0.50)	-0.135*** (0.016)	-0.005 (0.016)
Computer / Laptop	0.12 (0.32)	0.04 (0.19)	-0.081*** (0.006)	-0.001 (0.006)
Mobile internet	0.54 (0.50)	0.46 (0.50)	-0.075*** (0.016)	0.016 (0.016)
Flat screen TV	0.75 (0.44)	0.64 (0.48)	-0.11*** (0.016)	-0.003 (0.016)
Vacuum cleaner	0.43 (0.50)	0.26 (0.44)	-0.172*** (0.014)	0.004 (0.014)
Microwave	0.25 (0.43)	0.11 (0.31)	-0.143*** (0.01)	-0.003 (0.01)
Air conditioner	0.26 (0.44)	0.15 (0.35)	-0.115*** (0.012)	0.003 (0.012)
Scooter / Moped / Motorcycle	0.10 (0.31)	0.11 (0.31)	0.006 (0.01)	0.005 (0.01)
Car	0.51 (0.50)	0.32 (0.47)	-0.198*** (0.015)	-0.017 (0.015)

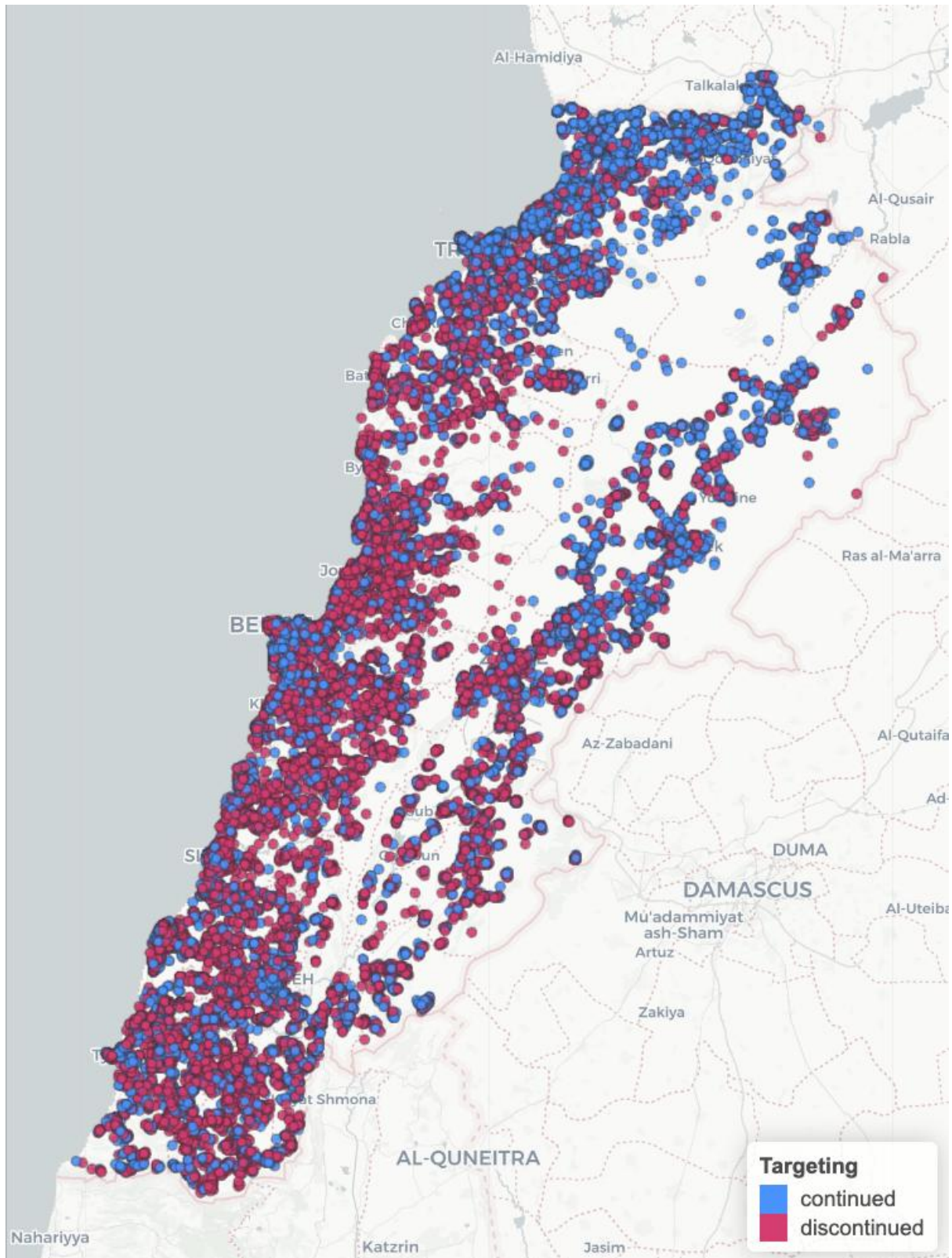
Note: Separate OLS regressions of selection into long notice versus control on covariate with robust standard error in parentheses. N=58,501. Statistical significance is denoted by: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A1.8: Balance in randomization across districts

Coefficient of selection (robust SE)				
District	Expenditure	FCS	CARI	Categorical
Akkar	0 (0.007)	0.002 (0.007)	-0.002 (0.007)	0 (0.007)
Aley	-0.001 (0.003)	-0.003 (0.002)	0.002 (0.003)	0.002 (0.003)
Baabda	0.003 (0.004)	0 (0.003)	-0.006 (0.004)	0.005 (0.005)
Baalbek	0.003 (0.005)	-0.003 (0.005)	-0.003 (0.004)	0.004 (0.006)
Bcharreh	0.001 (0.001)	0.002 (0.001)	-0.001 (0.001)	-0.001 (0.002)
Beirut	0.003 (0.003)	0 (0.003)	-0.003 (0.003)	0.001 (0.004)
Bint Jbeil	-0.001 (0.001)	0 (0.002)	-0.001 (0.002)	0.002 (0.003)
Chouf	-0.003 (0.002)	-0.003* (0.002)	0.005** (0.002)	0.002 (0.003)
El Batroun	0.001 (0.001)	-0.001 (0.001)	0.001 (0.001)	-0.002 (0.002)
El Hermel	-0.001 (0.002)	0 (0.002)	0 (0.002)	0 (0.003)
El Koura	0.003 (0.002)	-0.002 (0.002)	-0.001 (0.002)	0 (0.003)
El Meten	0.002 (0.003)	-0.001 (0.002)	0.001 (0.003)	-0.003 (0.004)
El Minieh-Dennie	-0.004 (0.004)	0.003 (0.004)	-0.001 (0.004)	0.002 (0.004)
El Nabatieh	-0.002 (0.003)	-0.004 (0.003)	0.004 (0.003)	0.003 (0.004)
Hasbaya	-0.001 (0.001)	0.003 (0.002)	0.001 (0.001)	-0.003 (0.002)
Jbeil	0.002 (0.001)	0 (0.001)	0 (0.001)	-0.002 (0.002)
Jezzine	0 (0.001)	0 (0.001)	0 (0.001)	-0.001 (0.001)
Keserwane	-0.002 (0.002)	-0.001 (0.001)	0.004** (0.002)	-0.002 (0.002)
Marjaayoun	-0.002 (0.002)	0.001 (0.002)	0.003 (0.002)	-0.003 (0.003)
Rachaya	0 (0.001)	0.001 (0.001)	-0.001 (0.001)	0 (0.001)
Saida	-0.005 (0.003)	0.007** (0.003)	0 (0.003)	-0.003 (0.003)
Sour	-0.001 (0.003)	-0.003 (0.003)	0.004 (0.003)	0.001 (0.004)
Tripoli	0.003 (0.005)	-0.001 (0.005)	-0.002 (0.005)	-0.001 (0.005)
West Bekaa	0.002 (0.002)	-0.004** (0.002)	0 (0.002)	0.003 (0.002)
Zahlé	-0.003 (0.003)	0.005* (0.003)	-0.002 (0.002)	-0.001 (0.003)
Zgharta	0.002 (0.003)	0.002 (0.003)	-0.003 (0.003)	-0.001 (0.003)

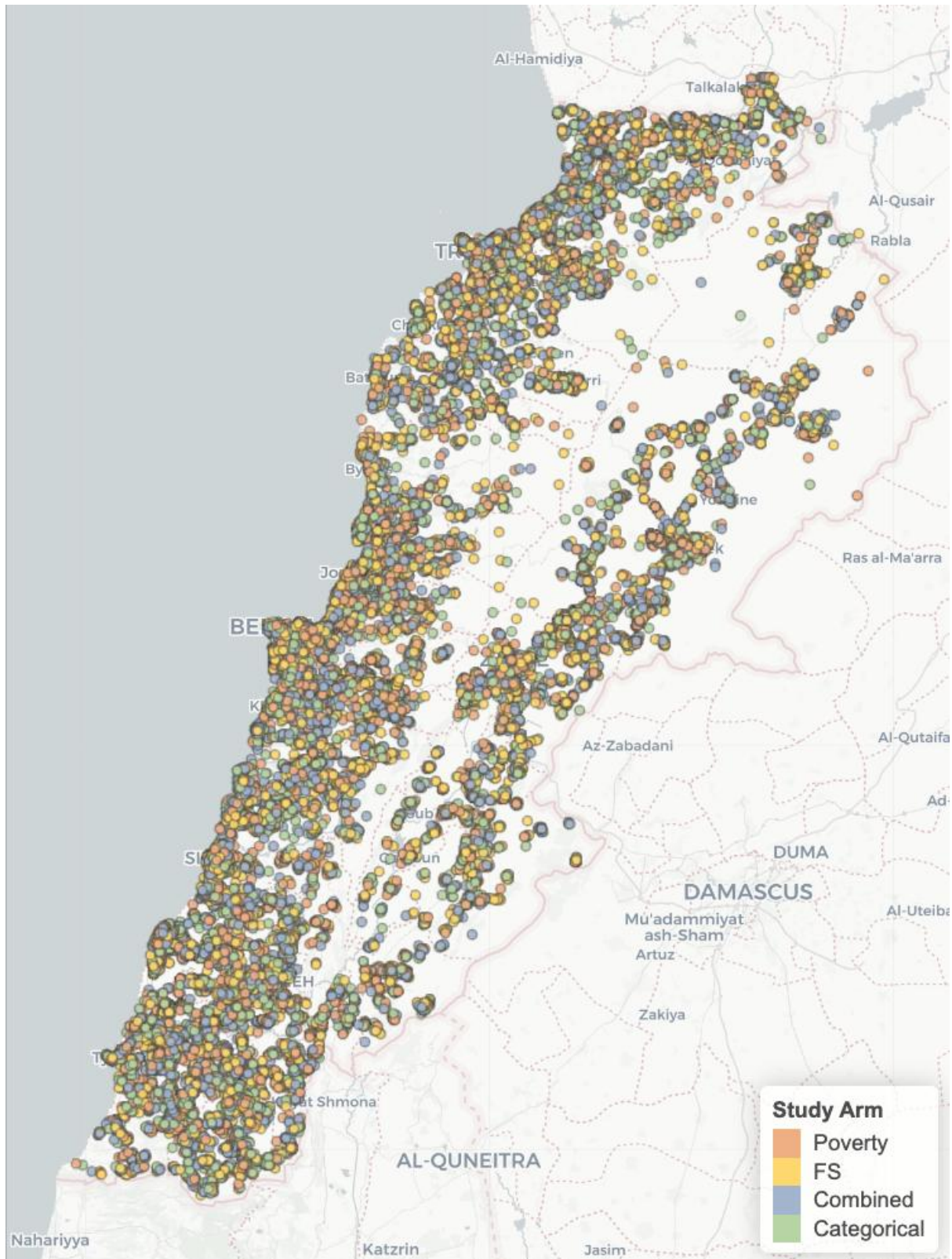
Note: Separate OLS regressions of study arm randomization on district with strata fixed effects and robust standard error in parentheses. N=58,501. Statistical significance is denoted by: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure A1.1: Geographical distribution of study sample households by selection status



Note: Own visualization. The figure shows the distribution of the study sample across Lebanon disaggregated by their assistance status after targeting.

Figure A1.2: Geographical distribution of households by study arm



Note: The figure shows the distribution of the households across Lebanon disaggregated by the targeting approach to which they were assigned to.

A5. Limitations

232. The study has some limitations:

233. **External validity of the study:** as in any empirical evaluation, the results hold for the specific population at hand. The fact that the sample was spread across the whole Lebanon and was representative of a large fraction of the country's population strengthens the external validity of the results. Nevertheless, more research on these topics shall be conducted across multiple contexts to enrich the evidence provided here.

234. **External validity of the targeting approach:** The four targeting methods examined in this study are based on existing approaches widely used in humanitarian and development settings. While their core logic is readily transferable across contexts, their specific implementation details are likely to vary. For example, the training data required for PMT-based methods will depend on the availability and quality of recent, representative household data. Although the training data is based on the 2022/2023 Lebanon Household Survey (LHS) -- which mirrors the World Bank's Living Standards Measurement Survey (LSMS) -- the dataset excludes Baalbek El-Hermel, El Nabatieh, and South Lebanon due to conflict issues. Likewise, the design and execution of expert-based surveys may differ depending on the availability and expertise of local specialists.

235. **External validity of the discontinuation analysis:** As outlined in the design section, the analysis of the consequences of assistance discontinuation focuses on 'marginal' households -- those that were eligible under at least one targeting method, but not all. This group represents over 60 percent of the target population. While the resulting estimates cannot be directly generalized to households consistently classified as vulnerable or non-vulnerable across all four targeting methods, the findings offer important insights also for these groups. Specifically, the analysis shows that both the likelihood of appeal and the severity of negative outcomes increased with the number of methods under which a household was deemed eligible. This indicates that the impact of discontinuation would likely be most severe for households identified as vulnerable by all four methods, and least severe for those excluded by all methods.¹¹⁵

236. **Limited power for the long notice analysis:** In round 1, only 31 percent of the intended sample of the 960 long notice households was successfully surveyed, due to an error in the household list provided to the survey firm. While this resulted in a substantially smaller sample, the surveyed subsample remained randomly selected and well balanced across treatment groups. By the endline, this issue had been resolved, and 93 percent of the original long notice sample was successfully surveyed. As a result, the analysis of long notice households at round 1 suffers from limited statistical power. This limitation should be considered when interpreting the findings. In the discussion of these results, greater

¹¹⁵ It is possible to further explore the consequences of the discontinuation of assistance for these two groups through a regression discontinuity analysis. However, this falls outside the scope of this report.

emphasis is placed on the direction and magnitude of the estimated, rather than on their statistical significance.

Emergency cash in conflict affected areas: In response to the conflict along Lebanon's southern border, WFP launched an emergency cash assistance programme targeting all registered households in priority areas – identified based on conflict exposure and displacement intensity – that were not already receiving food assistance. This support began in October 2024 and lasted for at least three months, with some locations continuing to receive transfers as late as May 2025. As a result, by the time round 2 data were collected, 74 percent of the households initially assigned to stop receiving food assistance had received at least three months of emergency cash transfers. While no systematic differences in conflict exposure were observed across targeting arms, the emergency cash likely affected households' trajectories and outcomes, thereby diluting the intended contrast between continued and discontinued assistance groups. This unexpected intervention reflects the complex and rapidly evolving nature of humanitarian contexts, where recurring emergencies frequently reshape programmatic priorities. However, it also poses a challenge to the evaluation design, reducing the ability to cleanly identify the medium-term effects of assistance discontinuation. As such, the estimates presented in this report should be interpreted as lower-bound effects of ending food assistance.¹¹⁶

¹¹⁶ It is possible to estimate the impact of discontinuation for households whose status was unaffected by the emergency cash intervention, using an instrumental variable (IV) approach. However, this falls outside the scope of this report.

A6. Additional details on outcomes

Table A1.9: Details on outcomes of interest

Outcome Group	Outcome	Description
Acceptance / satisfaction with targeting	Appeal rate	Formal appeals can be submitted by households that are excluded from assistance either through the LBCO's call centre or internet form. All formal appeals are recorded by the LBCO.
Food security	Food Consumption Score (FCS)	FCS is a well-established measure of food security, which captures dietary diversity and food consumption frequency across eight food groups over a seven-day recall period. It is calculated by multiplying the number of days each food group was consumed by a standard weight, resulting in a total score that ranges from 0 to 112.
	Household expenditure for essential food and non-food items	Expenditure data was collected following WFP's standard approach. Expenditure on a total of nine different food groups and 11 non-food groups used a standard 30-day recall period. For each food group, respondents report values for items purchased with cash or credit, received as in-kind assistance, or obtained through own production. A six-month recall period was instead used to capture seven less frequent, but often significant, non-food expenses (medical and health care, clothing and shoes, education costs, celebrations and social events, agricultural inputs, and construction or house repairs).
	reduced Coping Strategies Index (rCSI)	The rCSI is a common measure used to assess household food insecurity. It consists of five standardized questions measuring the frequency and severity of behaviours that households adopt to cope with food shortages. It includes relying on less preferred or less expensive foods, borrowing food or relying on help from friends or relatives, limiting portion sizes at mealtimes, limiting adult intake so children can eat, and reducing the number of meals eaten in a day. Each behaviour is assigned a severity weight (1 to 4), and the number of days each behaviour was used in the past seven days is multiplied by its respective weight. The weighted frequencies are then summed to produce the total rCSI score, which ranges from 0 to 56, with higher scores indicating greater food insecurity.
	Percentage of households using livelihood coping strategies	The Livelihood Coping Strategies Index (LCSI) is used by humanitarian organizations to assess the severity of livelihood stress caused by food insecurity or economic shocks. ¹¹⁷ It captures longer-term coping strategies that can undermine future household resilience. The index is composed of four stress, three crisis, and three emergency strategy questions. Usage of a "stress strategy" was then

¹¹⁷ The standard LCSI module uses "the past 30 days" as the timeframe for the questions; however, for this study, a significantly longer timeframe was used to match the study's timeline. For round 1, the timeframe was "since the start of the year" which coincided with the beginning of the new food assistance cycle and the re-targeting of households. For round 2, the timeframe was "since July 2024" which was the endpoint of round 1 household surveys.

		coded as binary based on whether a household reported using at least one type of strategy.
Psychological well-being	Life Satisfaction Index	The life satisfaction score is based on Brück et al. ¹¹⁸ It uses five Likert-scale questions (1 = Very satisfied to 4 = Very unsatisfied) to assess satisfaction with household nutrition, health, finances, community relations, and overall life. To calculate the score, ratings were inverted and then summed together, such that a higher score would indicate better life satisfaction. The index ranges from 5 to 20.
	Worriedness Index	The worriedness index is based on Hidrobo et al. ¹¹⁹ and uses ten Likert-scale questions (1 = Not worried at all to 4 = Very worried) to capture concerns about food, health care, education, debt, employment, family, extremism, conflict, natural disasters, and displacement. Individual responses were summed and the total was then standardized to generate an index. It ranges from 10 to 40.
	World Health Organization-Five Well-Being Index (WHO-5 Index)	Mental well-being is assessed using the WHO-5 Index, which includes five items measuring psychological well-being. ¹²⁰ The items cover feeling cheerful, calm, well-rested, active, and interested in daily life. Each item is rated on a 6-point Likert scale from 0 (At no time) to 5 (All of the time). ¹²¹ The five items are summed to produce a raw score ranging from 0 to 25, with higher scores indicating greater subjective well-being. This score is then rescaled to range from 0 to 100 by multiplying the total by 4, facilitating interpretation and comparison across studies.
	Post-Traumatic Stress Disorder (PTSD) Score	To understand the psychological impact of the conflict with Israel on the study population, a PTSD module was added to the questionnaire for round 2 of the survey. Questions were adapted from the International Trauma Questionnaire and fall under three categories of trauma-related behaviour: re-experiencing in the here and now; avoidance; and a sense of current threat. ¹²² Each item is rated on a 5-point Likert scale from 0 (Not at all) to 5 (Extremely). The responses are summed to produce a dimensional PTSD score which ranges from 0 to 24.
Community Cohesion	Social Cohesion Index	Social cohesion is a standardized index measured using three questions capturing community engagement/relations: membership in community groups, volunteering, and perceived tension in the community. ¹²³ Individual responses were first standardized to have mean zero and standard deviation one within each component of the index. These

¹¹⁸ Brück, T. et al. 2019. Assets for Alimentation? The Nutritional Impact of Assets-based Programming in Niger. *The Journal of Development Studies*, 55(sup1), 55–74. <https://doi.org/10.1080/00220388.2019.1687876>

¹¹⁹ Hidrobo, M., Karachiwalla, N., & Roy, S. (2023). *The Impacts of Cash Transfers on Mental Health and Investments: Experimental Evidence from Mali*. <https://doi.org/10.1016/j.jebo.2023.10.016>

¹²⁰ The standard WHO-5 Index timeframe uses the “previous two weeks” as a reference point; however, for the study, this was changed to the “previous week” to align with the other well-being questions.

¹²¹ WHO. 2024. *The World Health Organization-Five Well-Being Index (WHO-5)*.

<https://www.who.int/publications/m/item/WHO-UCN-MSD-MHE-2024.01>

¹²² Cloitre, M. et al. 2018. The International Trauma Questionnaire: Development of a self-report measure of ICD-11 PTSD and Complex PTSD. *Acta Psychiatrica Scandinavica*. DOI: 10.1111/acps.12956

¹²³ Bossuroy, T. et al. 2022. Tackling Psychosocial and Capital Constraints to Alleviate Poverty. *Nature*, 605(7909), 291–297. <https://doi.org/10.1038/s41586-022-04647-8>

		standardized values were then averaged without weights to form a composite index, which was itself restandardized. ¹²⁴
	Trust Index	The trust index is a standardized index which includes a generalized trust question adapted from the World Values Survey, ¹²⁵ as well as seven Likert-scale items on trust in specific institutions and groups such as family, neighbours, religious organizations, NGOs, local and national government, and migrants. Individual responses were first standardized to have mean zero and standard deviation one within each component of the index. These standardized values were then averaged without weights to form a composite index, which was itself restandardized.
Household Finances	Labour Income	Data on income was collected through a household roster. For each household member engaged in paid work, either through self-employment or a job outside the household, the respondent was asked to estimate the respective member's earnings from the activity in either Lebanese pounds or US dollars and the frequency for which they get paid for the work. All responses were then converted to US dollars and a monthly estimate was calculated by summing the earnings of each household member. Total labour income was then standardized to manage outliers.
	Debt	Households reported whether they had any debt with individuals or institutions, either in Lebanon or abroad. This is a binary indicator.
	Remittances	Households were asked if they had received any non-repayable cash remittances from abroad. This is a binary indicator.
	Perceived financial security	To measure how respondents feel about their level of need, they were asked to rate their level of financial security in comparison with the average Lebanese household on a scale from 0 to 10, where 0 meant that they were completely worse-off and 10 meant that they were completely better-off. These questions were only asked in round 2 of survey data collection.
Employment	Overall employment rate	Data on employment was collected through a household roster. For each household member, the respondent was asked whether they had engaged in any paid work in the previous seven days. For each job reported, respondents provided information on the type of job, start date, and number of hours worked.
	Male employment rate	
	Female employment rate	
Women's Agency	Household decision-making for food	Female labour force participation is a binary indicator that is coded as 1 if the household has at least one female household member who earns an income through paid

¹²⁴ 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>

¹²⁵ Delhey, J. and Welzel, C. 2012. Generalizing Trust: How Outgroup-Trust Grows Beyond Ingroup-Trust. *World Values Research*, WVR Volume 5, Number 3, 2012. <http://dx.doi.org/10.2139/ssrn.2390636>

	Household decision-making for major purchases	work. Women's household decision-making is measured, in round 1 of the survey, through two questions asking about who made decisions regarding: (1) food purchases and (2) major household purchases. For round 2, a third question was added on who made decisions regarding whether household members could engage in paid work. All agency questions are binary and coded as 1 if a woman was involved in decision-making.
	Household decision-making for employment	
Human Capital Development	Enrolment rate (compulsory schooling)	Data on enrolment was collected through a household roster. For each household member of schooling age (6-14 years), the respondent was asked whether the member was currently enrolled in school. The enrolment rate – children enrolled / school-age children – was then calculated at the household level.
	Health-seeking behaviour	Data on health-seeking behaviour was collected through a household roster. For each household member, the respondent was asked whether they had suffered from any illness/injury requiring medical attention since either January 2024 for round 1 or since July 2024 for round 2. The indicator is binary and is coded as 1 if at least one household member did not seek medical attention.

Table A1.10: Details on the construction of composite indices

Indicator	Components	Subcomponents	Possible answers	Weights	Score range
Food Consumption Score (FCS)	Weighted frequency of food groups consumed over past 7 days	Cereals, roots, and tubers	0-7 days	2	0-112
		Pulses	0-7 days	3	
		Vegetables	0-7 days	1	
		Fruits	0-7 days	1	
		Meat and fish	0-7 days	4	
		Milk	0-7 days	4	
		Sugar	0-7 days	0.5	
		Oil	0-7 days	0.5	

Reduced Coping Strategies Index (rCSI)	Frequency of 5 coping strategies used in past 7 days	Rely on less preferred or less expensive foods	0-7 days	1	0-56
		Borrow food or rely on help from a friend or relative	0-7 days	2	
		Limit portion size at mealtimes	0-7 days	1	
		Restrict consumption by adults so children can eat	0-7 days	3	
		Reduce number of meals eaten in a day	0-7 days	1	
Life Satisfaction Score	Sum of 5 questions from Brück et al. ¹²⁶	How satisfied are you with the household's nutrition/food security? (reversed)	1 = Very satisfied to 4 = Very unsatisfied; -999 = Prefer not to disclose	n.a.	0-20
		How satisfied are you with the household's overall health? (reversed)	1 = Very satisfied to 4 = Very unsatisfied; -999 = Prefer not to disclose	n.a.	
		How satisfied are you with the household's financial situation? (reversed)	1 = Very satisfied to 4 = Very unsatisfied; -999 = Prefer not to disclose	n.a.	
		How satisfied are you with the household's interaction with the community and neighbours? (reversed)	1 = Very satisfied to 4 = Very unsatisfied; -999 = Prefer not to disclose	n.a.	

¹²⁶ Brück, T. et al. 2019. Assets for Alimentation? The Nutritional Impact of Assets-based Programming in Niger. *The Journal of Development Studies*, 55(sup1), 55-74. <https://doi.org/10.1080/00220388.2019.1687876>

		How satisfied are you with your life in general? (reversed)	1 = Very satisfied to 4 = Very unsatisfied; -999 = Prefer not to disclose	n.a.	
WHO-5 Well-being Index	Sum of 5 items from WHO-5 well-being scale (multiplied by 4)	I have felt cheerful and in good spirits	1 = Never to 5 = All of the time; -999 = Prefer not to disclose	4	0-100
		I have felt calm and relaxed	1 = Never to 5 = All of the time; -999 = Prefer not to disclose	4	
		I have felt active and vigorous	1 = Never to 5 = All of the time; -999 = Prefer not to disclose	4	
		I woke up feeling fresh and rested	1 = Never to 5 = All of the time; -999 = Prefer not to disclose	4	
		My daily life has been filled with things that interest me	1 = Never to 5 = All of the time; -999 = Prefer not to disclose	4	
Worriedness Score	Sum of 10 questions from Hidrobo et al. ¹²⁷	How worried are you about money for basic food needs?	1 = Not worried at all to 4 = Very worried; -999 = Prefer not to disclose	n.a.	0-40
		How worried are you about money for health care?	1 = Not worried at all to 4 = Very	n.a.	

¹²⁷Hidrobo, M., Karachiwalla, N., & Roy, S. 2023. The Impacts of Cash Transfers on Mental Health and Investments: Experimental Evidence from Mali. *Journal of Economic Behavior and Organization*. Vol 216, Dec 2023. 608-630. <https://doi.org/10.1016/j.jebo.2023.10.016>

			worried; -999 = Prefer not to disclose	
		How worried are you about money for education?	1 = Not worried at all to 4 = Very worried; -999 = Prefer not to disclose	n.a.
		How worried are you about indebtedness?	1 = Not worried at all to 4 = Very worried; -999 = Prefer not to disclose	n.a.
		How worried are you about employment?	1 = Not worried at all to 4 = Very worried; -999 = Prefer not to disclose	n.a.
		How worried are you about family?	1 = Not worried at all to 4 = Very worried; -999 = Prefer not to disclose	n.a.
		How worried are you about extremism?	1 = Not worried at all to 4 = Very worried; -999 = Prefer not to disclose	n.a.
		How worried are you about conflict/sectarian tensions?	1 = Not worried at all to 4 = Very worried; -999 = Prefer not to disclose	n.a.
		How worried are you about natural disasters?	1 = Not worried at all to 4 = Very worried; -999 = Prefer not to disclose	n.a.

		How worried are you about migration/displacement?	1 = Not worried at all to 4 = Very worried; -999 = Prefer not to disclose	n.a.	
Post-Traumatic Stress Disorder Score	Sum of first six items of the International Trauma Questionnaire ¹²⁸	Having upsetting dreams that replay part of the experience or are clearly related to the experience?	0 = Not at all; 1 = A little bit; 2 = Moderately; 3 = Quite a bit; 4 = Extremely	n.a.	0-24
		Having powerful images or memories that sometimes come into your mind in which you feel the experience is happening again in the here and now?	0 = Not at all; 1 = A little bit; 2 = Moderately; 3 = Quite a bit; 4 = Extremely	n.a.	
		Avoiding internal reminders of the experience (e.g. thoughts, feelings, or physical sensations)?	0 = Not at all; 1 = A little bit; 2 = Moderately; 3 = Quite a bit; 4 = Extremely	n.a.	
		Avoiding external reminders of the experience (for example, people, places, conversations, objects, activities, or situations)?	0 = Not at all; 1 = A little bit; 2 = Moderately; 3 = Quite a bit; 4 = Extremely	n.a.	
		Being "super-alert", watchful, or on guard?	0 = Not at all; 1 = A little bit; 2 = Moderately; 3 = Quite a bit; 4 = Extremely	n.a.	

¹²⁸ Cloitre, M. et al. 2018. The International Trauma Questionnaire: Development of a self-report measure of ICD-11 PTSD and Complex PTSD. *Acta Psychiatrica Scandinavica*. DOI: 10.1111/acps.12956

		Feeling jumpy or easily startled?	0 = Not at all; 1 = A little bit; 2 = Moderately; 3 = Quite a bit; 4 = Extremely	n.a.	
Trust Index	Generalized trust	In general, do you think people can be trusted or that they cannot be trusted?	1 = Most people can be trusted; 2 = Need to be very careful; -99 = Don't know; -999 = No answer	n.a.	0-30
	Perception of trust on institutions	How much do you currently trust: Your family (reversed)	1 = No trust to 4 = Complete trust; -999 = Prefer not to disclose	n.a.	
		How much do you currently trust: Your neighbours (reversed)	1 = No trust to 4 = Complete trust; -999 = Prefer not to disclose	n.a.	
		How much do you currently trust: Religious organizations (reversed)	1 = No trust to 4 = Complete trust; -999 = Prefer not to disclose	n.a.	
		How much do you currently trust: International organizations (UN)/NGOs (reversed)	1 = No trust to 4 = Complete trust; -999 = Prefer not to disclose	n.a.	
		How much do you currently trust: Your local government (reversed)	1 = No trust to 4 = Complete trust; -999 = Prefer not to disclose	n.a.	

		How much do you currently trust: Your national government (reversed)	1 = No trust to 4 = Complete trust; -999 = Prefer not to disclose	n.a.	
		How much do you currently trust: Migrants (from outside of Lebanon) (reversed)	1 = No trust to 4 = Complete trust; -999 = Prefer not to disclose	n.a.	
Social Cohesion Index	Collective action Bossuroy et al. ¹²⁹	How many groups or associations (such as youth groups or tontines or religious groups) have you or any household member been a member of since July 2024?	Integer	n.a.	
		Has any member of your household volunteered in community/neighbourhood activities since July 2024?	1 = Yes; 0 = No	n.a.	
	Tensions Bossuroy et al.	Since July 2024, has there been tension between members of your community/neighbourhood? (reversed)	1 = Not at all; 2 = Very little; 3 = Somewhat; 4 = A lot; -99 = I don't know; -999 = Prefer not to disclose	n.a.	

¹²⁹ Bossuroy, T. et al. 2022. Tackling Psychosocial and Capital Constraints to Alleviate Poverty. *Nature*, 605(7909), 291–297. <https://doi.org/10.1038/s41586-022-04647-8>

A7. Additional tables and figures on the main findings

Table A10.1: Differences in selection across the targeting methods

	Coef of selection (robust SE)		
	FCS	CARI	Categorical
Household headed by a woman	0.076*** (0.008)	0.038*** (0.008)	0.286*** (0.008)
Age of the household head	2.872*** (0.252)	1.32*** (0.248)	10.279*** (0.248)
Household head divorced/separated or widowed	0.068*** (0.007)	0.032*** (0.007)	0.238*** (0.008)
Household head is married	-0.111*** (0.008)	-0.062*** (0.008)	-0.349*** (0.008)
Household head suffers from any disability	0.017** (0.007)	0.023*** (0.007)	0.19*** (0.008)
Household head never attended school	0.01 (0.008)	0.005 (0.008)	0.071*** (0.008)
Household head with tertiary education	-0.02*** (0.003)	-0.017*** (0.003)	0.008** (0.004)
Number of household members	-0.721*** (0.034)	-0.216*** (0.034)	-1.546*** (0.034)
Dependency ratio (non-productive/household members)	0.018*** (0.005)	0.007 (0.005)	0.071*** (0.006)
At least one household member has a disability	-0.007 (0.009)	0.024*** (0.009)	0.217*** (0.009)
At least one child attends private school	-0.02*** (0.005)	0.021*** (0.006)	-0.057*** (0.005)
Skilled regular job private	-0.006*** (0.001)	0.001 (0.002)	0.002 (0.002)
Business owner informal	-0.017*** (0.006)	-0.001 (0.006)	-0.032*** (0.006)
Unemployed	0.085*** (0.008)	0.015* (0.008)	0.327*** (0.008)
No Syrians in household	-0.016** (0.008)	0.005 (0.008)	-0.036*** (0.007)
Household does not own house	-0.018** (0.009)	0.054*** (0.009)	-0.047*** (0.009)
Dwelling area per household member	8.369*** (0.395)	2.339*** (0.334)	19.593*** (0.497)
Flush toilet	-0.024*** (0.008)	-0.003 (0.008)	0.035*** (0.007)
Piped water access	-0.032*** (0.008)	-0.009 (0.008)	0.003 (0.008)
Use gas heating	-0.012** (0.005)	-0.013** (0.005)	0.024*** (0.005)
Use gas for cooking	0.011 (0.008)	-0.008 (0.008)	0.054*** (0.008)
Use electricity from own solar panel	-0.014*** (0.004)	-0.007* (0.004)	0.007 (0.005)
WiFi internet (e.g. Cable / ADSL)	-0.127*** (0.008)	-0.056*** (0.009)	-0.04*** (0.009)
Computer / Laptop	-0.006** (0.002)	-0.003 (0.002)	0.02*** (0.003)
Mobile internet	0.051*** (0.009)	0.033*** (0.009)	0.006 (0.009)
Flat screen TV	-0.025*** (0.009)	-0.03*** (0.009)	0.043*** (0.009)
Vacuum cleaner	-0.045*** (0.006)	-0.066*** (0.006)	0.067*** (0.007)
Microwave	-0.017*** (0.004)	-0.025*** (0.004)	0.036*** (0.005)
Air conditioner	-0.024*** (0.005)	-0.034*** (0.004)	0.02*** (0.005)
Scooter / Moped / Motorcycle	-0.022*** (0.005)	-0.022*** (0.005)	-0.055*** (0.005)
Car	-0.013** (0.007)	0.014** (0.007)	0.049*** (0.007)

Note: PMT-Expenditure group serves as the reference category. Multivariate OLS regressions of selection under each study arm on covariate with robust standard error in parentheses. N=25,160 Statistical significance is denoted by: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A1.2: Targeting secondary outcomes – finances

	Labour income (standardized)		Percentage of households w/ debt		Percentage of households receiving remittances		Perceived financial security
	r1	r2	r1	r2	r1	r2	r2
PMT-FCS	-0.072 (0.094)	0.005 (0.053)	0.025 (0.026)	0.030 (0.026)	0.025 (0.017)	-0.009 (0.013)	-0.084 (0.091)
PMT-CARI	-0.069 (0.096)	-0.019 (0.040)	0.017 (0.026)	0.040 (0.026)	-0.008 (0.016)	-0.015 (0.013)	0.001 (0.098)
Categorical	-0.060 (0.093)	-0.060 (0.040)	0.040 (0.026)	0.025 (0.026)	-0.013 (0.016)	0.002 (0.014)	-0.209** (0.095)
Observations	6,358	6,465	6,325	6,450	6,342	6,462	6,464
R-squared	0.001	0.001	0.001	0.001	0.002	0.001	0.002

Note: PMT-Expenditure group serves as the reference category. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results use robust standard errors and apply sampling weights.

Table A1.3: Targeting secondary outcomes – employment

	Employment rate (overall)		Employment rate (male)		Employment rate (Ffemale)	
	r1	r2	r1	r2	r1	r2
PMT-FCS	-0.019 (0.016)	-0.023 (0.015)	-0.002 (0.026)	0.010 (0.026)	-0.023 (0.017)	-0.031* (0.016)
PMT-CARI	0.006 (0.016)	-0.009 (0.016)	0.018 (0.025)	0.016 (0.026)	-0.011 (0.018)	-0.016 (0.017)
Categorical	-0.022 (0.016)	-0.020 (0.016)	0.000 (0.026)	0.026 (0.027)	-0.031* (0.017)	-0.025 (0.017)
Observations	5,860	5,935	4,837	4,908	5,523	5,568
R-squared	0.002	0.001	0.000	0.000	0.002	0.002

Note: PMT-Expenditure group serves as the reference category. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results use robust standard errors and apply sampling weights.

Table A1.4: Targeting secondary outcomes – human capital and agency

Women's agency – household decider									
	Food purchases		Major purchases		Employed household members	Compulsory school enrolment rate		Sought medical attention when needed	
	r1	r2	r1	r2	r2	r1	r2	r1	r2
PMT-FCS	-0.004 (0.019)	0.030 (0.024)	0.009 (0.021)	0.062** (0.025)	0.032 (0.026)	-0.007 (0.028)	-0.023 (0.022)	0.062* (0.037)	0.002 (0.029)
PMT-CARI	-0.016 (0.019)	0.052** (0.024)	-0.002 (0.021)	0.067*** (0.025)	0.031 (0.026)	-0.012 (0.028)	0.012 (0.024)	-0.002 (0.037)	-0.038 (0.029)
Categorical	0.009 (0.019)	0.047** (0.024)	0.023 (0.020)	0.079*** (0.025)	0.030 (0.026)	-0.016 (0.028)	-0.010 (0.022)	0.013 (0.037)	0.012 (0.029)
Observations	6 358	6 465	6 358	6 465	6 465	2 565	2 618	3 176	4 953
R-squared	0.001	0.002	0.001	0.004	0.001	0.000	0.002	0.003	0.001

*Note: PMT-Expenditure group serves as the reference category. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results use robust standard errors and apply sampling weights.*

Table A1.5: Heterogeneity in targeting by gender of household head

	FCS		Expenditure food (USD)		Life Satisfaction Index	
	r1	r2	r1	r2	r1	r2
Female head of household	-4.482*** (1.307)	-2.544* (1.317)	-39.729*** (5.910)	-32.087*** (6.087)	-0.599** (0.247)	-0.910*** (0.217)
PMT-FCS * household headed by a woman	0.437 (1.853)	-0.514 (1.768)	4.018 (8.893)	-11.457 (8.605)	0.068 (0.338)	0.556* (0.297)
PMT-CARI * household headed by a woman	2.428 (1.863)	0.808 (1.879)	3.596 (8.773)	8.372 (8.676)	0.162 (0.334)	0.445 (0.308)
Categorical * household headed by a woman	3.101* (1.862)	1.578 (1.870)	-3.083 (8.723)	-10.296 (8.829)	-0.065 (0.333)	0.574* (0.300)
Observations	6 358	6 465	6 358	6 465	6 240	6 363
R-squared	0.028	0.022	0.056	0.076	0.026	0.024

Note: The reported estimates represent weighted means and the PMT-Expenditure group serves as the reference category. Vertical lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

Table A1.6: Heterogeneity in targeting by age of household head (minor/elderly)

	FCS		Expenditure food (USD)		Life Satisfaction Index	
	r1	r2	r1	r2	r1	r2
Minor/elderly head of household	-1.964 (1.367)	-0.650 (1.310)	-31.860*** (6.924)	-31.455*** (6.561)	-0.314 (0.237)	-0.066 (0.203)
PMT-FCS * minor/elderly head of household	0.354 (1.945)	-0.881 (1.864)	9.735 (10.016)	-5.970 (9.097)	0.204 (0.342)	-0.332 (0.306)
PMT-CARI * minor/elderly head of household	1.906 (1.980)	-0.105 (1.929)	-3.849 (9.666)	2.599 (9.310)	0.209 (0.348)	-0.366 (0.301)
Categorical * minor/elderly head of household	2.573 (1.897)	1.924 (1.881)	-2.809 (9.861)	7.406 (9.493)	0.668* (0.346)	0.094 (0.290)
Observations	6 358	6 465	6 358	6 465	6 240	6 363
R-squared	0.028	0.021	0.050	0.067	0.027	0.024

Note: The reported estimates represent weighted means and the PMT-Expenditure group serves as the reference category. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

Table A1.7: Heterogeneity in targeting by household head education (primary or lower)

	FCS		Expenditure food (USD)		Life Satisfaction Index	
	r1	r2	r1	r2	r1	r2
Head of household ≤ primary	-0.243 (1.280)	-1.215 (1.241)	-6.795 (6.312)	-17.941*** (6.337)	-0.059 (0.240)	-0.275 (0.216)
PMT-FCS * head of household ≤ primary	-1.877 (1.855)	-1.077 (1.769)	-14.721 (9.151)	-0.043 (9.110)	-0.123 (0.325)	-0.146 (0.292)
PMT-CARI * head of household ≤ primary	-1.628 (1.820)	-2.402 (1.778)	-12.787 (8.862)	-9.240 (8.520)	0.272 (0.336)	-0.126 (0.298)
Categorical * head of household ≤ primary	-1.178 (1.797)	-0.869 (1.764)	-9.631 (9.015)	2.766 (8.973)	0.279 (0.335)	0.269 (0.293)
Observations	6 358	6 465	6 358	6 465	6 240	6 363
R-squared	0.025	0.021	0.050	0.068	0.025	0.022

Note: The reported estimates represent weighted means and the PMT-Expenditure group serves as the reference category. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

Table A1.8: Heterogeneity in targeting by presence of member with a disability

	FCS		Expenditure food (USD)		Life Satisfaction Index	
	r1	r2	r1	r2	r1	r2
Householder with disability	-2.645** (1.314)	-1.489 (1.388)	-29.409*** (7.197)	-12.469 (7.631)	-0.488* (0.261)	-0.950*** (0.248)
PMT-FCS * householder with disability	2.414 (2.180)	1.334 (2.130)	18.824* (10.605)	3.313 (11.047)	0.368 (0.385)	0.373 (0.344)
PMT-CARI * householder with disability	1.439 (1.869)	-2.551 (1.967)	7.394 (9.979)	-12.121 (9.663)	0.339 (0.379)	0.411 (0.352)
Categorical * householder with disability	1.182 (1.936)	-1.651 (1.963)	2.091 (9.918)	-8.543 (10.339)	0.561 (0.378)	0.984*** (0.356)
Observations	6 358	6 465	6 358	6 465	6 240	6 363
R-squared	0.028	0.021	0.050	0.067	0.027	0.024

Note: The reported estimates represent weighted means and the PMT-Expenditure group serves as the reference category. Disability status is measured using the Washington Group (WG) Short Set of Questions. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

Table A1.9: Heterogeneity in targeting by household wealth (poorest 50%)

	FCS		Expenditure food (USD)		Life Satisfaction Index	
	r1	r2	r1	r2	r1	r2
Wealth (< median)	-4.645*** (1.210)	-3.306*** (1.192)	-27.541*** (6.213)	-14.348** (6.166)	-0.211 (0.231)	-0.354* (0.204)
PMT-FCS * wealth (< median)	-1.221 (1.736)	0.762 (1.706)	-1.540 (8.876)	-6.645 (8.775)	-0.019 (0.320)	0.044 (0.283)
PMT-CARI * wealth (< median)	-0.844 (1.735)	-2.107 (1.722)	-1.837 (8.662)	-8.045 (8.381)	0.392 (0.327)	-0.283 (0.285)
Categorical * wealth (< median)	-1.574 (1.705)	-0.915 (1.694)	4.146 (8.825)	3.756 (8.707)	0.089 (0.326)	-0.185 (0.280)
Observations	6 358	6 465	6 358	6 465	6 240	6 363
R-squared	0.033	0.024	0.062	0.076	0.026	0.024

Note: The reported estimates represent weighted means and the PMT-Expenditure group serves as the reference category. Below-median wealth refers to households in the bottom 50 percent of the baseline wealth index distribution. The index was constructed using principal component analysis (PCA) based on household assets and access to basic services. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

Table A1.10: Heterogeneity in targeting by dependency ratio

	FCS		Expenditure food (USD)		Life Satisfaction Index	
	r1	r2	r1	r2	r1	r2
Top dep. quartile	0.363 (1.525)	1.024 (1.383)	-14.286* (8.080)	-17.512** (7.113)	0.497** (0.248)	0.169 (0.223)
PMT-FCS * top dep. quartile	0.470 (2.042)	0.114 (1.991)	9.873 (11.115)	8.182 (10.123)	-0.238 (0.348)	-0.383 (0.327)
PMT-CARI * top dep. quartile	0.098 (2.141)	-0.153 (2.048)	-13.080 (10.834)	4.583 (10.069)	-0.805** (0.362)	-0.300 (0.321)
Categorical * top dep. quartile	0.677 (2.006)	-3.102 (1.934)	2.122 (11.217)	-9.575 (9.971)	-0.137 (0.363)	-0.257 (0.323)
Observations	6 358	6 465	6 358	6 465	6 240	6 363
R-squared	0.029	0.022	0.051	0.068	0.027	0.023

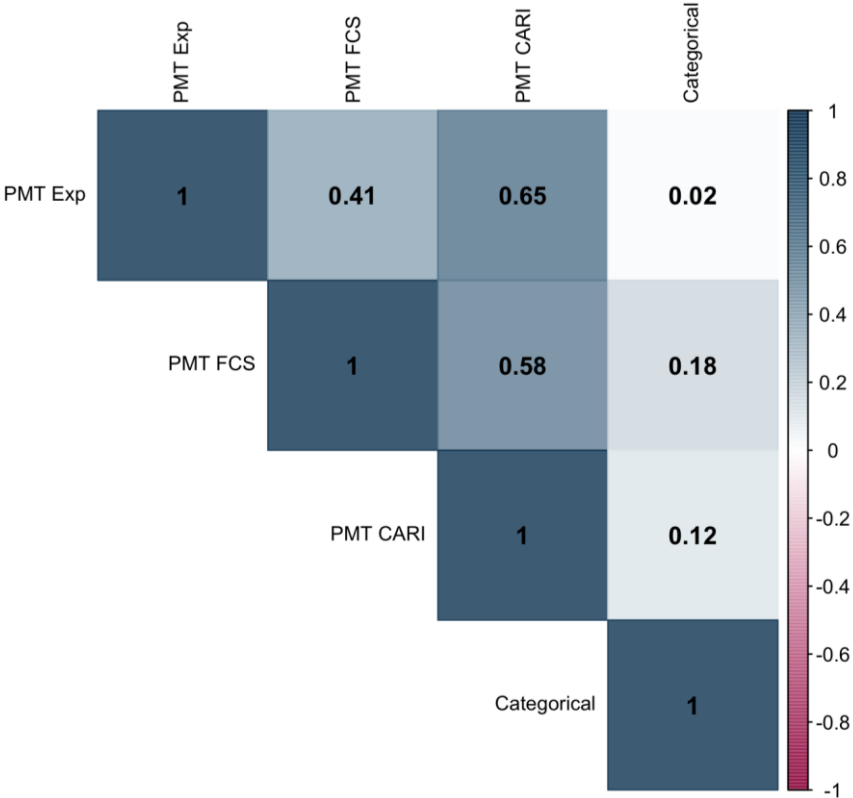
Note: The reported estimates represent weighted means and the PMT-Expenditure group serves as the reference category. The highest dependency ratio quartile captures the top 25 percent of households ranked by the share of non-working-age members, defined as those under 15 or over 64. The ratio was calculated as (children under 15 + adults 64 and over) / total household members of working age. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

Table A1.11: Heterogeneity in targeting by discontinuation status

	FCS		rCSI		Expenditure food (USD)		Worriedness	
	r1	r2	r1	r2	r1	r2	r1	r2
Discontinued assistance	1.071 (1.161)	0.964 (1.150)	-4.113*** (0.845)	-1.085 (0.777)	3.508 (5.850)	-2.701 (5.780)	-0.644 (0.418)	-1.267*** (0.387)
PMT-FCS *	3.518** (1.671)	0.893 (1.635)	1.794 (1.219)	1.191 (1.116)	14.197* (8.434)	15.057* (8.224)	0.724 (0.586)	1.355** (0.547)
PMT-CARI *	2.398 (1.663)	4.872*** (1.642)	0.861 (1.193)	-0.308 (1.102)	5.122 (8.259)	10.023 (7.817)	-0.102 (0.587)	0.313 (0.536)
Categorical * discontinued	1.137 (1.650)	1.378 (1.631)	4.842*** (1.223)	1.718 (1.149)	51.917*** (8.238)	44.340*** (8.139)	2.360*** (0.599)	3.536*** (0.549)
Observations	6 358	6 465	6 358	6 465	6 358	6 465	5 900	6 263
R-squared	0.010	0.010	0.014	0.004	0.031	0.018	0.008	0.017

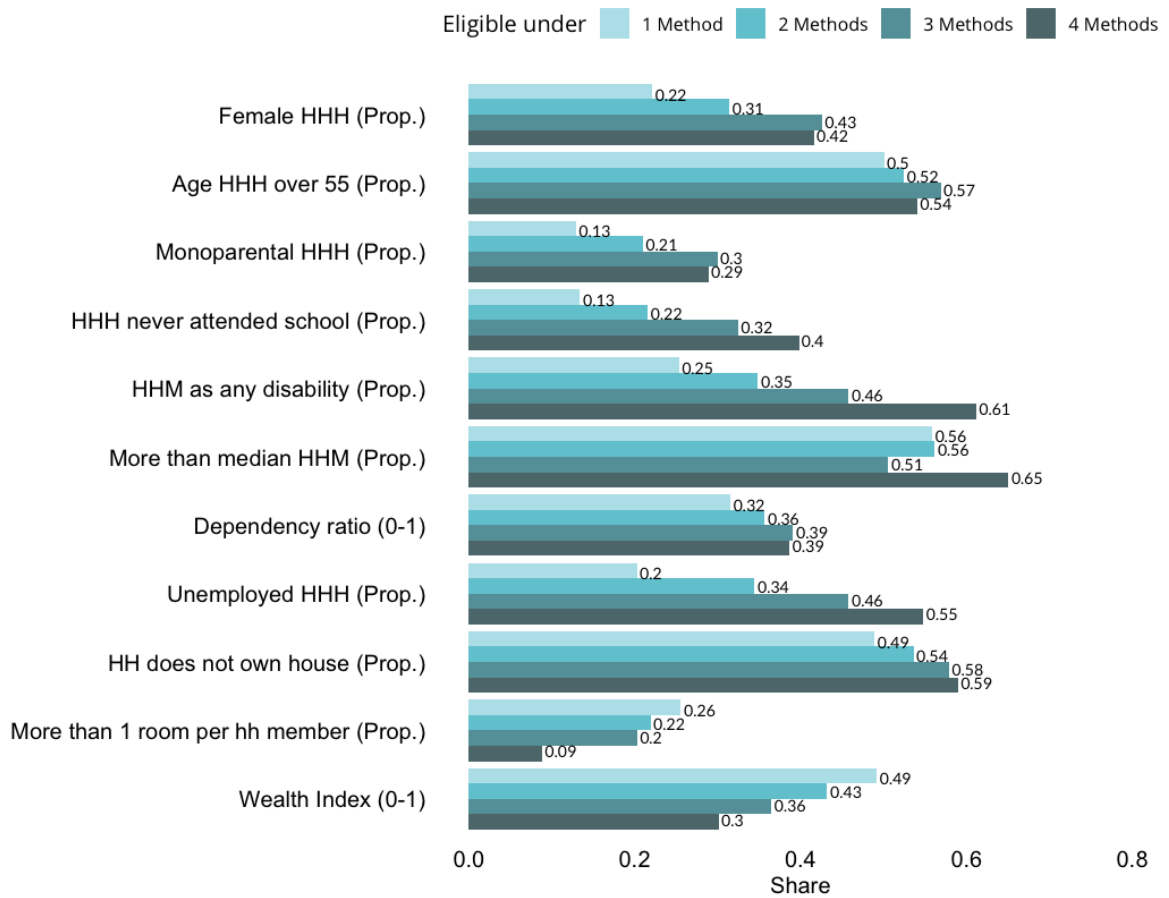
Note: The PMT-Expenditure group serves as the reference category. Statistical significance is indicated by asterisks:
 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Results apply sampling weights.

Figure A1.1: Correlation in selection between the four targeting approaches



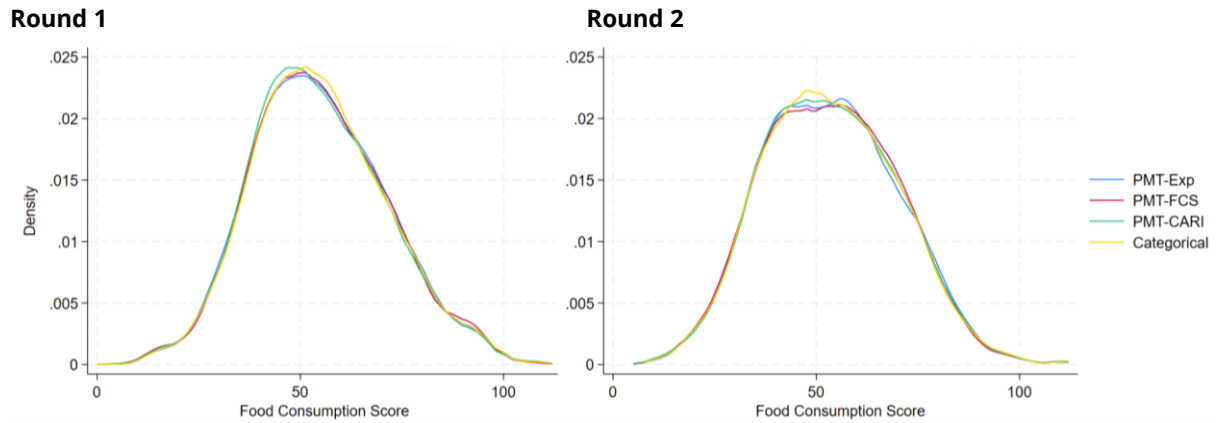
Note: The figure shows the correlation matrix between the selection of the four targeting methods. The colour intensity of each cell represents the strength and direction of the correlation, with darker shades indicating stronger relationships.

Figure A1.2: Predictors of selection of levels of eligibility (using full census)



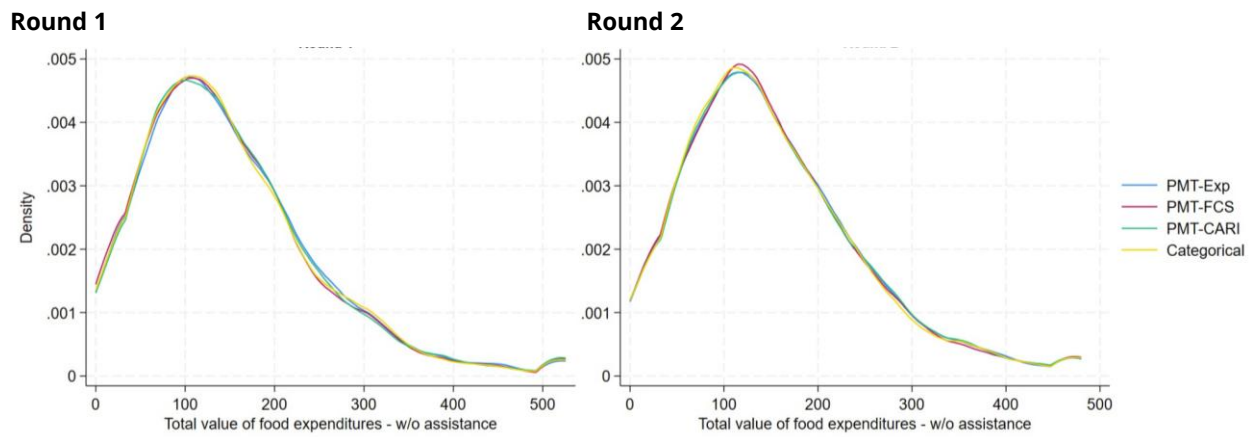
Note: Own visualization. The figure shows the proportion of households interviewed at baseline that fulfil the characteristic reported among households selected to continue receiving food assistance, disaggregated by level of eligibility.

Figure A1.3: Distribution of FCS by targeting arm



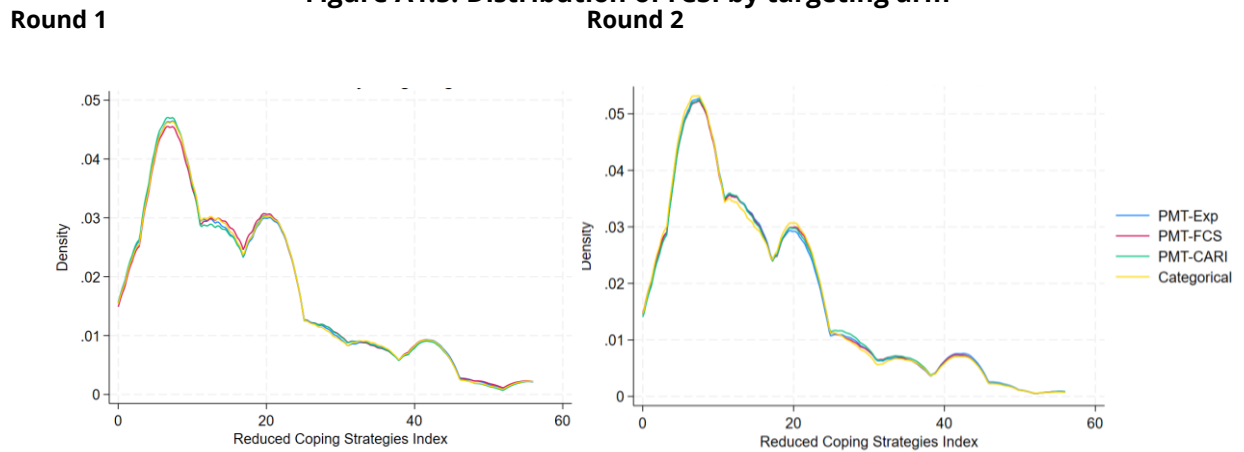
Note: The figures display the distribution of FCS for rounds 1 and 2 disaggregated by targeting method.

Figure A1.4: Distribution of food expenditure (USD) by targeting arm



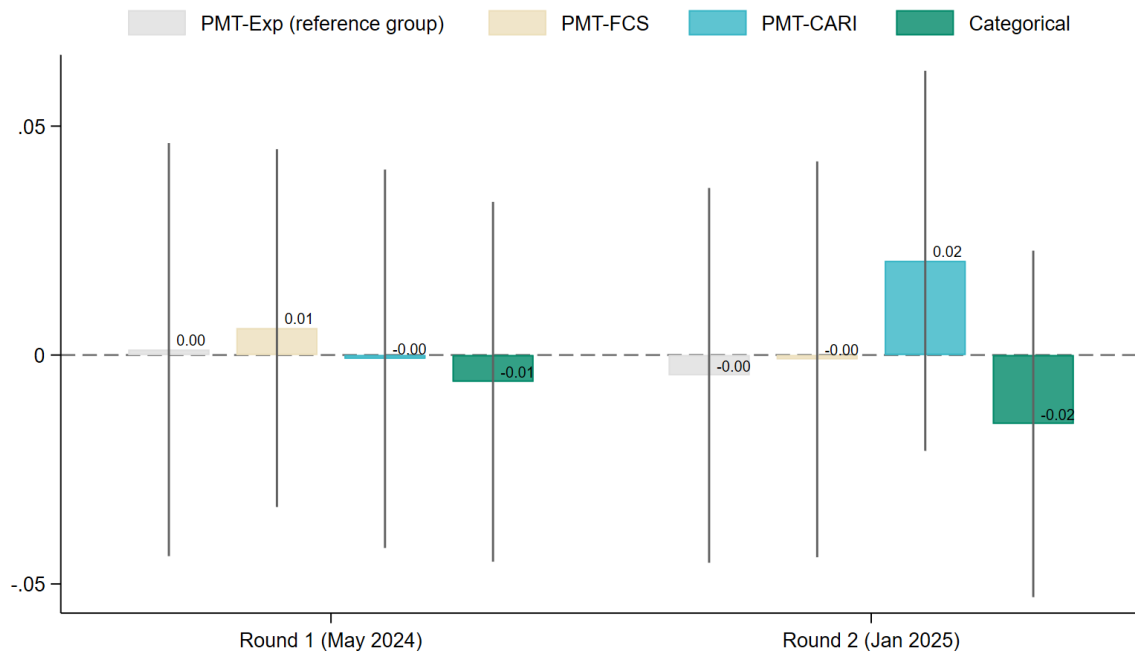
Note: The figures display the distribution of food expenditure in rounds 1 and 2 disaggregated by targeting method.

Figure A1.5: Distribution of rCSI by targeting arm



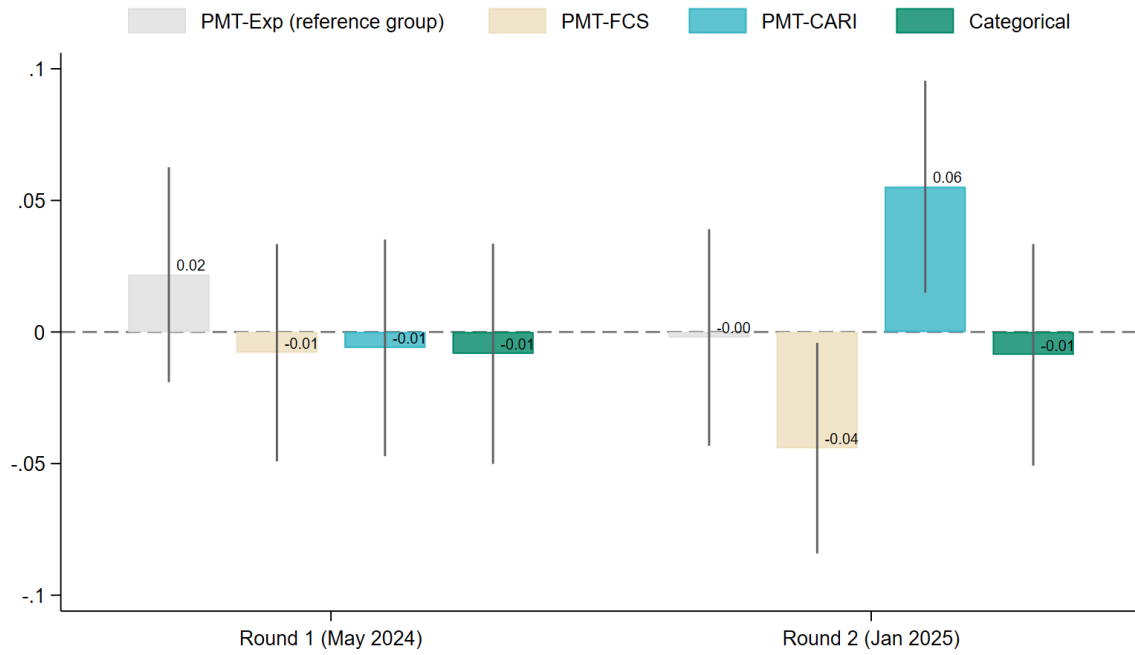
Note: The figures display the distribution of rCSI in rounds 1 and 2 disaggregated by targeting method

Figure A1.6: Social cohesion by targeting arm



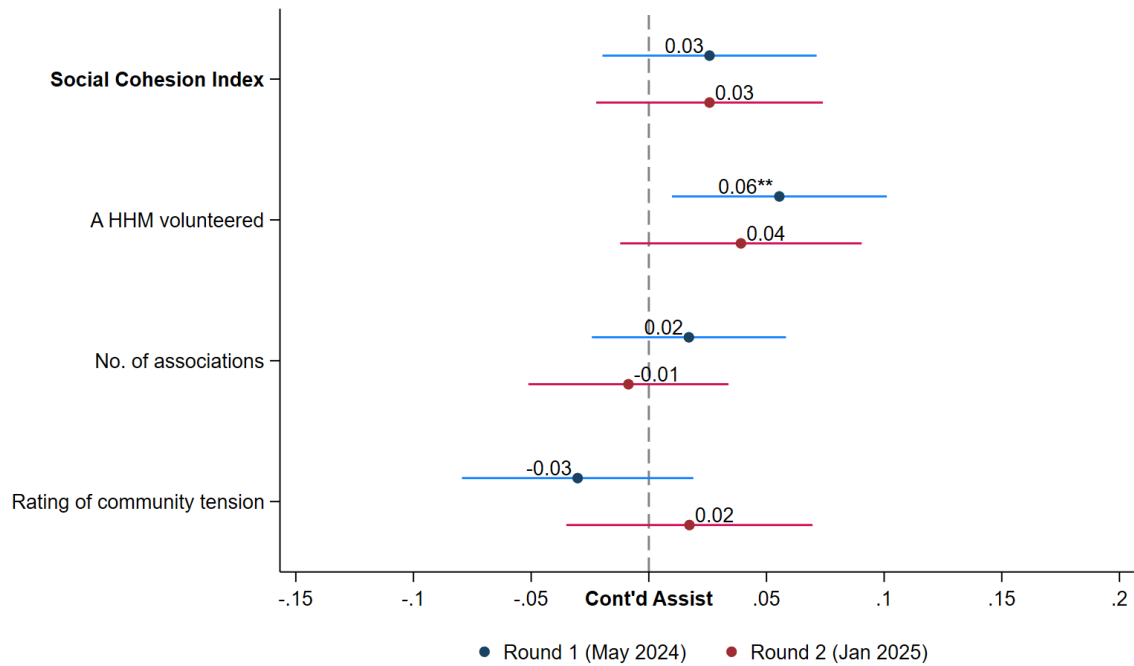
*Note: The Social Cohesion Index is a standardized measure constructed from ten questions capturing trust and community engagement, along with membership in community groups and volunteering activities following the approach of Bossuroy et al. Estimates represent standard deviations compared to the reference category, PMT-Expenditure. Vertical lines indicate 90 percent confidence intervals around the point estimates. Non-overlapping intervals suggest statistically significant differences at the 10 percent level or lower. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All significance tests reflect differences relative to the reference group. Results use robust standard errors.*

Figure A1.7: Trust index by targeting arms



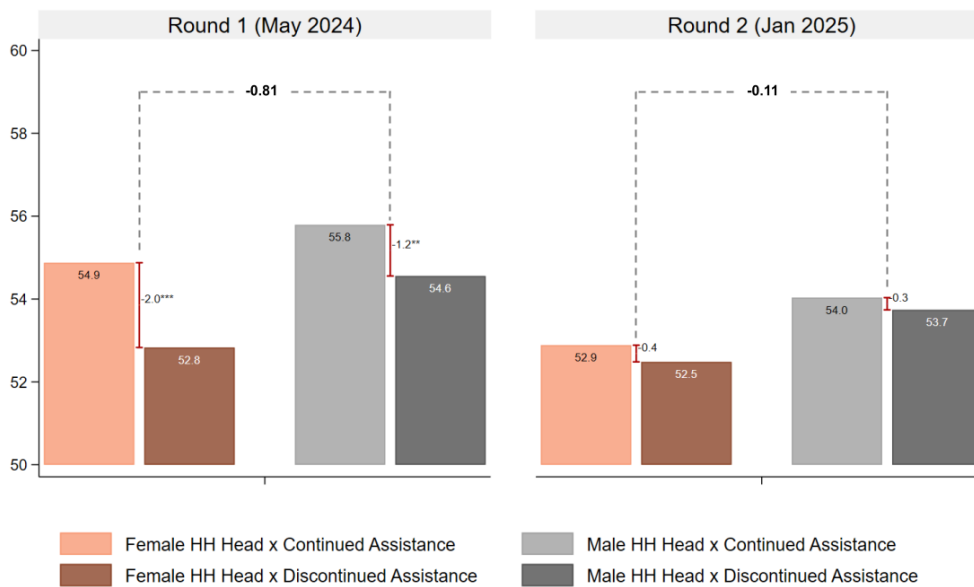
*Note: The Trust Index is a standardized index combining a generalized trust question adapted from the World Values Survey and seven Likert-scale items on trust in specific institutions and groups. Estimates represent standard deviations relative to the reference group, PMT-Expenditure. Vertical lines indicate 90 percent confidence intervals around the point estimates. Non-overlapping intervals suggest statistically significant differences at the 10 percent level or lower. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All significance tests reflect differences relative to the reference group. Results use robust standard errors.*

Figure A1.8: Impact of the discontinuation of assistance on social cohesion



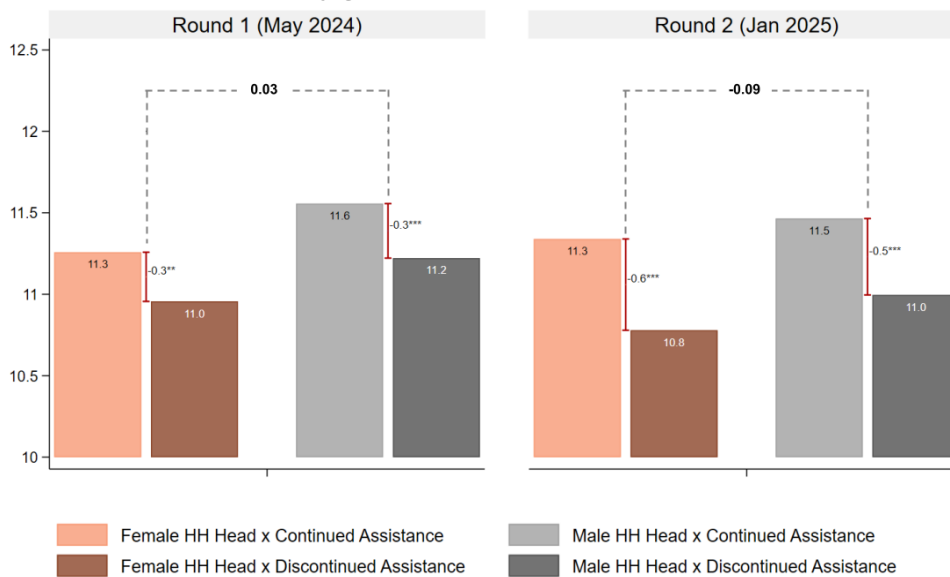
*Note: The figure presents the estimated effects of assistance discontinuation on standardized measures of social cohesion. The Social Cohesion Index is a composite standardized measure: "a household member volunteered" is a binary variable; while the number of associations to which the household belongs and the rating of community tension are both continuous variables. Points reflect changes in standard deviations, with values to the left indicating lower trust among discontinued households. Control variables were selected using post-double selection Lasso, and always include eligibility-space fixed effects and a control for long notice households. For round 2, additional controls include the random assignment of survey modules and survey quality. Horizontal lines indicate 90 percent confidence intervals around the point estimates. Statistical significance is indicated by asterisks: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.*

Figure A1.9: Differential impacts of discontinuation on FCS by gender of the household head



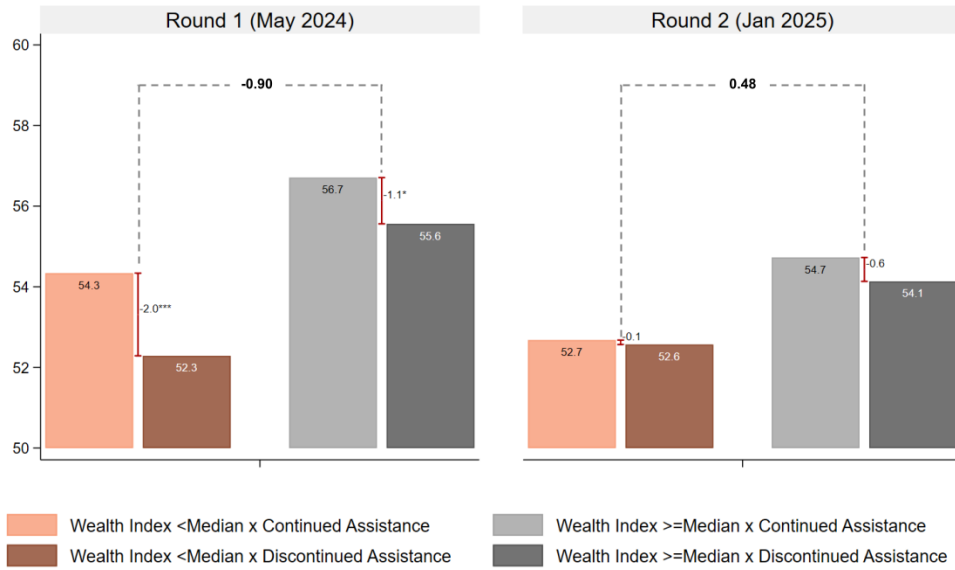
Note: The figures show mean values for combinations of household head gender (female or male) and programme status (continued or discontinued) for each round. The point estimates within each gender group indicate the group-specific average treatment effect. Differences between gender groups indicate heterogeneous treatment effects. Statistical significance is denoted by asterisks: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Figure A1.10: Differential impacts of discontinuation on life satisfaction by gender of the household head



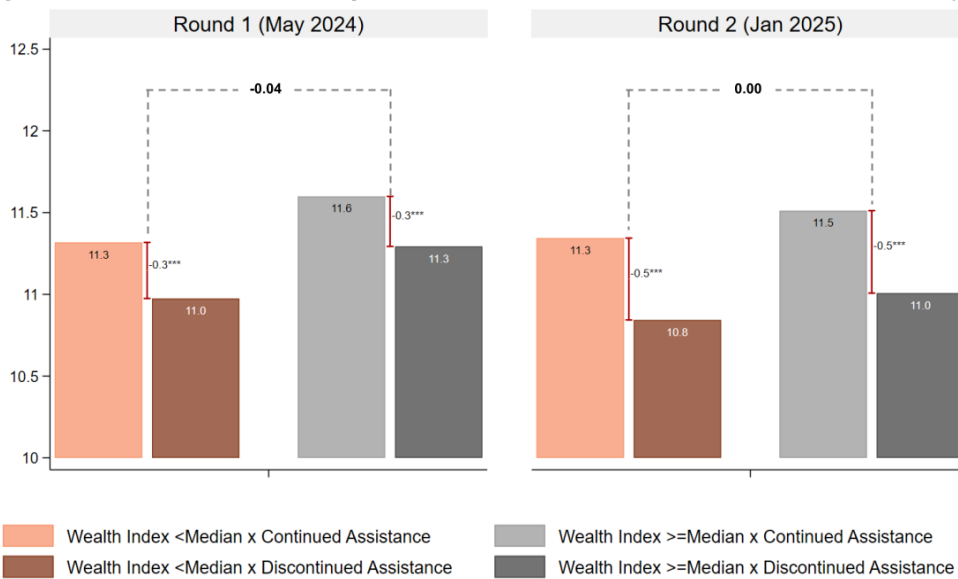
Note: The figures show mean values for combinations of household head gender (female or male) and programme status (continued or discontinued) for each round. The point estimates within each gender group indicate the group-specific average treatment effect. Differences between gender groups indicate heterogeneous treatment effects. Statistical significance is denoted by asterisks: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Figure A1.11: Differential impacts of discontinuation on FCS by wealth



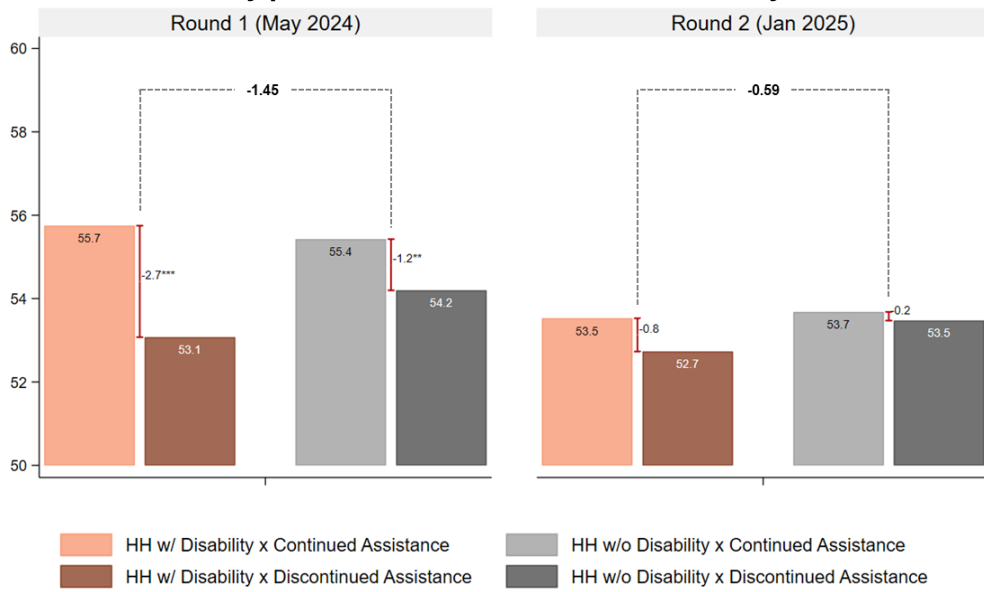
Note: The figures show mean values for combinations of household wealth (below or at/above median) and programme status (continued or discontinued) for each round. The point estimates within each wealth group indicate the group-specific average treatment effect. Differences between wealth groups indicate heterogeneous treatment effects. Statistical significance is denoted by asterisks: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Figure A1.12: Differential impacts of discontinuation on life satisfaction by wealth



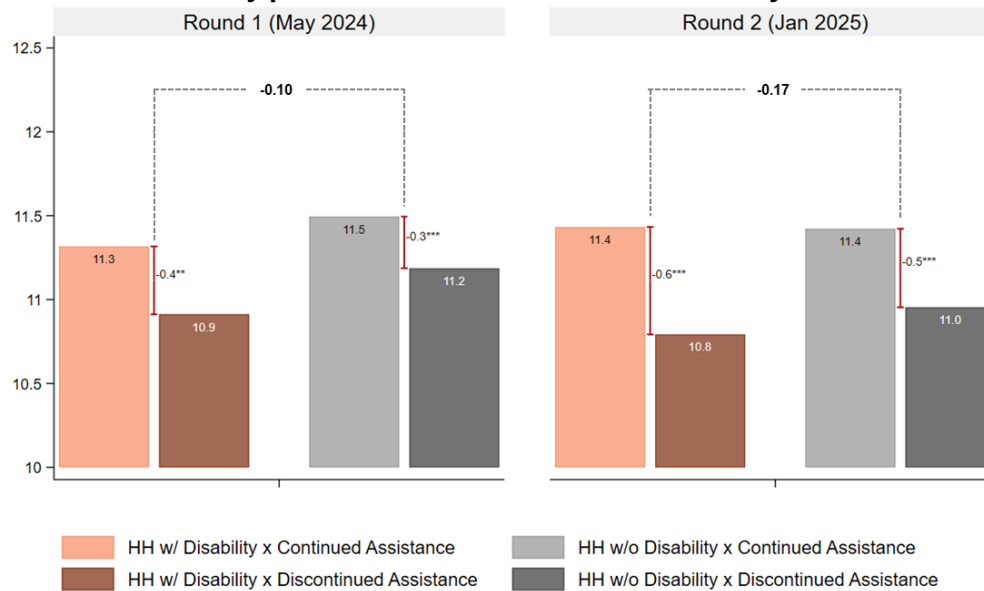
Note: The figures show mean values for combinations of household wealth (below or at/above median) and programme status (continued or discontinued) for each round. The point estimates within each wealth group indicate the group-specific average treatment effect. Differences between wealth groups indicate heterogeneous treatment effects. Statistical significance is denoted by asterisks: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Figure A1.13: Differential impacts of discontinuation on FCS by presence of a member with a disability



Note: The figures show mean values for combinations of the presence of disability (having or not having a household member with a disability/chronic illness) and programme status (continued or discontinued) for each round. The point estimates within each disability group indicate the group-specific average treatment effect. Differences between disability groups indicate heterogeneous treatment effects. Statistical significance is denoted by asterisks: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Figure A1.14: Differential impacts of discontinuation on life satisfaction by presence of a member with a disability



Note: The figures show mean values for combinations of the presence of disability (having or not having a household member with a disability/chronic illness) and programme status (continued or discontinued) for each round. The point estimates within each disability group indicate the group-specific average treatment effect. Differences between disability groups indicate heterogeneous treatment effects. Statistical significance is denoted by asterisks: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

A8. Data collection and ethical considerations

237. The survey data used in this study was collected by the Central Arab Statistical Network (CASN), an independent survey firm with extensive experience in large-scale data collection across Lebanon. Overall, more than 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). This popular software is well-suited for handling confidential data, featuring end-to-end encryption accessible only to authorized team members.
238. During data collection, the research team and field coordinator supported the survey firm with planning, logistics, and quality control. The field coordinator conducted daily monitoring of the submitted survey data, running a high-frequency check code to identify any inconsistencies, outliers, duplicate cases, missing data, or other errors present in the data. The field coordinator maintained constant communication with the field team to keep them informed about the data situation and to promptly relay any significant information to all enumerators.
239. Evaluations conducted by WFP must conform to the 2020 United Nations Evaluation Group (UNEG) ethical guidelines. Accordingly, the Office of Evaluation and International Security and Development Center (ISDC) are responsible for safeguarding and ensuring ethics at all stages of the evaluation cycle. This includes, but is not limited to, ensuring informed consent, protecting privacy, confidentiality, and anonymity of participants, ensuring cultural sensitivity, respecting the autonomy of participants, ensuring fair recruitment of participants and that the evaluation results in no harm to participants or their communities.
240. This study also adheres to the note on [Ethical Guidelines for Implementing Impact Evaluations in WFP](#) which outlines the ethical considerations for each phase of an impact evaluation. In addition, ISDC has adopted the [Ethical Principles for Collecting Primary Data and Conducting Fieldwork](#). The Ethical Principles complement ISDC's [Code of Conduct](#), its [Guideline for Ensuring Good Scientific Practice](#) and its [Data Protection Policy](#). The Ethical Principles are binding for all staff and subcontractors of ISDC, and the research team strictly follows these practices and guidelines.
241. All households that participated in the study provided informed consent to participate in the study and be surveyed. Consent was obtained separately for each survey round. Participation in the study was entirely voluntary and had no impact on eligibility for any form of assistance. Respondents could decline to answer any question or end the interview at any time, and they always retained the right to withdraw from the study without penalty. Participants were clearly informed that all data would be anonymized, used solely for learning purposes aligned with the study's objectives.
242. The evaluation team followed a strict "do no harm" approach,¹³⁰ prioritizing the safety, dignity, and well-being of all participants and staff. Enumerators received at least one week

¹³⁰ Anderson, M. B. 1999. *Do No Harm: How Aid Can Support Peace-Or War*. Lynne Rienner Publishers.

of training, including field piloting, and were instructed in ethical protocols, privacy safeguards, and context-appropriate interviewing. To support respondent comfort and enumerator safety, teams were gender-balanced to the extent possible. No incidents were reported during the study period.

243. The study was approved by the ethics committees of the Humboldt-Universität zu Berlin (#2023-08-W) and by HML International Review Board (#2766).

A9. Stakeholder analysis

244. The stakeholder analysis identifies those who may influence or be influenced by the study results. Stakeholders include both internal and external actors, including households included in the programme. While the primary user of this evaluation is the WFP Country Office in Lebanon, the findings are intended to inform a broader audience within the humanitarian sector.

245. Stakeholders are categorized as follows:

- **Internal stakeholders in Lebanon:** Key personnel within the WFP Lebanon Country Office.
- **Internal stakeholders outside Lebanon:** Includes the WFP Office of Evaluation, the WFP regional bureau, and relevant headquarters divisions.
- **External stakeholders:** Comprising international NGOs, donors, United Nations agencies, and local coordination forums.
- **National stakeholders:** Encompassing government entities at both national and subnational levels, as well as local NGOs.
- **Academia and researchers:** Including local universities and research institutions, where the evaluation's design and findings will be shared to contribute not only to this specific topic but also to broader humanitarian research.

246. This engagement ensures that diverse perspectives are considered and that the evaluation's findings are effectively utilized by stakeholders. Given the limited rigorous evidence on the scaling down of humanitarian assistance, this study also aims to contribute to the broader literature. The results will serve as public goods, providing a foundation for further analysis and policy development on this topic.

A10. SMS reminder messages

247. The SMS messages to households read as follows:

Continued assistance households: "Based on the latest data collection exercise conducted in 2023, the World Food Programme would like to inform you that your family will continue to receive in-kind food assistance in 2024 on a monthly basis (from January to December 2024). However, please note that assistance is subject to funding availability. For more information, please call WFP Call Centre on 1526 from Monday till Saturday from 8 a.m. to 7 p.m."

Discontinued households: "Based on the latest data collection exercise conducted in 2023, the World Food Programme has prioritized families in dire need of food assistance and regrets to inform you that, after careful consideration, your family did not qualify for assistance in 2024. To appeal for reconsideration, please submit a claim through the Call Centre on 1526. In case you could not, please use the following link: <https://grm.beneficiaryregistration.lbn.wfp.org/> (valid until 15 February 2024)."

Long notice households: "Based on the latest data collection exercise conducted in 2023, the World Food Programme would like to inform you that your family will continue to receive in-kind food assistance on a monthly basis, for a period of 6 months (from January to June 2024). For more information, please call WFP Call Centre on 1526 from Monday till Saturday from 8 a.m. to 7 p.m."

Second SMS to long notice households: "Based on the last communication by WFP and as per the data collection exercise that was conducted in 2023 in which the World Food Programme has prioritized families in dire need of food assistance, this is to remind you that your family will continue receiving in-kind food assistance on a monthly basis only until June 2024. As of July, assistance will stop. For more information, please call WFP Call Centre on 1526 from Monday till Saturday from 8 a.m. to 7 p.m."

Office of Evaluation

World Food Programme

Via Cesare Giulio Viola 68/70

00148 Rome, Italy

T +39 06 65131 **wfp.org**