Oxford Poverty & Human Development Initiative (OPHI)
Oxford Department of International Development
Queen Elizabeth House (QEH), University of Oxford

OPHI *Working Paper No. 142*

Distributional Impacts of Cash Transfers on the Multidimensional Poverty of Refugees: The ESSN Programme in Turkey

Matthew Robson[†], Frank Vollmer[‡], Basak Berçin Doğan[§] and Nils Grede[§]

August 2022

Abstract

Most impact evaluations of humanitarian cash transfer programmes use traditional metrics of poverty and study average effects of outcomes separately. We analyse the impact of the Emergency Social Safety Net (ESSN) cash programme on the multidimensional poverty of refugees in Turkey, using a purpose-built Refugee Multidimensional Poverty Index (RMPI). We conduct a causal analysis of both average and distributional impacts of the ESSN on the incidence and intensity of multidimensional poverty, and decompose effects for separate dimensions of poverty. Results show that the ESSN significantly reduced the incidence and intensity of multidimensional poverty amongst its beneficiaries. Significant reductions are found in the dimensions of food security, living standards and education. This supports emerging claims that these types of programmes, still relatively new in humanitarian contexts, can be transformative for their beneficiaries to achieve multiple outcomes simultaneously. Reductions in the deprivation scores of the more deprived households stand out as a finding that outcome-specific evaluations and multidimensional impact evaluations focusing on estimating average treatment effects would have missed, demonstrating the added value of the distributional analyses used. By learning from the largest humanitarian cash programme in the world, results provide important lessons for cash programmes on multidimensional poverty of refugees elsewhere.

Keywords: Cash transfers, Refugees, Multidimensional poverty, Impact evaluation, Humanitarian intervention, Turkey

JEL Classification: I38, C21, D63.

[†]Oxford Poverty and Human Development Initiative (OPHI), University of Oxford, Oxford, UK, and Department of Health Sciences, University of York, York, UK. Email: matthew.robson@york.ac.uk.

[‡]Oxford Poverty and Human Development Initiative (OPHI), University of Oxford, Oxford, UK. Email: frank.vollmer1@gmx.de.

[§]World Food Programme (WFP), Turkey.

This study has been prepared within the OPHI theme on multidimensional measurement.

Acknowledgements
We would like to thank Sabina Alkire, İlgi Bozdağ, Owen O'Donnell, Lena Hohfeld, Corinne Mitchell, Hector Moreno, Ricardo Nogales, Nicolai Suppa and Keeyoul Yoon for their helpful comments and collaboration. Comments from participants at the 'Viewing Multidimensional Poverty From Many Angles' seminar and the 'Multidimensional Poverty among Refugees and Forcibly Displaced Populations' workshop, University of Oxford, are greatly appreciated. The contents of the paper are the responsibility of the authors and do not reflect the views of OPHI, WFP, the Turkish government and other institutions, or individuals who were involved in the study. All errors remain our own.

Funder: Data used in this study was provided by the World Food Programme (WFP). Funding for this study from the Directorate-General for European Civil Protection and Humanitarian Aid Operations (DG ECHO) is greatly acknowledged. Matthew Robson is also grateful for part funding from the Wellcome Trust (205427/Z/16/Z).

Citation: Robson, M., Vollmer, F., Doğan, B., and Grede, N. (2022): 'Distributional Impacts of Cash Transfers on the Multidimensional Poverty of Refugees: The ESSN Programme in Turkey', OPHI Working Paper 142, University of Oxford.

The Oxford Poverty and Human Development Initiative (OPHI) is a research centre within the Oxford Department of International Development, Queen Elizabeth House, at the University of Oxford. Led by Sabina Alkire, OPHI aspires to build and advance a more systematic methodological and economic framework for reducing multidimensional poverty, grounded in people's experiences and values.

Oxford Poverty & Human Development Initiative (OPHI)
Oxford Department of International Development
Queen Elizabeth House (QEH), University of Oxford
3 Mansfield Road, Oxford OX1 3TB, UK

Tel. +44 (0)1865 271915 Fax +44 (0)1865 281801
ophi@qeh.ox.ac.uk http://www.ophi.org.uk

ISSN 2040-8188 ISBN 978-1-912291-33-5

1 Introduction

Cash transfer programmes are well established in national poverty reduction and social protection strategies (Farrington and Slater, 2006; Fiszbein and Schady, 2009) and are increasingly being applied in refugee and humanitarian contexts (Bagstagli et al., 2016; World Bank, 2016).[1] While the impacts of cash transfers on individual outcomes (monetary poverty, for example) are well understood (Galiani and McEwan, 2013; Golan et al., 2017),[2] an emerging research agenda looks at the impacts of cash transfers on multiple outcomes simultaneously, as captured through a multidimensional poverty metric (Song and Imai, 2019; Seth and Tutor, 2021). This is grounded in the understanding of poverty as multidimensional and allows for the evaluation of the multi-purpose designs of most poverty, social and humanitarian programmes (Vaz et al., 2019). While such studies point to the effectiveness of cash-based programming in anti-poverty and social contexts (Masset and García-Hombrados, 2021; Mitchell and Macció, 2021), rigorous findings on the impacts of multi-purpose programmes on simultaneous outcomes in humanitarian contexts are only nascent (Loschmann et al., 2015). Therefore, given that most impact evaluations of humanitarian cash transfer programmes still use traditional metrics of poverty and study average effects of outcomes separately (Tappis and Doocy, 2018), there is a lack of evidence on multidimensional impacts of such multi-purpose programmes in refugee settings.

To study such impacts, we focus on the Emergency Social Safety Net (ESSN) programme. The ESSN was launched in Turkey in 2016 and is now the largest humanitarian cash transfer programme in the world. The ESSN provides monthly unconditional cash transfers to poor and vulnerable refugees, enabling them to decide how to cover costs of living to meet basic needs such as rent, transport, bills, food and medicine (Cuevas et al., 2019; Cetinoglu and Yilmaz, 2021; Gibárti et al., 2021). While previous evaluation exercises have looked at a range of outcomes of the programme separately (Maunder et al., 2018; Aygün et al., 2021; Özler et al., 2021), there is interest in understanding how this type of programme might impact the beneficiaries' overall multidimensional poverty (WFP, 2019a). The programme is therefore an ideal candidate for such a study as it used cash-based multi-purpose assistance

[1] For further papers on poverty reduction and social protection contexts, see: Skoufias et al. (2001); Slater (2011); Haushofer and Shapiro (2016); Handa et al. (2018); Bastagli et al. (2019); Molyneux et al. (2016); Hjelm et al. (2017). For refugee and humanitarian contexts see Hagen-Zanker et al. (2016); Doocy and Tappis (2017); Ulrichs et al. (2017); Hagen-Zanker et al. (2018); Ulrichs and Sabates-Wheeler (2018); Özler et al. (2021); MacPherson and Sterck (2021); Salti et al. (2021).

[2] Such programmes create income multiplier effects (Sadoulet et al., 2001), are flexible (Blattman and Niehaus, 2014) and have proven economic and administrative advantages as they are less likely to distort prices than in-kind transfer programmes (Tabor, 2002).

on an unparalleled scale in a humanitarian context.

This paper takes a new approach by applying a tailored Refugee Multidimensional Poverty Index (RMPI) (OPHI and WFP, 2022), based on the Alkire–Foster method (Alkire and Foster, 2011). This index is used to evaluate the impact of the ESSN on multidimensional poverty of refugees in Turkey. We estimate causal effects by using a doubly-robust inverse-probability-weighted regression-adjustment and rely on three waves of a repeated cross-sectional dataset (2018–20), the Comprehensive Vulnerability Monitoring Exercise (CVME), which is representative of refugees at the national level (n=4,042). Moreover, we utilise distributional methods which go beyond estimating average treatment effects. By combining these methods with the RMPI, we allow for the estimation of causal effects on the incidence and intensity of multidimensional poverty and the identification of heterogeneous effects across the quantiles of the deprivation score. In addition, we use a novel approach which estimates the entire distribution of probabilistic potential outcomes. We use the beta-binomial distribution to separately estimate marginal distributions of each dimension of poverty to provide a joint estimate of deprivation scores.[3] This allows for more precise predictions of outcomes, can take account of multiple dimensions and enables the prediction of convenient statistics (e.g. expected intensity, headcount ratios).

The paper thus contributes to the literature in three ways. First, we apply a purpose-built Multidimensional Poverty Index for refugees that allows for the measurement of the incidence and intensity of multidimensional poverty amongst refugees in Turkey. Second, we utilise distributional analysis which, combined with this purpose-built index, allows us to go beyond the mean and estimate treatment effects across the entire distribution of the deprivation score. This enables a more nuanced analysis of the heterogeneous effects of interventions on the incidence and intensity of multidimensional poverty. Third, we use these methods to evaluate the impact of the world's largest humanitarian cash transfer programme on the multidimensional poverty of refugees. The results will inform ongoing debates on poverty reduction potential of cash programmes in social and humanitarian contexts.

The structure of this paper is as follows: Section 2 reviews the existing literature on multidimensional poverty evaluation design and articulates how this paper makes novel contributions to current knowledge, as well as describing the background of the ESSN programme; Section 3 explains the methods and data used in this paper for both the RMPI and the impact evaluation; Section 4 presents the results of the RMPI and impact evaluation; and Section 5

[3]This approach is generally applicable to any discrete, evenly spaced, bounded outcome (e.g. MPI deprivation score, income rank).

discusses how these results link to the broader humanitarian programme and impact evaluation contexts. The paper ends with some concluding remarks.

2 Literature

2.1 Multidimensional Poverty Indices in Impact Evaluations

A growing focus in the academic literature on impact evaluations of anti-poverty programmes is placed on reusing counting approaches (Atkinson, 2003; Alkire and Foster, 2011). This aims to better analyse the overlapping simultaneous impacts of anti-poverty programmes, such as conditional and unconditional cash transfers, that increasingly tend to have multiple outcome targets. These impact evaluations rely primarily on three types of multidimensional poverty indices: programmatic, general and population-specific.

Seth and Tutor (2021) applied a 'programmatic' multidimensional impact evaluation framework to Pantawid Pamilyang Pilipino in the Philippines, a conditional cash anti-poverty programme. By evaluating multiple outcomes simultaneously, their approach went beyond prior impact evaluations which focused on programme outcomes and non-compliance separately. They found that for 'the poorest families, the cash grants may only be enough to marginally improve their consumption, but are not sufficient to alleviate other associated deprivations' (Seth and Tutor, 2021). Similarly, Vaz et al. (2019) evaluated the conditional cash transfer programme Progresa, in Mexico, using a multidimensional evaluative framework. The authors used the objectives of the programme to construct a multidimensional index to assess impact. The greatest contributors to the multiple objectives of Progresa were found to be school attendance and health visits, indicating that the programme had significant multidimensional effects.[4]

With a multidimensional evaluation of unconditional cash transfers (UCTs), Song and Imai (2019) evaluated Kenya's Hunger Safety Net Programme (HSNP). Song and Imai developed the HSNP MPI, a poverty measure designed to study programme status and outcomes consisting of thirteen indicators across three dimensions (consumption poverty, food insecurity and asset accumulation). They showed that the programme significantly reduced multidimensional poverty, primarily through the food insecurity dimension. Their study shines light on the impact of the programme on the *simultaneous deprivations* that poor people face

[4]For more applications of the Alkire–Foster method to impact evaluations, see Mitchell and Macció (2021); Wang et al. (2021).

and clearly identifies the main driver out of poverty.

These examples of conditional and unconditional cash transfers study the impact on indicators targeted by the programme; thus, they are designed to capture anticipated (or assumed) outcomes of these cash programmes, making these studies and their findings highly contextual.

This can be contrasted with evaluations where 'general' multidimensional poverty indices (MPIs) were used. Masset and García-Hombrados (2021) assessed how well the Savannah Accelerated Development Authority (SADA) Northern Ghana Millennium Village Project (MVP) performed in an impact evaluation using the global Multidimensional Poverty Index (global MPI). They assessed the programmes' joint impact on the incidence and intensity of poverty and on the ten indicators of the index separately. They identified a significant reduction in the incidence of poverty, but found that the intensity of poverty only marginally decreased. The impact was driven by two of the ten indicators only, namely, school attendance and sanitation. The advantage of using the latter approach is that it made use of established poverty measures, thus producing findings that are highly relevant and coherent for a general context. The authors explicitly suggest to avoid building ad hoc indices based on the expected outcomes of specific programmes, but highlight the disadvantage of their approach as 'the global MPI is very sensitive to changes in some deprivations and that it does not include important welfare dimensions affected by development programmes' (Masset and García-Hombrados, 2021, p.14).

A third option is to use a specifically designed 'poverty' MPI for certain *population* groups, such as refugees. One such example is offered by Loschmann et al. (2015), who adopted a multidimensional approach to assess the impact of the UNHCR post-return shelter assistance programme in Afghanistan 2009—11. They designed an MPI to measure wellbeing for returning refugees reintegrating in Afghanistan consisting of three dimensions (economic welfare, health and education, and basic services) and nine indicators, including dietary diversity and food security, that can also be considered crucial indicators in refugee contexts in general. They found that significant reductions in the MPI were primarily driven by indicators in dietary diversity, food security and heating. The authors demonstrated the added value and applicability of a tailored MPI in a refugee context, albeit more narrow in its focus on those returning, and by doing so were able to design a specific MPI that captures both the (rather immediate) needs and vulnerabilities of refugees, but also the (rather longer term) overall wellbeing of recipient households across several socio-economic indicators.

We contribute to this literature, where multidimensional poverty measures are used to evaluate impacts of social programmes, by first, presenting and applying a specifically designed MPI that moves towards the measurement of *poverty in dimensions and indicators that are specifically crucial to refugees*, and second, by using the index to measure and evaluate the impact of an unconditional cash programme (the ESSN) in a dynamic humanitarian context (Turkey). Study findings will contribute to the ongoing debate on how cash programmes for refugees best target their immediate humanitarian needs and longer-term livelihood potential and protection.

2.2 Treatment Effects

A long line of research in the applied econometrics literature is concerned with estimating causal treatment effects. Since Rubin's seminal paper, the *potential outcomes* framework has been used as the theoretical framework to estimate average treatment effects in many settings (Rubin, 1974). However, the literature has long since moved beyond estimating average effects alone (Heckman and Robb Jr, 1985; Cornelissen et al., 2016; Bedoya et al., 2017). Conditional average treatment effects allow for the identification of heterogeneity by observable characteristics. Quantile treatment effects, both conditional (Koenker and Bassett Jr, 1978) and unconditional (Firpo et al., 2009), and distributional treatment effects (Chernozhukov et al., 2013) reveal how the treatment changes the distribution of the outcome variable. And inequality treatment effects (Firpo and Pinto, 2016) reveal how overall inequality changes, according to an index (e.g. the Gini coefficient (Gini, 1912) or Atkinson index (Atkinson, 1970)).

This literature allows for more nuance in the estimation of effects of interventions on the distribution and inequality in outcomes. However, these methods are primarily applied to outcomes separately. To our knowledge, they have not been applied to the analysis of a multidimensional poverty index. This, in part, is due to the structure of the MPI itself: an index specifically developed to measure the prevalence and intensity of poverty. By estimating average treatment effects of particular statistics (i.e. headcount ratios, intensity, and the MPI, which is the combination of both), important information on poverty is already provided. These are particularly convenient to estimate because of the additively separable form of the index.[5] However, by combining these methods with the purpose-built structure of the MPI, gains can be made in the ability to detect the impact of interventions on the incidence and, importantly, distribution of the deprivation score.

[5]Unlike the Gini coefficient, for example, which often requires more complex RIF regressions to estimate (Firpo et al., 2009).

Therefore, as will be outlined in the methods section, in addition to estimating average treatment effects (on the treated) for measures of incidence, deprivation score and the RMPI, we will also go beyond these traditional evaluation measures. First, we estimate (unconditional) quantile treatment effects to identify the effect of the ESSN on the intensity of those at the top and bottom of the distribution (of intensity). Second, we estimate the entire probabilistic potential outcome distributions, allowing the effect of the ESSN on the entire distribution of intensity to be estimated. In doing this, we develop a 'Headcount Function', which is a modified distributional treatment effect, allowing the change in headcount at any cutoff threshold of poverty to be estimated.

2.3 The ESSN programme in Turkey

We apply the RMPI and impact evaluation methods to the Emergency Social Safety Net (ESSN) programme, which is the flagship humanitarian assistance programme of the European Union (originally budgeted at €6 billion for 2016–19). Launched in December 2016 in response to the influx of mostly Syrian refugees fleeing to Turkey[6] in the wake of the Syrian civil war that started in 2011, the ESSN is an unconditional multi-purpose cash transfer scheme that provides cash assistance via debit cards. It reached over 1.8 million refugees in Turkey, making it the largest humanitarian cash transfer programme in the world (Cuevas et al., 2019; Ark-Yıldırım and Smyrl, 2021a,b; Maunder et al., 2018). It is important to note that the majority of refugees in Turkey live amongst the host population (i.e. not in camps).

Forming part of the EU Facility for Refugees in Turkey (FRiT), which sets the framework for the EU contribution to the refugee response in Turkey, the ESSN is an integral part of a joint coordination mechanism that focuses on humanitarian assistance, education, migration management, health, municipal infrastructure and socio-economic support, implemented under the leadership of the Government of Turkey (Sida, 2018). Beneficiary households of the ESSN programme receive, as of 2021, 155 Turkish Lira (approximately €16) monthly per family member.[7] Beneficiary households are those who are *eligible* and who *applied* for the ESSN. Eligibility status is predominantly determined by six eligibility criteria: households with disabled family member(s) or high dependency ratios, households headed by a single

[6]With 4 million refugees, Turkey hosts the largest refugee population worldwide: 4 million, of which 3.6 million are Syrians. Only a minority, about 5% of the refugees, live in camps (Cuevas et al., 2019, 5).

[7]The amount was gradually increased from the initial transfer value of TRY 100, first to TRY 120 in 2017, where it remained until 2020. Depending on household size, beneficiary households also receive quarterly top-ups of up to TRY 600. Changes were made in part because Turkey underwent sharp macroeconomic changes during the study period that led to a reduction in the purchasing power of refugees (OPHI and WFP, 2022).

parent or elderly person, single females and large families are eligible to apply.[8] These eligibility criteria were selected as proxy measures of household poverty and vulnerability to meet basic needs (Cuevas et al., 2019).[9]

The programme is ongoing as of June 2021, but monitoring and evaluation (M&E) results from phases I and II (2016–20) have shown positive results on various outcomes related to the nine outcome indicators of the programme's M&E plan, including indebtedness, food consumption, schooling and livelihood-based coping strategies, inter alia (Maunder et al., 2018).[10] The 2018 mid-term evaluation found 'sufficient evidence to confirm both the needs and the appropriateness of using cash-based transfers to reduce the poverty and vulnerability of the refugees' (Maunder et al., 2018, p.v). Beneficiaries were found on average 'to be better off after the transfer, more food secure, had lower debt levels and were less likely to resort to negative coping strategies. In comparison the welfare of non-beneficiaries has declined according to most of the analysed measures of welfare' (Maunder et al., 2018, p.iv). Özler et al. (2021) found a moderate increase in the diversity and frequency of food consumption among eligible households; a decline in the poverty headcount of more than 50% after one year, using the monetary poverty line of $3.20/day; and meaningful gains in school enrolment among the most vulnerable beneficiary households. Aygün et al. (2021) found that the ESSN had a large effect on school enrolment.

While these studies monitored and evaluated the outcomes of the programme on individual indicators by focusing on monetary poverty, the World Food Programme in Turkey (WFP VAM) analysed poverty profiles of ESSN beneficiaries across multiple dimensions of poverty. In 2019, WFP VAM designed a multidimensional poverty index based on data collected in wave 3 (March 2018) of its Comprehensive Vulnerability Monitoring Exercise (CVME), the CVME MPI. Multidimensional poverty among the ESSN beneficiaries was found to be 54% based on data from March 2018 (WFP, 2019a, p.9). When reapplied to data from wave 4 (September 2018) and wave 5 (October 2019), this changed to 38% and 43%, respectively (**?**, p.31). While a valuable attempt in understanding multidimensional poverty among refugees

[8]Applicants to the programme need to be households with: (a) at least one disabled family member with a 40% minimum disability threshold; (b) a single female aged 18 or above; (c) a single parent family, comprising one adult 18 to 59 and at least one child; (d) four or more children; (e) an elderly headed family, i.e. one or more adults aged 60 or over with no adults aged 18 to 59; or (f) a household dependency ratio above or equal to 1.5 (OPHI and WFP, 2022; Cuevas et al., 2019).

[9]An important ineligibility criterion of the programme was for those households with any member enrolled in social security (a result of formal employment) (Cuevas et al., 2019, p.28).

[10]The Theory of Change identified immediate outcomes covering food consumption and dietary diversity, food and livelihood coping strategies, per-capita expenditure on the minimum expenditure basket, reduced debt, improved enrolment and attendance in formal education, and aid effectiveness (Maunder et al., 2018).

in Turkey, the CVME MPI suffers from several methodological weaknesses (OPHI and WFP, 2022).[11]

We build upon this to better understand the impact of the ESSN on the multidimensional poverty of refugees, moving beyond single indicator analyses by using an improved Refugee MPI that was developed specifically to this end. In applying the novel impact evaluation methods developed in the paper, we will add a robust causal analysis of the effectiveness of the ESSN on multidimensional poverty.

3 Methods

3.1 Refugee Multidimensional Poverty Index

We use the Refugee Multidimensional Poverty Index (RMPI), an index specifically designed to reflect the needs and rights of refugees in Turkey. The RMPI was developed in collaboration between OPHI and the WFP (OPHI and WFP, 2022), is based on the Alkire–Foster (AF) method (Alkire and Foster, 2011), and relies on the data available in the CVME survey.

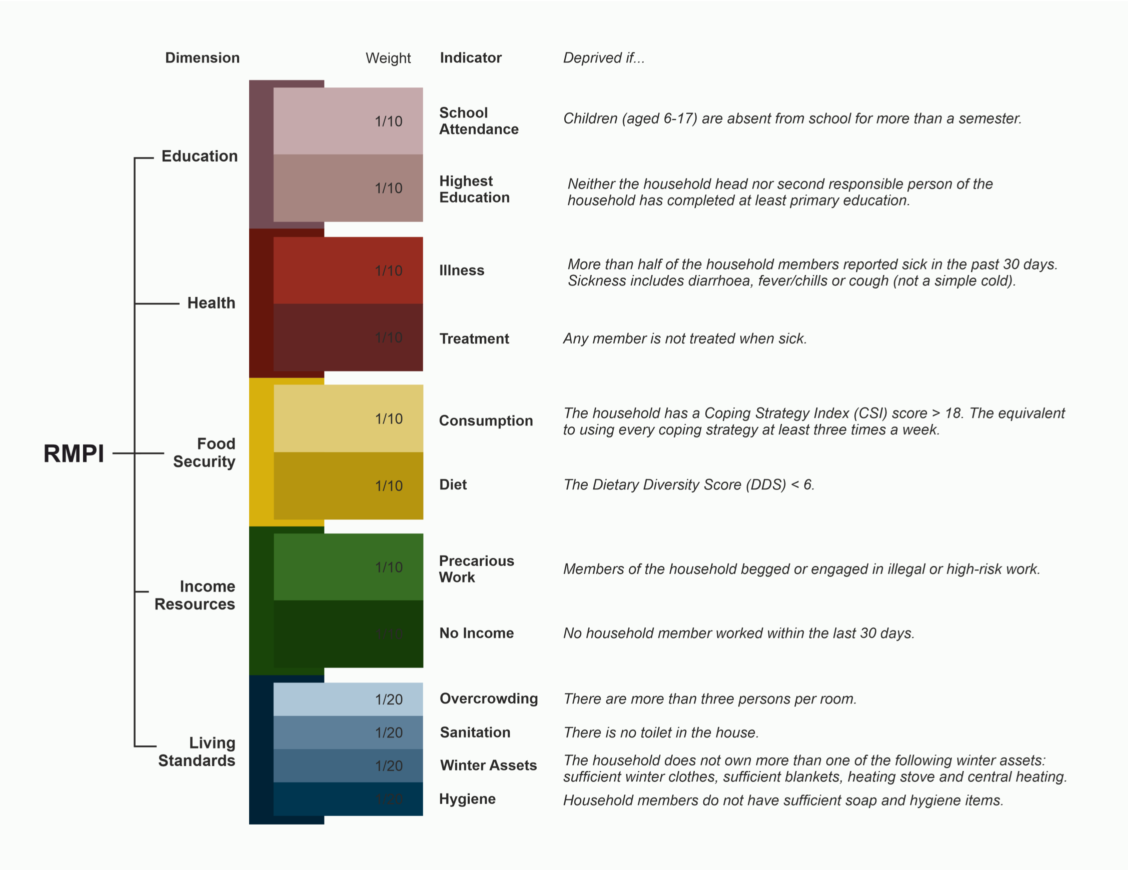
Figure 1 summarises the dimensions, indicators and weights of the RMPI. Poverty is measured in twelve indicators across five dimensions, namely education, health, food security, income resources and living standards, which are weighted with a nested weighting structure (where dimensions, and indicators within dimensions, are weighted equally).[12] The RMPI is formulated in the next section, and the process of developing the RMPI is described in the section following.

3.1.1 Formulation

The RMPI is based on the Alkire–Foster (AF) method, a flexible approach that allows poverty to be measured across multiple dimensions and summarised in one index. The method requires the selection of relevant dimensions, indicators and weights; the setting of a poverty cutoff to identify who is poor; and, finally, the computation of the summary measures: the headcount ratio, the intensity of poverty and the RMPI.

[11] For example, the index used statistical weights that depend on the eigen decomposition of the corresponding survey(s). Although the CVME MPI was reapplied to waves 4 and 5 with the same statistical weights as first published in 2019, the weights are relative to the eigen decomposition of wave 3. Thus, even if anchored on relative weights, absolute comparisons across waves are not possible.
[12] For a table presentation, see Appendix A.2.

Figure 1: Refugee Multidimensional Poverty Index Structure



Formally, let i denote a person in a population of n, and let j denote an indicator of poverty in the set of d indicators, so that $i \in [1, \ldots, n]$ and $j \in [1, \ldots, d]$. Then let g^0 be a deprivation matrix where $g^0_{ij} = 1$ if person i is deprived in indicator j and $g^0_{ij} = 0$ otherwise. Each indicator j is assigned a weight denoted by w_j such that $w_j > 0, \forall j \in [1, \ldots, d]$, and $\sum_j^n w_j = 1$. The *deprivation score* of each person i is, then, denoted as c_i, which is the sum of their weighted deprivations across all indicators:

$$c_i = \sum_{j=1}^{d} w_j g^0_{ij} \tag{1}$$

A poverty cutoff, k, is then used to identify the multidimensionally poor. The poverty cutoff is the minimum deprivation score a person needs to exhibit to be identified as poor. This poverty cutoff is implemented using an identification function $\rho_i(k)$ which depends upon each person's deprivation score, c_i, and the poverty cutoff, k, and takes the value of 1 if the person is poor, that is if $c_i \geq k$, and 0 otherwise:

$$\rho_i(k) = I(c_i \geq k) \tag{2}$$

The AF measure can, then, be expressed in terms of two partial indices. The first partial index H is the percentage of population that is poor, also called the multidimensional headcount ratio. The second is the intensity of poverty, A, which is the average deprivation score across the poor. Thus, the RMPI is the breadth-adjusted headcount ratio.

$$H = \sum_i^n \frac{\rho_i(k)}{n}, \qquad A = \sum_i^n \frac{\rho_i(k)c_i}{Hn}, \qquad RMPI = H \times A \tag{3}$$

3.1.2 Development of the RMPI

Following a review of existing measures of multidimensional poverty of refugees, the Refugee Multidimensional Poverty Index (RMPI) was developed through a collaborative process of normative discussion and statistical analysis, as well as the computation of five alternative trial measures with different indicator specifications, dimensional structures and weights. These were formulated, statistically assessed and extensively debated in various consultation rounds between OPHI and the WFP in 2021 (see Appendix A.1 and OPHI and WFP (2022)).[13] Normative discussions surrounding these trial measures aimed to ensure that the final RMPI is comprehensively informed by the needs of refugees, such as health, nutrition and shelter (WFP, 2020b), but also by their rights.[14] In particular, the right to education, both for children and adults, was considered in the indicator selection for the education dimensions, and the right to work is reflected in the income resource dimension. Where possible, refugees' own voices were accounted for as captured through thematic focus group discussions (e.g. on shelter, see WFP, 2020c). Statistical analysis also ensured that the measures were based on a stringent application of the AF-method, resulting in an RMPI that is parsimonious, inclusive of different reference populations in each dimension, and robust to reasonable alternatives in key selection criteria of the index, such as setting the poverty cutoff line to 20%. The RMPI thus moves towards a better approximation of the dimensions and indicators that are specific to the needs and rights of refugees.

[13] These indices were subject to the data possibilities and constraints of the CVME dataset used for our analysis (see Appendix A.1).
[14] Based on the 1951 Refugee Convention and its 1967 Protocol, as well as the 2006 Refugee Act (UNHCR, 2017).

3.2 Treatment Effects Estimation

We estimate causal treatment effects using several approaches which rely on inverse-probability-weighting regression adjustments (IPWRA). First, average treatment effects on the treated (ATETs) are estimated using standard IPWRA methods. Then, unconditional quantile treatment effects (UQTEs) are estimated across the quantiles of the deprivation score, using methods developed by Firpo et al. (2009). Finally, we develop a method, using the beta-binomial distribution, to estimate marginal distributions of deprivations in each dimension of poverty, which jointly provides an estimate of the entire distribution of the deprivation score.

By using the RMPI, the estimated ATETs can provide summaries of the average effects of an intervention on the incidence and intensity of multidimensional poverty, alongside separate effects on each dimension (or indicator). The UQTE and beta-binomial approaches provide further information on the distributional consequences of an intervention. The UQTE uncovers heterogeneity in effects for those at different parts of the distribution (i.e. those with higher or lower (ranked) deprivation scores). The beta-binomial method allows for the estimation of probabilistic potential outcomes and treatment effects on the entire distribution of the deprivation score. This allows for the identification of changes to the incidence and intensity of poverty, at any possible cutoff threshold. The framework and methods are described below.

3.2.1 Deterministic and Probabilistic Potential Outcomes

We use the potential outcomes framework; however, we extend this framework beyond its typical *deterministic* conceptualisation to focus on *probabilistic* potential outcomes. This is a shift towards modelling the entire distribution of the outcome variable, rather than expectations; which is particularly important when analysing poverty. Moreover, our focus is on a *multidimensional* poverty index, rather than a *unidimensional* outcome variable.

In the standard framework, for a population of individuals $i \in N$ we observe actual outcomes y_i. A treatment dummy D_i identifies whether i receives the treatment ($D_i = 1$) or not ($D_i = 0$). The *potential* outcome, $Y_i(.)$, is that outcome an individual would have had if they were treated, $Y_i(1)$, or not, $Y_i(0)$. The outcome we observe is, therefore:

$$y_i = Y_i(1)D_i + Y_i(0)(1 - D_i) \tag{4}$$

The effect of the treatment on an individual would be $\tau_i = Y_i(1) - Y_i(0)$. The average treatment effect, $\tau_A = E[Y_i(1) - Y_i(0)]$, provides a summary measure of the expected effectiveness of the treatment. This is the focus of much applied research.[15]

One issue with these approaches is the focus upon *expectations* and the subsequent relegation of differences from these expectations to 'noise'. In the context of poverty measurement, this 'noise' is actually the individuals and households who are in the most extreme depths of poverty and often those whom interventions are targeting.

To shift focus away from expectations, we can, instead, focus upon the entire distribution, by modelling probabilistic potential outcomes and identifying treatment effects at different parts of the distribution. The probabilistic potential outcomes are:

$$f(y_i) = f(Y_i(1))D_i + f(Y_i(0))(1 - D_i) \tag{5}$$

where $f(.)$ is a probability mass function. This identifies the probability, $p_i(y)$, of observing each level of the outcome y. The probabilistic distribution of potential outcomes an individual would have with treatment is $f(Y_i(1))$, and that without treatment is $f(Y_i(0))$. The cumulative mass functions, $F(y)$, can also be identified using $f(y)$, allowing the identification of the probability that $Y \leq y$. Using these functions, many summary statistics can be identified which allow for the identification of the expected treatment effects, quantile treatment effects and distributional treatment effects.

Practically, we cannot observe the entire distribution of an outcome for an individual. Therefore, we focus on estimating conditional probabilistic potential outcomes using observable characteristics of interest $\mathbf{X}$. This allows an approximation of individual level potential outcomes. To model the distribution of the outcome variable across the sample, we aggregate the probability mass functions, estimated for all individuals, to identify the entire distribution of potential outcomes and treatment effects.

3.2.2 Inverse-probability-weighted regression adjustment

Inverse-probability-weighted regression adjustment (IPWRA) is a doubly-robust method which first, estimates propensity weights for the probability of treatment; second, performs a weighted

[15]The fundamental problem of causal inference – we cannot simultaneously observe an individual with and without treatment – means that we cannot observe individual treatment effects directly. Therefore, expected *counterfactual* potential outcomes need to be predicted using experimental (e.g. randomised control trials), quasi-experimental (e.g. difference-in-differences, regression discontinuity) and matching approaches (e.g. nearest neighbour, propensity-score methods).

least-squares regression of the outcome on the treatment dummy, the controls and their interaction, using the inverse of the estimated weights; and third, uses the predicted outcomes to estimate statistics of interest (e.g. Average Treatment Effects).

The propensity score is denoted by:

$$p(x) = P(D = 1|X = x) \tag{6}$$

This is the probability of receiving treatment given the characteristics x. This propensity score is estimated using a logit regression. Once estimated, propensity scores are used to define weights:

$$\hat{\omega}(x) = \frac{D}{\hat{p}(x)} + \frac{1-D}{1-\hat{p}(x)} \tag{7}$$

Weights for the treated are the inverse of the propensity weights, whilst the weights for the control group are the inverse of $1-\hat{p}(x)$. These inverse-probability weights are used across all analysis conducted to enable the identification of causal effects.

For the IPWRA, a weighted regression is run:

$$\sqrt{\hat{\omega}(x)}y = \sqrt{\hat{\omega}(x)}\left(\beta_0 + D_i\beta_\tau + \sum_j x_{ij}\beta_j + \sum_j D_i x_{ij}\beta_{\tau j} + u_i\right) \tag{8}$$

Estimated coefficients are used to predict (deterministic) potential outcomes that each individual is expected to have with and without treatment, and hence enables the predictions of treatment effects. The weighted average of these predicted treatment effects, for the treated group, gives the Average Treatment on the Treated (ATET).[16] The identification of the ATET relies on two crucial assumptions: conditional independence and common support (Rosenbaum and Rubin, 1983).[17]

In addition, we go beyond estimating average treatment effects. By running unconditional quantile regressions (Firpo et al., 2009), with inverse-probability-weights and control variables, Unconditional Quantile Partial Effects of the treatment (Unconditional Quantile Treat-

[16]The `teffects ipwra` command, in Stata, is used for estimation.

[17]The conditional independence assumption (CIA) can be formally expressed as $Y_i(0), Y_i(1) \perp D_i|X_i$. CIA indicates that potential outcomes are independent of treatment, conditional on the observable characteristics X_i. This is satisfied if there are no unobservable variables which contribute to hidden bias in our estimates. Common support requires that $0 < P(D_i = 1|X_i) < 1$, i.e. that every individual has a positive probability of being treated. This is further discussed in Section 3.5.

ment Effects) can be estimated.[18] This allows for the effects of the treatment on the (unconditional) quantiles of the deprivation score to be estimated, and hence uncovers whether effects are greater for the most or least deprived part of the distribution.

3.3 Beta-Binomial Distribution

Probabilistic potential outcomes can be modelled by using an appropriate probability distribution. The distribution chosen here is a beta-binomial. This is a family of discrete probability distributions[19] with flexible shape parameters, which can be adapted to model the distribution of weighted deprivations. Although we apply this to the RMPI, it can easily be applied to any evenly-spaced, discrete outcome.

We begin formulating the beta-binomial by defining T. Typically, this is the number of trials; here, it is the number of discrete values that the outcome can take. For the deprivation score of the RMPI, $T = 20$, as the nested weighted structure allows for values in evenly spaced increments of 0.05, or $1/T$. For the weighted deprivations in living standards, $T = 5$, as with four indicators, values in increments of 0.2 are possible.

Each discrete value is indexed by t, with $t \in 0, ..., T$. We normalise the distribution, focusing on $t/T = y$ as our measure of a discrete outcome, bounded between 0 and 1, in increments of $1/T$. Therefore, we model the probability mass function of our outcome as:

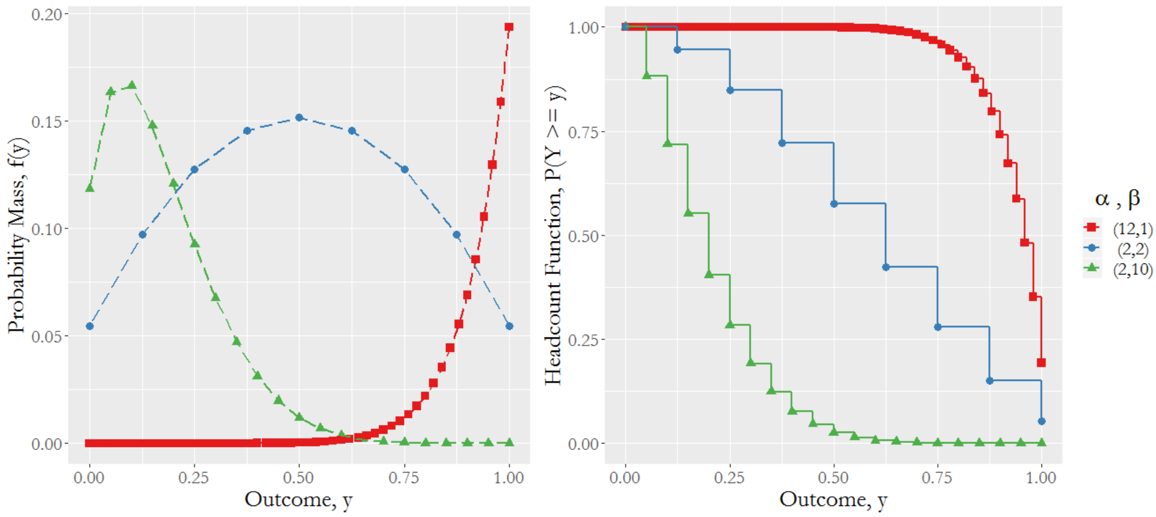
$$f\left(\frac{t}{T}\middle| T, \alpha, \beta\right) = \frac{\Gamma(T+1)}{\Gamma(t+1)\Gamma(T-t+1)} \frac{\Gamma(t+\alpha)\Gamma(T-t+\beta)}{\Gamma(T+\alpha+\beta)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \tag{9}$$

This formulation allows for the probability of each discrete observation $y = t/T$ to be modelled for any T number of discrete values, using the flexible shape parameters α and β. Figure 2 illustrates this flexibility, both in T and the shape parameters.[20]

The right panel of Figure 2 shows the 'Headcount Function', which shows the probability of observing an outcome equal to or greater than a value of y, $P(Y \geq y)$.[21] It identifies the

[18]The `rifhdreg` command, in Stata, is used for estimation (Rios-Avila, 2020).
[19]Typically used to identify the probability of success over a number of Bernoulli trials.
[20]The left panel of Figure 2 shows the probability mass function, $f(y) = P(Y = y)$. The red curve shows a heavily left-skewed distribution, with $(\alpha, \beta) = (12, 1)$ and a high number of possible discrete values, $T = 50$. The blue curve shows a symmetric distribution, centred on 0.5, with $(\alpha, \beta) = (2, 2)$ and $T = 8$. The green curve shows a right-skewed distribution, with a modal probability at 0.1, and $T = 20$.
[21]The concept is similar to the cumulative density function, which shows the probability of observing an outcome below or equal to a value of y, $F(y) = P(Y \leq y)$. This is particularly useful in the context of poverty measurement.

Figure 2: Beta-Binomial Distribution



proportion of the population classified as poor at each possible poverty cutoff (k).

It is the shape parameters, α and β, that we will estimate to model the distribution of the outcome variables. These will be estimated as additive functions of estimated coefficients α_j and β_j, the observable characteristics x_{ij}, the treatment dummy D_i and their interaction $D_i x_{ij}$, as:[22]

$$\alpha_i = \exp\left(\alpha_0 + D_i\alpha_\tau + \sum_j x_{ij}\alpha_j + \sum_j D_i x_{ij}\alpha_{\tau j}\right) \tag{10}$$

$$\beta_i = \exp\left(\beta_0 + D_i\beta_\tau + \sum_j x_{ij}\beta_j + \sum_j D_i x_{ij}\beta_{\tau j}\right) \tag{11}$$

These parameters will estimated using weighted maximum-likelihood estimation,[23] where:

$$LL = \sum_i \omega_i \ln(f(y_i|T, \alpha_i, \beta_i)) \tag{12}$$

Once estimated, these shape parameters allow probabilistic potential outcomes to be estimated at the individual level. This allows for important statistics of interest to be estimated, including the expectations of the outcome and the probabilities of the deprivation score. In our main specification, we estimate the marginal distributions of each dimension of poverty

[22]As α_i and β_i are strictly positive, the exponential function is a natural function to use.
[23]The weights are the estimated inverse probability of treatment weights, using control variables to balance control and treated groups.

separately. We then estimate the joint distribution of these marginals to identify the probabilistic potential outcomes of the deprivation score, for each individual.[24]

3.4 Data

The analysis uses data collected through the Comprehensive Vulnerability Monitoring Exercise (CVME). The CVME is a cross-sectional survey, collected in five waves, and provides in-depth information on the determinants of refugee vulnerability in Turkey. We pool data from waves 3–5; this data was gathered between March 2018 and February 2020 and was collected using a two-staged sampling approach that allows the sample to be representative for refugees across Turkey.[25] Sampling weights were then calculated such that the final sample is representative at the national level.[26]

Descriptive statistics for the data are shown in Table 1 at the sample level and split into subgroups according to ESSN application and eligibility status. Statistics for the household composition, household head's characteristics, arrival time in Turkey and region are shown, alongside the ESSN eligibility criteria. The total sample size is 4,106, with 2,230 eligible applicants, 1,216 ineligible applicants and 545 non-applicants.[27]

There are several notable differences between the subgroups. Eligible applicants have higher household size, number of children and dependency ratios than ineligible applicants and non-applicants. The household head is more likely to be female and the household is more likely to have arrived earlier (3–6 years ago). As expected, eligible applicants are more likely to fit each of the ESSN criteria: 80% have a dependency ratio ≥ 1.5 and 55% have four or more children. Ineligible applicants tend to have older household heads, household heads with a skilled or highly skilled profession, and are less likely to come from the south-east (near the border with Syria). Non-applicant household heads are less likely to be female and more likely to be Afghan, and almost a quarter of non-applicant households have arrived within the last year, compared to less than 5% of applicant households.

[24] In Appendix A.9, we also show results where we directly estimate the deprivation score.

[25] First, to achieve spatial representation, GPS points were selected based on the density of the refugee population residing in Turkey. In a second step, respondent-driven sampling was used to select 25 households for each GPS point (WFP, 2019b).

[26] For the first two rounds of the CVME, the sampling was not representative, and for most of the data collection (except for ESSN beneficiaries), enumerators used a snowball method. This implies that there are no weights for the first rounds to make them representative, due to the sampling framework used (WFP, 2019b).

[27] In addition, there are 115 households with their ESSN status 'Pending'.

Table 1: Descriptive Statistics by ESSN Eligibility Status

	Sample		Applicants				Non-Applicants	
			Eligible		Ineligible			
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Household								
Household Size	5.64	(2.54)	6.37	(2.55)	4.89	(1.92)	4.45	(2.46)
Number of Children	3.09	(2.00)	3.86	(1.86)	2.04	(1.31)	2.13	(1.83)
Dependency Ratio	1.70	(1.33)	2.10	(1.37)	1.19	(1.13)	1.12	(0.84)
Household Head								
Female Headed	0.17	(0.37)	0.21	(0.41)	0.14	(0.35)	0.05	(0.22)
Age (Years)	40.13	(12.39)	40.33	(11.13)	41.06	(13.85)	38.56	(14.10)
Syrian	0.87	(0.34)	0.87	(0.33)	0.87	(0.34)	0.83	(0.37)
Afghan	0.05	(0.21)	0.04	(0.19)	0.04	(0.19)	0.10	(0.30)
Iraqi	0.08	(0.27)	0.09	(0.28)	0.09	(0.29)	0.06	(0.24)
Highly Skilled/Skilled	0.16	(0.37)	0.13	(0.33)	0.25	(0.43)	0.13	(0.34)
Arrival Time								
<12 months	0.07	(0.26)	0.05	(0.22)	0.03	(0.16)	0.23	(0.42)
1-3 years	0.22	(0.41)	0.18	(0.39)	0.22	(0.41)	0.34	(0.47)
3-6 years	0.59	(0.49)	0.66	(0.47)	0.58	(0.49)	0.41	(0.49)
Before conflict	0.12	(0.32)	0.10	(0.31)	0.18	(0.38)	0.02	(0.15)
Region								
Istanbul	0.06	(0.23)	0.05	(0.21)	0.08	(0.28)	0.06	(0.23)
Aegean	0.10	(0.30)	0.10	(0.30)	0.12	(0.33)	0.07	(0.26)
Mediterranean	0.16	(0.37)	0.17	(0.38)	0.15	(0.35)	0.17	(0.38)
Anatolia	0.32	(0.47)	0.30	(0.46)	0.37	(0.48)	0.33	(0.47)
South-east	0.36	(0.48)	0.38	(0.49)	0.28	(0.45)	0.36	(0.48)
Eligibility Criteria								
Dependency Ratio $\geq$ 1.5	0.57	(0.49)	0.80	(0.40)	0.29	(0.45)	0.32	(0.47)
Single Female	0.02	(0.14)	0.03	(0.18)	0.00	(0.07)	0.00	(0.06)
Elderly Headed	0.01	(0.11)	0.01	(0.12)	0.01	(0.11)	0.01	(0.08)
Single Parent	0.06	(0.24)	0.08	(0.27)	0.01	(0.11)	0.04	(0.21)
Disabled Member	0.12	(0.32)	0.11	(0.31)	0.17	(0.38)	0.04	(0.20)
Num. Children $\geq$ 4	0.36	(0.48)	0.55	(0.50)	0.12	(0.32)	0.16	(0.37)
N	4106		2230		1216		545	

Note: Descriptive statistics shown by recorded ESSN eligibility status. Sample weights are used throughout. All variables, bar household size, number of children, dependency ratio and age of household head, are dummy variables, with means denoting proportions.

3.5 Treatment Groups and Controls

Of these groups, it is the eligible applicants who receive the ESSN, and hence are the treated group; together, the ineligible applicants, non-applicants and pending households are the control group. When estimating treatment effects, the treatment and control groups are balanced across observable characteristics, using inverse probability of treatment weights, to reduce selection bias and enable the estimation of causal treatment effects.

To estimate propensity weights, observable characteristics used as controls are:

- *Household*: region, arrival date, number of children, adults and elderly;
- *Household head*: nationality, age, gender, profession;

- *ESSN eligibility criteria*: dependency ratio, single female, elderly headed, single parent, disabled members, large family.

The inclusion of ESSN eligibility criteria as controls alongside demographic characteristics is essential as this ensures that the control and treated groups are balanced in terms of their vulnerability. This extensive list of characteristics, combined with the inclusion of the ineligible applicant, non-applicant and pending households in the control group, reduces the concerns of confounding caused by unobservables (CIA) and lack of overlap (common support). More detail on the inverse probability weighting, balance and overlap are found in Appendix A.6.

Within the inverse-probability-weighted regression adjustment, these controls are again used as regressors within the regression adjustment. For the beta-binomial estimation, however, a more parsimonious set of controls are used, with selected dummy variables within each of the categories above.

4 Results

4.1 Refugee MPI

Table 2 presents the headline figures of the Refugee MPI (RMPI), across all refugees. The incidence or headcount ratio of multidimensional poverty (H) describes the proportion of people who are identified as poor. The intensity (A) describes the average proportion of (weighted) deprivations faced by the people identified as multidimensionally poor. The MPI is the product of the headcount ratio (H) and the average intensity (A).

Table 2: RMPI, Headcount Ratio (H) and Intensity (A)

RMPI (95% CI)	**Headcount Ratio (H)** (95% CI)	**Intensity (A)** (95% CI)
0.143 (0.126 , 0.160)	**50.0%** (44.4% , 55.6%)	**28.7%** (27% , 30.3%)

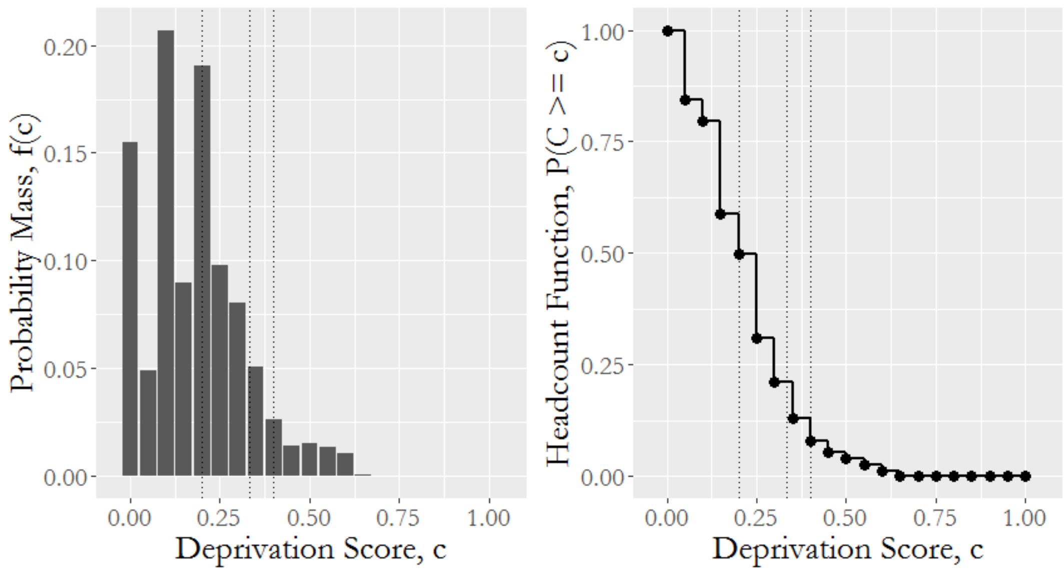
As can be seen in Table 2, the incidence of multidimensional poverty is 50%. In other words, half of the refugee population is identified as multidimensionally poor. The average intensity of poverty is 28.7%. Thus, multidimensionally poor people face, on average, deprivations in nearly one-third of the dimensions included in the RMPI. Finally, the RMPI has a value of 0.143. This means that multidimensionally poor refugees in Turkey experience 14.3% of the total deprivations that would be experienced if everyone was fully deprived in all indicators.

Maps providing further information on geographical variation in the RMPI, incidence and intensity across Turkey are shown in Appendix A.4, alongside a regional summary table.

4.1.1 Distribution of Incidence and Deprivation Scores

Figure 3 provides more detail on the distribution of the deprivation score that refugees in Turkey face. The left panel shows the empirical distribution of deprivation scores, where the proportion of refugees who face each deprivation score is shown by the height of the bar. At the extremes, we observe that 15.5% of the population are deprived in no indicators ($c = 0$), and no households are deprived in more than 65% of indicators ($c > 0.65$).

Figure 3: Distribution of Deprivation Scores and the Headcount Function



The right panel of Figure 3 shows the headcount function. This shows the proportion of refugees classed as poor by any cutoff threshold, k, at any deprivation score, c. We observe that 50% of refugees are classified as poor at a threshold of 0.2 (similarly shown in Table 2). If higher thresholds were chosen, 0.333 or 0.4, for example, we can see how the poverty headcounts would change to 21.2% and 8% of refugees, respectively.[28]

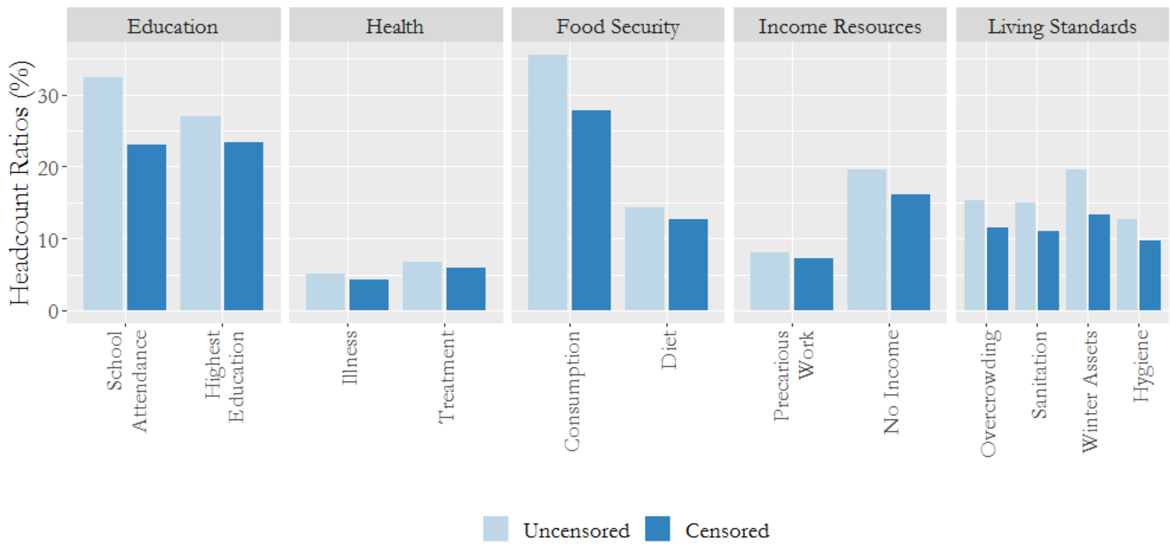
[28]This is an expected result of the dual-cutoff approach of the AF method, as with higher (or more severe) poverty cutoffs, the headcount ratio will fall (Alkire and Foster, 2011; Alkire et al., 2015). Note that robustness

The distribution of the deprivation scores in the left panel are particularly informative for the poor. By censoring the bars to the left of 0.2, we observe a right-skewed distribution, where the majority of those classified as poor have lower levels of intensity. Of those classified as poor, 38% have a deprivation score of 0.2, 19.6% have a score of 0.25 and 16.2% a score of 0.3. In other words, the majority of those classified as poor have intensity levels close to the poverty cutoff of 20%.

4.1.2 Uncensored and Censored Headcount Ratios

Figure 4 shows results of the uncensored headcount ratios (the percentage of the population that is deprived in each indicator) and censored headcount ratios (the percentage of the population that is poor and deprived in each indicator). We see that the education indicators and the consumption indicator have the highest uncensored and censored headcount ratios, whilst the health indicators have the lowest. These results are also reflected in the percentage contributions that are presented in Appendix A.3.

Figure 4: RMPI uncensored and censored headcount ratios (%)



test results on the poverty cutoff choice were also conducted (OPHI and WFP, 2022, pp.14–15) and showed that poverty rankings are robust to reasonable changes in the k value (ranging from 10% to 50%) for a population share of 98.9% of the CVME3 when results were disaggregated by arrival time. In the context of refugees, different arrival times help to understand why certain groups of refugees are poorer than others, so having established robust results by this aggregation adds confidence in the use of the identification function of the RMPI.

4.1.3 Decomposing the RMPI by ESSN Status

Table 3 presents the disaggregated results of the RMPI by ESSN status. Non-applicants, and those with pending applications, are the poorest groups, with RMPI values of 0.198 and 0.221, respectively, compared to 0.138 for beneficiaries and 0.110 for ineligible populations.

Table 3: RMPI, Headcount Ratio and Average Intensity of Poverty, by ESSN Status

ESSN Status	Population Share (Percentage)	Sample size (Frequency)	RMPI (95% CI)	H (95% CI)	A (95% CI)
Beneficiary	54.3%	2,230	0.138	49.3%	28%
			(0.117, 0.159)	(42.4%, 56.2%)	(26.0%, 30.0%)
Ineligible	29.6%	1,216	0.110	41.1%	26.8%
			(0.084, 0.136)	(31.1%, 51.1%)	(24.6%, 29.0%)
Non-applicant	13.3%	545	0.198	60.7%	32.6%
			(0.152, 0.244)	(47.1%, 74.4%)	(26.0%, 39.2%)
Pending	2.8%	115	0.221	78.2%	28.2%
			(0.171, 0.270)	(57.8%, 98.6%)	(24.7%, 31.8%)

While these results show significant differences between individuals grouped by ESSN status, these differences are not necessarily *caused* by the ESSN. These groups are fundamentally different in their composition (see Table 1), as there is selection into these groups. The next section, therefore, uses inverse-probability weights to address this selection to uncover the causal effects of the ESSN on the eligible applicants.

4.2 Treatment Effects on the Treated

In the following sections, we estimate the impact of the ESSN on the treated population: the eligible applicants. We estimate average treatment effects on the treated (ATET) for headcount ratios, deprivation scores, RMPI and weighted dimensions, using an inverse-probability-weighted regression adjustment. Further, we estimate unconditional quantile treatment effects (UQTEs) across the quantiles of the deprivation score. Finally, we estimate the entire probabilistic potential outcome distribution of the deprivation score, revealing the impact of the ESSN on the entire distribution of the deprivation score and headcount function.

4.2.1 Average Treatment on the Treated

Table 4 shows the effects of the ESSN programme on RMPI headcounts, deprivation scores and the RMPI, using standard inverse-probability-weighted regression-adjustment methods.[29]

Table 4: RMPI: Average Treatment on the Treated

	(1) **H: 20%** Coef./S.E.	(2) **H: 33.3%** Coef./S.E.	(3) **H: 40%** Coef./S.E.	(4) **c_i** Coef./S.E.	(5) **RMPI** Coef./S.E.
Treatment Effect	-0.1833***	-0.1052**	-0.0980**	-0.0577***	-0.0729***
	(0.0437)	(0.0433)	(0.0440)	(0.0139)	(0.0155)
PO: No Treatment	0.6689***	0.2147***	0.1730***	0.2328***	0.2090***
	(0.0351)	(0.0412)	(0.0430)	(0.0147)	(0.0163)
N	4042	4042	4042	4042	4042

Note: Inverse-probability-weighted regression adjustments are used to estimate coefficients, the Average Treatment on the Treated and the expected potential outcome without treatment. Models (1), (2) and (3) show headcount ratios, with different cutoff thresholds, k; (4) shows the deprivation score, c_i; and (5) shows the RMPI ($k = 20\%$). p-values: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The first three models show the effect of the ESSN programme on the headcount ratio at the 20% threshold of the RMPI and at more *severe* thresholds of 33.3% and 40%. The headcount ratio that the eligible applicant group would have had without the ESSN programme is 0.669 (i.e. 67% of eligible applicants would have been in poverty) at the 20% threshold, 0.215 at the 33.3% threshold and 0.173 at the 40% threshold. The ESSN is found to significantly reduce the incidence of poverty, at each level. At the 20% level, there is a significant reduction of 0.183, reducing the headcount ratio to 0.486. At the 33% level the reduction is 0.105 and at the 40% level it is 0.098.

The expected level of the deprivation score, c_i, that the eligible applicants' group would have had without the ESSN programme is 0.233. The causal effect of the ESSN programme on the deprivation score is a significant reduction of 0.058, down to 0.175. The RMPI ($k = 20\%$) is also significantly reduced, from 0.209 to 0.136. Together these results show that the ESSN significantly reduced the incidence and intensity of poverty amongst eligible applicants.

Table 5 shows the effect of the ESSN programme on the dimensions of the RMPI, separately. These are weighted, so that a household deprived in all indicators in a particular dimension would score 1, and a household with no deprivations, 0. Results show that if the eligible applicants had not received the ESSN assistance, deprivation in the dimension of food secu-

[29]Appendix A.6 provides more detail on the balance of the control and treatment groups before and after the inverse-probability weights are applied and the degree of overlap. After applying the weights, there are no significant differences in the observed characteristics and there is a high degree of overlap.

rity (0.375) would have been be the greatest, followed by education (0.366), living standards (0.235) and the ability to generate income (0.116). They would have been least deprived in health (0.072). The ESSN programme caused significant reductions in deprivation in food security (-0.156), living standards (-0.078) and education (-0.063), but there were no significant effects on health or the ability to generate income.

Table 5: RMPI Weighted Deprivations: Uncensored

	(1) **Education** Coef./S.E.	(2) **Health** Coef./S.E.	(3) **Food** Coef./S.E.	(4) **Income** Coef./S.E.	(5) **Living** Coef./S.E.
Treatment Effect	-0.0630^{*}	-0.0134	-0.1555^{***}	0.0215	-0.0782^{***}
	(0.0354)	(0.0190)	(0.0343)	(0.0200)	(0.0283)
PO: No Treatment	0.3657^{***}	0.0721^{***}	0.3753^{***}	0.1160^{***}	0.2348^{***}
	(0.0312)	(0.0177)	(0.0322)	(0.0196)	(0.0236)
N	4042	4042	4042	4042	4042

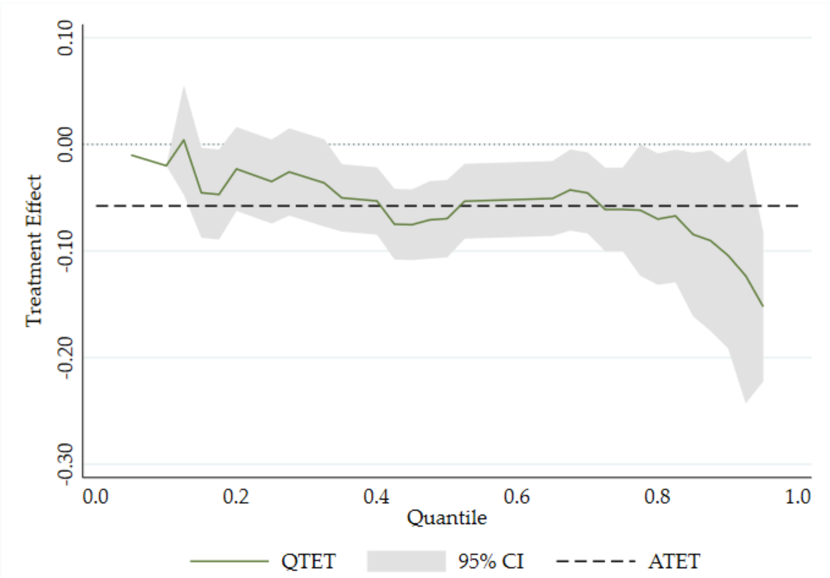
Note: Inverse-probability-weighted regression adjustments are used to estimate coefficients, the Average Treatment on the Treated and the expected potential outcome without treatment. Models show the weighted deprivations in each of the five dimensions, bounded between 0 (deprived in no indicators in that dimension) and 1 (deprived in all indicators in that dimension). p-values: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Further analysis on individual indicators is provided in Appendix A.7. Results show that the school attendance indicator drives the effects on education, both the Coping Strategy Index (CSI) and the Dietary Diversity Score (DDS) indicators are reduced for food security, and deprivations in winter assets are significantly reduced for living standards. We also find that, while there are no significant effects on the health and income resources dimensions, there is a significant reduction in those not receiving treatment, but an increase in those not working. Additional analysis in Appendix A.8 also shows conditional average treatment effects of the ESSN on subgroups defined by: arrival time; sex, age and occupation of the household head; dependency criteria and eligibility criteria. Households who arrived more recently, were female-headed and met more simultaneous eligibility criteria were found to have higher expected levels of poverty without treatment, and the ESSN had greater effects on reducing poverty in households with skilled household heads and meeting more simultaneous eligibility criteria.

4.2.2 Unconditional Quantile Treatment Effects

Figure 5 plots the Unconditional Quantile Treatment Effects of the ESSN on the deprivation score. Results show significant treatment effects which increase from the lower to the higher quantiles of the deprivation score. At the 10th percentile, the effect size is -0.020; this increases to -0.069 at the 50th percentile and is -0.123 by the 90th percentile. At these percentiles, the sample mean of the deprivation score is 0.061, 0.219 and 0.423, respectively. This shows that the ESSN reduced deprivation scores to a greater extent for the more deprived.

Figure 5: Unconditional Quantile Treatment Effects: Deprivation Score



Note: Unconditional Quantile Treatment Effects on the Treated (QTET) are shown by the green line, with 95% confidence intervals, with robust standard errors. These show the treatment effect on the treated of the ESSN at different (unconditional) quantiles. The dotted line shows the Average Treatment on the Treated (ATET) for reference.

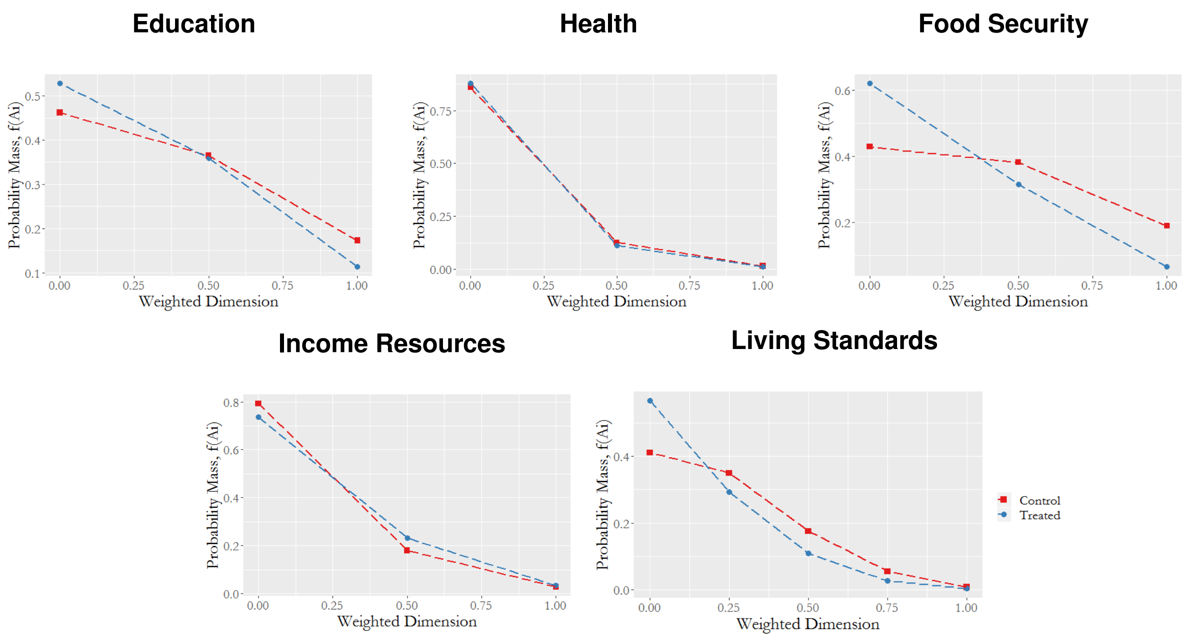
4.2.3 Probabilistic Potential Outcomes

This section moves to the estimation of the entire probabilistic potential outcome distribution. This reveals the causal effects of the ESSN on the distribution of the deprivation score, for eligible applicants. The probabilistic potential outcomes with (blue) and without (red) treatment, for eligible applicants, are plotted.

First, the estimated marginal distributions of each dimension are shown in Figure 6. These plots show the estimated probability mass for each level of the weighted dimension. Education, for example, has 0, 1 or 2 deprivations, corresponding to 0, 0.5 and 1. These distributions are probabilistic potential outcomes for the eligible applicants, both with (blue) and without (red) treatment.

In the education dimension, there was a notable reduction of the probability of being deprived on both indicators and an increase in the probability of being deprived in none. For

Figure 6: Marginal Distributions by Dimension



food security, we observe a large increase in the probability of being deprived in no indicators and decreases in both one or both indicators. For living standards, there is a similar increase in the probability of no deprivations, with the main decrease in probabilities from two, three or four simultaneous deprivations. For health and income resources, we see little change. In both of these dimensions, however, we already observe high probabilities of no deprivations and very low probabilities of both deprivations.[30]

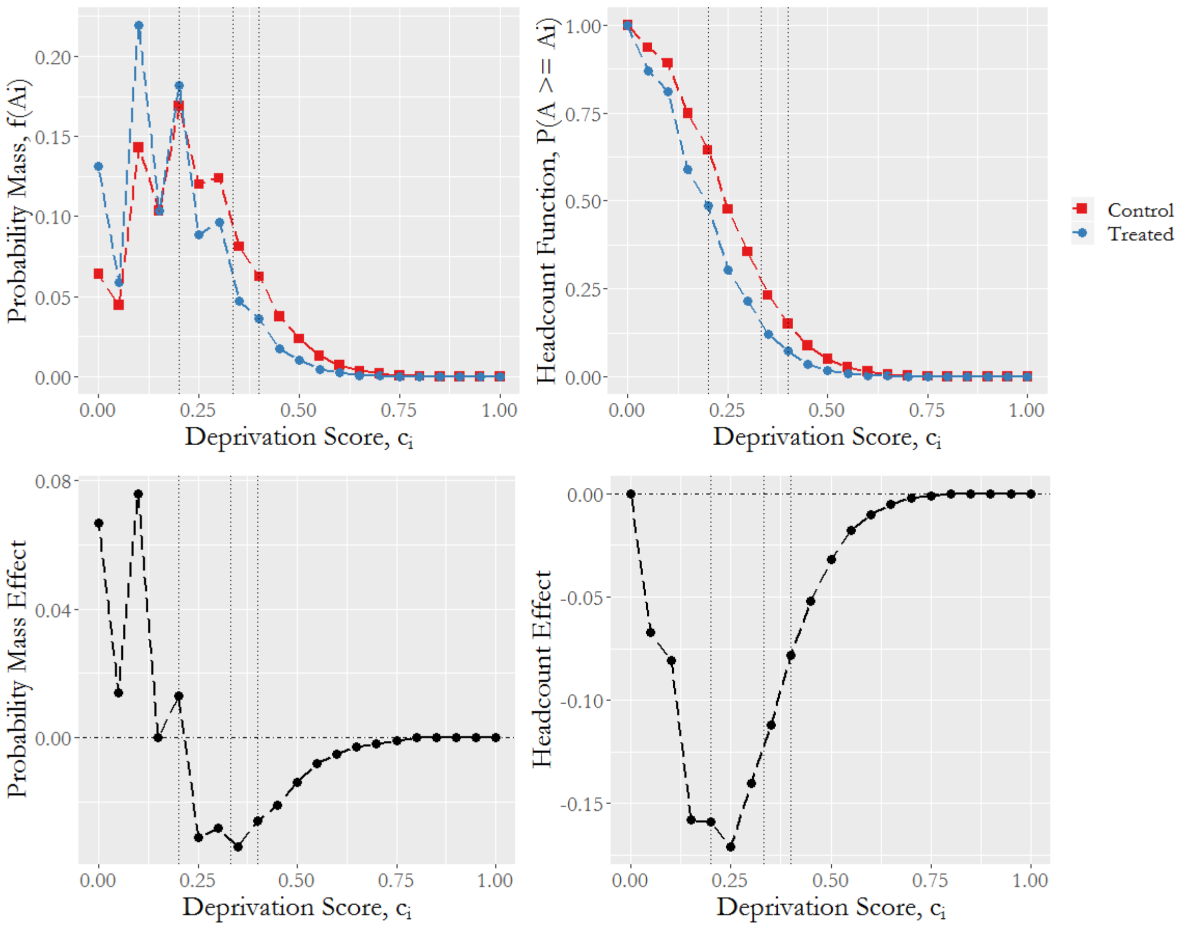
Once these marginal distributions have been estimated, the joint distribution can also be estimated. Figure 7 shows these estimated probabilistic potential outcomes. The left panel shows the probability mass function; the right shows the headcount function.

Results show that the ESSN has caused a downward shift in deprivation scores. There is a reduction in the probability of high deprivation scores and an increase in lower scores. Similar decreases in headcount ratios are found across different cutoff thresholds.

The bottom panel shows the differences between the probabilistic potential outcomes, i.e. the treatment effects. The left graph shows how the distribution of the deprivation score has changed, with the ESSN increasing the probability of having lower intensities of poverty

[30]Using these marginal plots, the expected potential outcomes of weighted deprivations in each dimension can be estimated. For education, this changes from 0.356 to 0.293, without and with the ESSN. For health this is 0.078 to 0.067, for food security 0.380 to 0.223, for income resources 0.118 to 0.149, and for living standards 0.226 to 0.153. These results are similar to those in Table 5.

Figure 7: Probability Mass and Headcount Functions and Effects



(below 0.25) and a lower probability of higher intensities of poverty (above 0.25). The headcount effect shows that at each threshold, the number of households classified as poor has reduced. This effect is greatest between the 20% and 30% threshold. At the 20% threshold, the number of poor households has reduced by 15.9 percentage points. At the 33.3% threshold the reduction is 14.0 percentage points, and at the 40% threshold it is 7.8 percentage points.

5 Discussion

Contributing to current debates on the poverty reduction potential of cash programmes in social and humanitarian contexts (Song and Imai, 2019; Loschmann et al., 2015; Seth and Tutor, 2021; Bastagli et al., 2019; Salti et al., 2021), we observed that the ESSN had a statistically significant impact on the overall multidimensional poverty of its beneficiaries. The finding confirms results of Song and Imai (2019) and Loschmann et al. (2015), who also found similar impacts of (unconditional) cash transfers on multidimensional poverty of beneficiaries in Kenya and Afghanistan. This supports emerging claims that these types of programmes, which in humanitarian contexts are still considered to be relatively new (World Bank, 2016), can be transformative for their beneficiaries when the intention is to achieve multiple outcomes simultaneously. However, this finding may depend on a number of conceptual and context-specific factors, including the purpose of the MPI used in the evaluation and the indicators included in the index, both of which we discuss below.

Moving beyond separate outcome evaluation exercises, which impact evaluations traditionally rely on (Tappis and Doocy, 2018), we also contribute to a growing body of academic literature that focuses the evaluative space on poverty intensities as measured through *simultaneous deprivations* across *multiple* outcome indicators (Vaz et al., 2019; Mitchell and Macció, 2021; Masset and García-Hombrados, 2021). We find that those with higher deprivation scores are within the group that the ESSN affects the most. In other words, the effects of the ESSN are greatest on the most deprived. Not only is this an encouraging finding when compared to other multidimensional poverty evaluations, such as offered by Seth and Tutor (2021), who found that the Pantawid Pamilyang Pilipino cash programme only marginally improved the consumption of the poorest families, it is also a finding that both the traditional indicator-specific evaluations and the newly emerging multidimensional poverty impact evaluations would have missed, and were only unearthed due to the distributional methods utilised in this paper, as discussed below.

5.1 Purpose of an MPI

The RMPI was designed as a 'population' MPI, with the purpose of measuring the poverty of refugees in Turkey. The index relies on the data available in the CVME survey and aims to provide a comprehensive understanding of multidimensional poverty that was informed not only by the needs of refugees but also by their rights and, where possible, their own voices (OPHI and WFP, 2022).[31] The indicators were selected to be applicable for the whole refugee population (in terms of age, gender, household composition, etc.), irrespective of their arrival time, length of stay (unlike the MPI developed in Loschmann et al. (2015)) or ESSN eligibility status. This contrasts with the approach taken by the other papers reviewed, which

[31] As captured through participatory thematic focus group discussions, for instance on shelter, which informed indicator selection on the living standards dimension (e.g. winter assets).

either designed 'programmatic' multidimensional indices based specifically on outcome indicators for programme beneficiaries, where the indicators and dimensions of the index may be selected based on their likely responsiveness to the programme and the modality of the assistance (see Song and Imai (2019); Seth and Tutor (2021); Vaz et al. (2019)), or relied on established 'general' indices of multidimensional poverty, such as the global MPI (Masset and García-Hombrados, 2021), which may not appropriately reflect the poverty of refugees.

Conceptual considerations regarding the purpose of the MPI are not trivial, and they need to be accounted for in any impact evaluation using a multidimensional perspective. Impacts of multi-purpose unconditional cash transfers on multidimensional poverty are more appropriate to be evaluated using a broader lens of poverty across the entire refugee population. Conditional cash transfers with closer theories of change might need to consider a comparative approach between a 'programmatic' and 'poverty' MPI.

5.2 Dimension and Indicator Level Effects

One advantage of measuring the impact of cash transfer programmes within a multidimensional framework is the ability to decompose effects into separate dimensions and indicators. This allows for a better understanding of the estimated effects, given the context and the choice of indicators.[32]

Our results show that the ESSN led to significant reductions in both deprivation scores and the incidence of multidimensional poverty. By decomposing effects to the dimension level, we observe that the largest effects were found on the food security dimension. Smaller, though still significant, effects were found on living standards and education, while no significant impact was identified for the health and income resources dimensions. These effects can, in turn, be explained by indicator level results, shown in Appendix A.7, where treatment effects are decomposed for each (uncensored) indicator of the RMPI.

Improvements in *education* are shown to be driven by improvements in school attendance,[33] but not by the highest education of the household.[34] For *health*, which has no significant

[32] Masset and García-Hombrados (2021), for example, found that the impact of the MVP on the global MPI was driven by two (of 10) indicators only, namely school attendance and sanitation.
[33] These findings are in line with those of Aygün et al. (2021).
[34] Due to data limitations of the CVME, the RMPI 'highest education' indicator only approximates the international standard indicator of 'years of schooling'. This indicator is used to proxy the learning outcomes (or educational attainments) of the household, and in the context of refugees, the right to education and continued learning for refugees of all ages. For reference, see the Refugee Right to Education, Article 22 of the 1951 Convention. Adult lifelong learning (vocational and technical training completed) can be accounted as equivalent

effect overall, we do observe a significant (albeit small) reduction in those not receiving treatment when sick, but no significant effect on the number of those reporting sick. For *food security*, deprivations in both the Coping Strategy Index and Dietary Diversity Score are significantly reduced. For *income resources*, while there is no significant effect overall, deprivation in household members in work has actually increased, whilst the effect on precarious work, while negative, is insignificant. Finally, for *living standards*, we observe negative effects for each of the four indicators. However, this effect is only significant for winter assets. The differences in these indicator level effects, both across and within each dimension, highlight potential reasons for the estimated overall effects on poverty.

One important consideration, is how *receptive* the chosen indicators are to cash transfers. Food security, for example, as measured by consumption and dietary diversity, is likely to be more receptive to cash assistance in market economies, such as in Turkey, even with relatively modest amounts of cash (Seidenfeld et al., 2014) – as in the accumulation of winter assets – than various other indicators, such as the highest education achieved of the household, where socio-economic, cultural and attitudinal barriers exist and are known to challenge even conditional cash transfers to have a comprehensive long-term impact (Millán et al., 2019).[35] This indicates, however, that for indicators that are less responsive to cash in the short and medium term in a given dimension, cash transfers alone might not be the most effective modality (particularly if unconditional) and may need to be complemented by other means to erase the aforementioned barriers. Further research is needed, first, to estimate medium- to long-term effects, and second, to identify whether isolated cash transfer programmes are as effective.

These indicator level effects may, however, also be particular to the Turkish context, where there is an ecosystem of assistance that exists for refugees.[36] For example, whilst effects on health are limited, this may be explained by the low baseline levels of deprivation in the health indicators (see Figure 4), which may be explained by the free access to public care which the Turkish government provides for both beneficiaries and non-beneficiaries.[37] In

to primary education achieved, in line with the with SDG Target 4 and the framework of the Global Compact on Refugees (UN, 2018). The indicator featured in a similar multidimensional index for internally displaced populations, with a similar rationale for its inclusion where, due to data limitations, only adults and adolescents were assessed (Admasu et al., 2021, 8).

[35]Even if adult lifelong learning equivalent to primary education in form of vocational and technical training completed were to be accounted for (Morrice, 2021).

[36]Which includes additional programmes on education, vocational training and other socio-economic support, but also sector capacity support to cover teacher salaries under the EU Facility for Refugees in Turkey (FRiT) (Sida, 2018).

[37]Under temporary protection for refugees. Schooling is also provided for free, limiting supply-side constraints

this context, it is encouraging that the ESSN forms part of a bigger ecosystem of assistance (as part of the FRiT and with additional assistance provided by the Government of Turkey) that complements the ESSN to fill the gaps. On the other hand, future research should further investigate the possible disincentive of the programme in picking up formal employment. The only indicator which is positive (indicating an increase in deprivation) is that which indicates that no household member has worked in the last 30 days. This may indicate that the ESSN has acted as a disincentive for household members to find formal work, given that households with any member enrolled in social security (a result of formal employment) are not eligible for the assistance (Cuevas et al., 2019, p.28).

In sum, given that the ESSN was designed as a humanitarian assistance fund to cover basic costs of living, related to the basic needs of refugees, we might not realistically expect to observe significant impacts in all 12 indicators of the multidimensional poverty index. However, whilst the RMPI is useful in allowing for the engagement in debates of the most effective modalities separately by indicator, it is important that this is within the overarching framework of a multidimensional index that ensures that the holistic view of poverty of refugees remains accounted for. In this context, the ESSN can be seen as a programme that successfully helped to reduce the incidence and intensity of poverty amongst its beneficiaries.

5.3 Simultaneous Deprivations

One of the main advantages of using a multidimensional measure of poverty based on the AF method is the ability to assess multiple indicators *simultaneously*. We find that the ESSN reduces the deprivation score of the more deprived *to a greater extent*, see Figure 5. This is a finding that indicator-specific evaluations would have missed, but, moreover, this finding would also have been missed by analysis which estimated *average* treatment effects on an MPI or on any of its statistics. The analysis of Table 4, which is common amongst other MPI impact evaluation papers (e.g. Song and Imai, 2019; Seth and Tutor, 2021), cannot provide this information. It is only through the econometric focus on the entire distribution of the deprivation score, which our analysis provides, that these effects can be revealed. It is, therefore, the combination of a multidimensional poverty index, specifically developed to capture simultaneous deprivations, with advanced distributional analysis which provides a more nuanced understanding of the effects of interventions on those individuals in the depths of poverty.

and potentially aiding the observed improvements in school attendance.

6 Concluding Remarks

In this paper, we have applied a refugee-specific MPI, the RMPI (OPHI and WFP, 2022), which was specifically developed to account for the needs, rights, and, where possible, own voices of refugees, to explore the impacts of the ESSN cash transfer programme in Turkey on the overall multidimensional poverty of its beneficiaries. We advocate for broader research into the impact of unconditional cash transfers on refugees, by first using a population MPI, such as the RMPI, to evaluate outcomes *jointly*, rather than separately, and then moving beyond estimating *average* treatment effects of MPIs to reflect the effects on the entire *distribution* of the deprivation score. This can uncover hitherto hidden findings which enrich current debates over the poverty-reduction potential of cash programmes in humanitarian contexts.

Using the RMPI and an inverse-probability-weighted regression adjustment method, we find that the ESSN significantly reduced the incidence and intensity of multidimensional poverty amongst its beneficiaries. We find that these reductions are largely driven by improvements in the dimensions of food security, living standards and education. Through conducting further distributional analysis, we find that reductions in the incidence of poverty are robust to the choice of cutoff threshold. Moreover, we find that the effects of the ESSN on deprivation scores are greatest amongst the most deprived.

Taken together, these findings suggest that multi-purpose cash transfer programmes in humanitarian contexts have an impact on the simultaneous deprivations refugees face. Using an MPI specifically tailored for refugees can help to provide this understanding of such joint impacts. Further research on other cash transfer programmes using the proposed methods in this paper would be useful to consolidate these findings more broadly, as would research on the sustainability of such impacts and the pathways out of social assistance.

References

Admasu, Y., Alkire, S., Ekhator-Mobayode, U.E., Kovesdi, F., Santamaria, J. and Scharlin-Pettee, S. (2021). 'A multi-country analysis of multidimensional poverty in contexts of forced displacement', *World Bank Policy Research Working Paper*, vol. 9826.

Alkire, S. and Foster, J. (2011). 'Counting and multidimensional poverty measurement', *Journal of Public Economics*, vol. 95(7-8), pp. 476–487, doi:10.1016/j.jpubeco.2010.11.006.

Alkire, S., Roche, J., Ballon, P., Foster, J., Santos, M. and Seth, S. (2015). *Multidimensional Poverty Measurement and Analysis*, Oxford University Press, USA, doi:10.1093/acprof:oso/9780199689491.001.0001.

Ark-Yıldırım, C. and Smyrl, M. (2021a). 'Cash transfer and humanitarian assistance', *Social Cash Transfer in Turkey*, pp. 89–113, doi:10.1007/978-3-030-70381-3_5.

Ark-Yıldırım, C. and Smyrl, M. (2021b). 'The consequences of ambiguity: designing and implementing the ESSN', *Social Cash Transfer in Turkey*, pp. 115–142, doi:10.1007/978-3-030-70381-3_6.

Atkinson, A.B. (1970). 'On the measurement of inequality', *Journal of Economic Theory*, vol. 2(3), pp. 244–263, doi:10.1016/0022-0531(70)90039-6.

Atkinson, A.B. (2003). 'Multidimensional deprivation: contrasting social welfare and counting approaches', *The Journal of Economic Inequality*, vol. 1(1), pp. 51–65, doi:10.1023/A:1023903525276.

Aygün, A.H., Kïrdar, M.G., Koyuncu, M. and Stoeffler, Q. (2021). 'Keeping refugee children in school and out of work: evidence from the world's largest humanitarian cash transfer program', *IZA Institute of Labor Economics Discussion Paper Series*, vol. 14513.

Bagstagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G. and Schmidt, T. (2016). 'Cash transfers: what does the evidence say? A rigorous review of programme impact and the role of design and implementation features', *Overseas Development Institute*.

Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G. and Schmidt, T. (2019). 'The impact of cash transfers: a review of the evidence from low-and middle-income countries', *Journal of Social Policy*, vol. 48(3), pp. 569–594, doi:10.1017/S0047279418000715.

Bedoya, G., Bittarello, L., Davis, J. and Mittag, N. (2017). *Distributional Impact Analysis: Toolkit and Illustrations of Impacts beyond the Average Treatment Effect*, The World Bank.

Blattman, C. and Niehaus, P. (2014). 'Show them the money: why giving cash helps alleviate poverty', *Foreign Affairs*, vol. 93(3), pp. 117–126.

Cetinoglu, T. and Yilmaz, V. (2021). 'A contextual policy analysis of a cash programme in a humanitarian setting: the case of the Emergency Social Safety Net in Turkey', *Disasters*, vol. 45(3), pp. 604–626, doi:10.1111/disa.12438.

Chernozhukov, V., Fernández-Val, I. and Melly, B. (2013). 'Inference on counterfactual distributions', *Econometrica*, vol. 81(6), pp. 2205–2268, doi:10.3982/ECTA10582.

Cornelissen, T., Dustmann, C., Raute, A. and Schönberg, U. (2016). 'From LATE to MTE: alternative methods for the evaluation of policy interventions', *Labour Economics*, vol. 41, pp. 47–60, doi:10.1016/j.labeco.2016.06.004.

Cuevas, P.F., Inan, O.K., Twose, A. and Celik, C. (2019). *Vulnerability and Protection of Refugees in Turkey: Findings from the Rollout of the Largest Humanitarian Cash Assistance Program in the World*, World Bank.

Doocy, S. and Tappis, H. (2017). 'Cash-based approaches in humanitarian emergencies: a systematic review', *Campbell Systematic Reviews*, vol. 13(1), pp. 1–200, doi:10.4073/csr.2017.17.

Farrington, J. and Slater, R. (2006). 'Introduction: cash transfers: panacea for poverty reduction or money down the drain?', *Development Policy Review*, vol. 24(5), pp. 499–511, doi:10.1111/j.1467-7679.2006.00344.x.

Firpo, S., Fortin, N.M. and Lemieux, T. (2009). 'Unconditional quantile regressions', *Econometrica*, vol. 77(3), pp. 953–973.

Firpo, S. and Pinto, C. (2016). 'Identification and estimation of distributional impacts of interventions using changes in inequality measures', *Journal of Applied Econometrics*, vol. 31(3), pp. 457–486, doi:10.1002/jae.2448.

Fiszbein, A. and Schady, N.R. (2009). *Conditional Cash Transfers: Reducing Present and Future Poverty*, World Bank Publications, doi:10.1596/978-0-8213-7352-1.

Galiani, S. and McEwan, P.J. (2013). 'The heterogeneous impact of conditional cash transfers', *Journal of Public Economics*, vol. 103, pp. 85–96, doi:10.1016/j.jpubeco.2013.04.004.

Gibárti, S. et al. (2021). 'Aiding Syrian refugees in Turkey: international approaches and domestic policies', *Security and Defence Quarterly*, vol. 33(1), pp. 57–72, doi:10.35467/sdq/132027.

Gini, C. (1912). 'Variabilità e mutabilità (variability and mutability)', *Cuppini, Bologna*, vol. 156.

Golan, J., Sicular, T. and Umapathi, N. (2017). 'Unconditional cash transfers in China: who benefits from the rural minimum living standard guarantee (dibao) program?', *World Development*, vol. 93, pp. 316–336, doi:10.1016/j.worlddev.2016.12.011.

Hagen-Zanker, J., Bastagli, F., Harman, L., Barca, V., Sturge, G. and Schmidt, T. (2016). 'Understanding the impact of cash transfers: the evidence', *Overseas Development Institute*.

Hagen-Zanker, J., Ulrichs, M. and Holmes, R. (2018). 'What are the effects of cash transfers for refugees in the context of protracted displacement? Findings from Jordan', *International Social Security Review*, vol. 71(2), pp. 57–77, doi:10.1111/issr.12166.

Handa, S., Natali, L., Seidenfeld, D., Tembo, G., Davis, B., Team, Z.C.T.E.S. et al. (2018). 'Can unconditional cash transfers raise long-term living standards? Evidence from Zambia', *Journal of Development Economics*, vol. 133, pp. 42–65, doi:10.1016/j.jdeveco.2018.01.008.

Haushofer, J. and Shapiro, J. (2016). 'The short-term impact of unconditional cash transfers to the poor: experimental evidence from Kenya', *The Quarterly Journal of Economics*, vol. 131(4), pp. 1973–2042, doi:10.1093/qje/qjw025.

Heckman, J.J. and Robb Jr, R. (1985). 'Alternative methods for evaluating the impact of interventions: an overview', *Journal of Econometrics*, vol. 30(1-2), pp. 239–267, doi:10.1016/0304-4076(85)90139-3.

Hjelm, L., Handa, S., de Hoop, J., Palermo, T., Zambia, C. and Teams, M.E. (2017). 'Poverty and perceived stress: evidence from two unconditional cash transfer programs in Zambia', *Social Science & Medicine*, vol. 177, pp. 110–117, doi:10.1016/j.socscimed.2017.01.023.

Koenker, R. and Bassett Jr, G. (1978). 'Regression quantiles', *Econometrica: Journal of the Econometric Society*, pp. 33–50, doi:10.2307/1913643.

Loschmann, C., Parsons, C.R. and Siegel, M. (2015). 'Does shelter assistance reduce poverty in Afghanistan?', *World Development*, vol. 74, pp. 305–322, doi:10.1016/j.worlddev.2015.05.022.

MacPherson, C. and Sterck, O. (2021). 'Empowering refugees through cash and agriculture: a regression discontinuity design', *Journal of Development Economics*, vol. 149, p. 102614, doi:10.1016/j.jdeveco.2020.102614.

Masset, E. and García-Hombrados, J. (2021). 'Sensitivity matters. Comparing the use of multiple indicators and of a multidimensional poverty index in the evaluation of a poverty eradication program', *World Development*, vol. 137, p. 105162, doi:10.1016/j.worlddev.2020.105162.

Maunder, N., Seyfert, K., Aran, M., Baykal, G., Marzi, M. and Smith, G. (2018). 'Evaluation of the DG-ECHO funded Emergency Social Safety Net (ESSN) in Turkey. November 2016–February 2018: Volume I: Final evaluation report', *World Food Programme*.

Millán, T.M., Barham, T., Macours, K., Maluccio, J.A. and Stampini, M. (2019). 'Long-term impacts of conditional cash transfers: review of the evidence', *The World Bank Research Observer*, vol. 34(1), pp. 119–159, doi:10.1093/wbro/lky005.

Mitchell, A. and Macció, J. (2021). 'Using multidimensional poverty measures in impact evaluation: emergency housing and the "declustering" of disadvantage', *Journal of Human Development and Capabilities*, pp. 1–24, doi:10.1080/19452829.2020.1847052.

Molyneux, M., Jones, W.N. and Samuels, F. (2016). 'Can cash transfer programmes have 'transformative' effects?', *The Journal of Development Studies*, vol. 52(8), pp. 1087–1098, doi:10.1080/00220388.2015.1134781.

Morrice, L. (2021). 'The promise of refugee lifelong education: a critical review of the field', *International Review of Education*, pp. 1–19, doi:10.1007/s11159-021-09927-5.

OPHI and WFP (2022). 'Meta-analysis of the impact and lessons learned for implementation of the Emergency Social Safety Net (ESSN) programme in Turkey (2016-–20). Part 2: Focus areas 2 and 3', *OPHI Special Publications and Reports*.

Özler, B., Çelik, Ç., Cunningham, S., Cuevas, P.F. and Parisotto, L. (2021). 'Children on the move: progressive redistribution of humanitarian cash transfers among refugees', *Journal of Development Economics*, vol. 153, p. 102733, doi:10.1016/j.jdeveco.2021.102733.

Rios-Avila, F. (2020). 'Recentered influence functions (RIFs) in Stata: RIF regression and RIF decomposition', *The Stata Journal*, vol. 20(1), pp. 51–94, doi:10.1177/1536867X20909690.

Rosenbaum, P.R. and Rubin, D.B. (1983). 'The central role of the propensity score in observational studies for causal effects', *Biometrika*, vol. 70(1), pp. 41–55, doi:https://doi.org/10.1093/biomet/70.1.41.

Rubin, D.B. (1974). 'Estimating causal effects of treatments in randomized and nonrandomized studies.', *Journal of Educational Psychology*, vol. 66(5), p. 688, doi:10.1037/h0037350.

Sadoulet, E., De Janvry, A. and Davis, B. (2001). 'Cash transfer programs with income multipliers: PROCAMPO in Mexico', *World Development*, vol. 29(6), pp. 1043–1056, doi:10.1016/S0305-750X(01)00018-3.

Salti, N., Chaaban, J., Moussa, W., Irani, A., Al Mokdad, R., Jamaluddine, Z. and Ghattas, H. (2021). 'The impact of cash transfers on Syrian refugees in Lebanon: evidence from a multidimensional regression discontinuity design', *Journal of Development Economics*, p. 102803, doi:10.1016/j.jdeveco.2021.102803.

Seidenfeld, D., Handa, S., Tembo, G., Michelo, S., Harland Scott, C. and Prencipe, L. (2014). 'The impact of an unconditional cash transfer on food security and nutrition: the Zambia child grant programme', *Institute of Development Studies Special Collection 1473*.

Seth, S. and Tutor, M.V. (2021). 'Evaluation of anti-poverty programs' impact on joint disadvantages: insights from the Philippine experience', *Review of Income and Wealth*, vol. 67(4), pp. 977–1004, doi:10.1111/roiw.12504.

Sida, L. (2018). *Strategic Mid-term Evaluation of the Facility for Refugees in Turkey, 2016–2019/2020*, Landell Mills Ltd.

Skoufias, E., Parker, S.W., Behrman, J.R. and Pessino, C. (2001). 'Conditional cash transfers and their impact on child work and schooling: evidence from the PROGRESA program in Mexico', *Economia*, vol. 2(1), pp. 45–96.

Slater, R. (2011). 'Cash transfers, social protection and poverty reduction', *International Journal of Social Welfare*, vol. 20(3), pp. 250–259, doi:10.1111/j.1468-2397.2011.00801.x.

Song, S. and Imai, K.S. (2019). 'Does the Hunger Safety Net Programme reduce multidimensional poverty? Evidence from Kenya', *Development Studies Research*, vol. 6(1), pp. 47–61, doi:10.1080/21665095.2019.1582347.

Tabor, S.R. (2002). 'Assisting the poor with cash: dsign and implementation of social transfer programs', *World Bank Social Protection Discussion Paper*, vol. 223, pp. 79–97.

Tappis, H. and Doocy, S. (2018). 'The effectiveness and value for money of cash-based humanitarian assistance: a systematic review', *Journal of Development Effectiveness*, vol. 10(1), pp. 121–144, doi:10.1080/19439342.2017.1363804.

Ulrichs, M., Hagen-Zanker, J. and Holmes, R. (2017). 'Cash transfers for refugees: an opportunity to bridge the gap between humanitarian assistance and social protection', *Overseas Development Institute*.

Ulrichs, M. and Sabates-Wheeler, R. (2018). 'Social protection and humanitarian response: what is the scope for integration?', *IDS Working Paper 516*.

UN (2018). 'Global compact for safe, orderly and regular migration (the Global Compact for Migration)', Accessed: https://refugeesmigrants.un.org/sites/default/files/180711_final_draft_0.pdf.

UNDP and OPHI (2019). *How to Build a National Multidimensional Poverty Index (MPI): Using the MPI to Inform the SDGs*, United Nations Development Programme.

UNHCR (2017). *A Guide to International Refugee Protection and Building State Asylum Systems, 2017, Handbook For Parliamentarians No. 27*, UN High Commissioner for Refugees.

Vaz, A., Malaeb, B. and Quinn, N. (2019). 'Evaluation of programs with multiple objectives: multidimensional methods and empirical application to PROGRESA in Mexico', *OPHI Research in Progress*, vol. 55a.

Wang, J.S.H., Malaeb, B., Ssewamala, F.M., Neilands, T.B. and Brooks-Gunn, J. (2021). 'A multifaceted intervention with savings incentives to reduce multidimensional child poverty: evidence from the Bridges study (2012–2018) in rural Uganda', *Social Indicators Research*, vol. 158(3), pp. 947–990, doi:10.1007/s11205-021-02712-9.

WFP (2019a). 'Comprehensive vulnerability monitoring exercise multidimensional poverty index', Accessed: https://reliefweb.int/sites/reliefweb.int/files/resources/WFP%20CVME3%20Multidimensional%20Poverty%20Index_May2019.pdf.

WFP (2019b). 'Reaching hidden populations with an innovative two-stage sampling method: a case study from the refugee population in Turkey', Accessed: https://docs.wfp.org/api/documents/WFP-0000104292/download/?iframe.

WFP (2020a). 'Comprehensive vulnerability monitoring exercise (CVME) round 5', Accessed: https://www.ecoi.net/en/file/local/2035471/CVME5_03072020.pdf.

WFP (2020b). 'Essential needs assessment. Guidance note', Accessed: https://docs.wfp.org/api/documents/WFP-0000074197/download/.

WFP (2020c). *The ESSN Shelter Security Focus Group Discussions Report*, World Food Programme.

World Bank (2016). *Cash Transfers in Humanitarian Contexts: Strategic Note*, World Bank Group Washington, DC.

CRediT authorship contribution statement

Robson, Matthew: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – Original Draft, Writing – Review & Editing, Visualisation, Project administration

Vollmer, Frank: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – Original Draft, Writing – Review & Editing, Visualisation, Project administration

Doğan, Basak Berçin: Resources, Writing – Review & Editing, Project administration

Grede, Nils: Resources, Funding acquisition, Writing – Review & Editing, Project administration

Declaration of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

A Appendix

A.1 Trial Measures

The identification of the RMPI followed normative and empirical considerations (OPHI and WFP, 2022, 1-40). Empirically, five trial measures were computed, based on the work of the World Food Programme (WFP) Turkey team and their proposed CVME MPI (WFP, 2019a). Using the same sample as the CVME MPI (wave 3, 1,301 refugee households, March–August 2018), the subsequent trial measures made use of statistical and normative considerations, including results of an association test, called redundancy, that analyses the association between indicator pairs developed by Alkire et al. (2015, 228–232) (see also UNDP and OPHI, 2019, 77). Results were prepared by a group of core researchers in OPHI and the WFP, and assessed in collaboration with contributors from OPHI, WFP Regional Bureau Cairo, Country Offices from Jordan, Chad and Turkey, as well as WFP Headquarters, to whom we express our gratitude. Table 6 presents a summary overview of the CVME MPI and trials 1–5, while trial 5 became the RMPI that we present in Figure 1 (and Appendix A.2). For a full description of the trial measure analysis, please see the Supplementary Information.

The RMPI moves towards a better approximation of the dimensions and indicators that are specific to the needs and rights of refugees, although we disclose some data concerns that other researchers and practitioners may wish to address in their application of the index. A suitable indicator on access to clean water, both a clear need and a right for refugees, could not be constructed, as objective information on the quality of the water source within an acceptable radius of 30-minute walk from home (roundtrip) was missing in the CVME dataset; if such indicator can be constructed, we advocate for its inclusion as a fifth indicator in the living standards dimension. Educational outcomes in terms of years of schooling could not be constructed for all household members; if such data were available, we advocate to replace the highest education achieved indicator, currently only assessing the household head and second responsible person of the household, to account for educational outcome information for all household members. The winter assets indicator was constructed based on results from the thematic FGD on shelter in Turkey, where respondents reported living in poorly insulated department buildings and suffering from the cold in the winter months (WFP, 2020c); when the index is applied in other contexts (e.g. in situations of south-south migration), the actual list of items could be adjusted. Conceptually, while in practice many overlaps exist between a needs-based and rights-based framework for refugees (e.g. on health and shelter), some compromises were needed to combine a needs and rights framework in one poverty index. For example, while needs are rather immediate, a rights-based focus will need data with longer recall periods (e.g. on educational attainments of children and adults beyond the last semester, or on income resources and sicknesses beyond the past 30 days).

While further indicator options were assessed considering the proposed safety dimension of the WFP's multidimensional deprivation index (MDDI), currently under development at WFP Headquarters (WFP, 2020b), and relevant literature, such as on the rights of refugees of freedom of movement, indicators on 'insecurity' were not included due to data limitations; they could be reconsidered with better data availability on safety and 'forced displacement'. Further, adding 'negative coping mechanisms' such as accumulating an unsustainable level of debt, an indicator included in the post-return refugee and displacement index of Loschmann et al. (2015), or not seeking healthcare when ill (to name just two examples) was also considered, but eventually not included at this stage as this would require a larger debate on the typologies of 'good' versus 'bad' coping strategies in refugee contexts. At this stage, note, though, that the food security indicator is already based on a Coping Strategy Index (CSI). Further debates on electricity and shelter indicators were held, also featured in Loschmann et al. (2015), but eventually no indicator was included because most refugees in Turkey are urban and live outside of camps, which explains why only 49 households (3.8%) in CVME3 reported living in tents (barns, makeshift shelter), most likely living in informal settings, while 98.46% used electricity for lighting. These extremes rendered the indicator options less suitable to measure poverty in the Turkish refugee context, while it may be more meaningful in contexts where refugees live in worse conditions.

Table 6: Trial Indicators and Weights

Dimension	Indicator	CVME MPI	T1	T2	T3	T4	T5
Education	Absence from school because children need to work and/or assist family	2.056	1/10	.	.	.	.
	Absence because family cannot afford to send children to school	2.729	.	.	.	.	.
	Absence from school for more than a semester	1.108	1/10	1/5	1/10	1/15	1/10
	Neither the household head nor the second responsible of the household (if applicable) has completed at least primary school (equivalent to six years of schooling)	.	.	.	1/10	.	1/10
Health	More than half of the household reported sick	2.785	1/10	1/10	1/10	1/15	1/10
	Any member not treated when sick	2.418	1/10	1/10	1/10	1/15	1/10
Food Security	Household with unacceptable food consumption	2.854	1/15	.	.	.	.
	Household with CSI>18	1.019	1/15	1/10	1/10	1/10	1/10
	DDS <6	2.153	1/15	1/10	1/10	1/10	1/10
Income Resources	No income source other than ESSN/other assistance or no income at all	2.303	1/20	.	.	.	.
	Begged	3.223	1/20	1/15	1/15	1/15	.
	Accepted high risk, illegal, socially degrading or exploitative temporary jobs	3.103	1/20	1/15	1/15	1/15	.
	No household member worked within last 30 days	1.979	1/20	1/15	1/15	1/15	1/10
	Begged or engaged in illegal or high-risk work	.	.	.	.	.	1/10
Living Standards	Crowding above 3	1.127	1/30	1/25	1/30	1/30	1/20
	No kitchen in the house	3.129	1/30	.	.	.	.
	No toilet in the house	2.075	1/30	1/25	1/30	1/30	1/20
	Bad quality apartment	0.418	1/30	1/25	1/30	1/30	.
	No sufficient winter clothes	0.753	1/30	1/25	1/30	1/30	.
	Insufficient access to any of the items below water, hygiene items, cooking fuel for cooking	1.496	1/30	1/25	1/30	.	
	Household does not own more than one of: television, telephone, refrigerator, motorbike, computer, stove, and does not own a car or truck	.	.	.	1/30	1/30	.
	Does not own more than one of the following winter assets: sufficient winter clothes, sufficient blankets, heating stove and central heating	.	.	.	.	.	1/20
	Household members do not have sufficient soap and hygiene items	.	.	.	.	.	1/20

Note: Frequencies (Uncensored Headcount Ratios) in % of CVME Indicators, from top to bottom of the table: 15.2, 2.9, 36.1, 5.7, 10.6, 11.2, 29.7, 16.4, 15.7, 7.5, 6.3, 22.7, 19.1, 8.1, 16.4, 60.3, 31.2, 23.4

A.2 RMPI Table Presentation

Table 7 presents the information in a tabular form, including more information for the deprivation criteria.

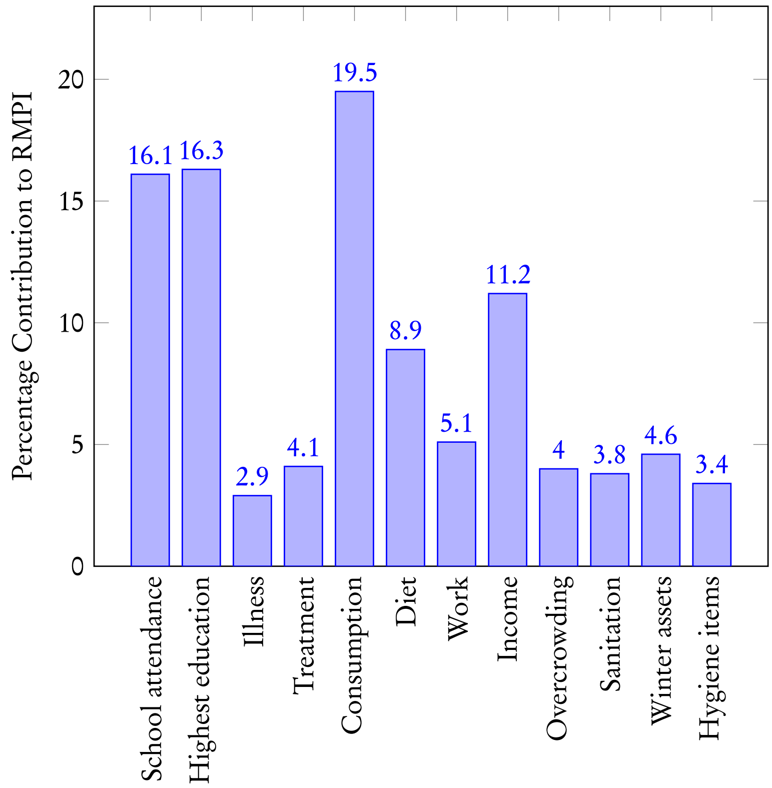
Table 7: RMPI: Dimensions, Indicators, Deprivation Thresholds and Weights

Dimension	Indicator	A household is deprived if...	Weight
Education	School Attendance	A household is deprived if children (girls and boys aged 6–17) are absent from school more than a semester.	$\frac{1}{10}$
	Highest Education	A household is deprived if neither the household head nor the second responsible person of the household (if applicable) has completed at least primary school.	$\frac{1}{10}$
Health	Illness	A household is deprived if more than half of the household members reported sick in the past 30 days. Sickness includes diarrhoea, fever/chills, or cough (i.e. not a simple cold).	$\frac{1}{10}$
	Treatment	A household is deprived if any member is not treated when sick.	$\frac{1}{10}$
Food Security	Consumption	A household is deprived if the household has a Coping Strategy Index (CSI) Score >18 (equating to using every consumption coping strategy at least three times per week).	$\frac{1}{10}$
	Diet	A household is deprived if the Dietary Diversity Score DDS is <6	$\frac{1}{10}$
Income Resources	Precarious work	A household is deprived if members of the household begged or engaged in illegal or high-risk work.	$\frac{1}{10}$
	No income	A household is deprived if no household member worked within the last 30 days.	$\frac{1}{10}$
Living Standards	Overcrowding	A household is deprived if there are more than 3 persons per room.	$\frac{1}{20}$
	Sanitation	A household is deprived if it does not have a toilet in the house.	$\frac{1}{20}$
	Winter assets	A household is deprived if it does not own more than one of the following winter assets: sufficient winter clothes, sufficient blankets, a heating stove, and central heating.	$\frac{1}{20}$
	Hygiene	A household is deprived if its members do not have sufficient soap and hygiene items.	$\frac{1}{20}$

A.3 Percentage Contributions

Figure 8 shows the composition of poverty, or how much each indicator is contributing to the RMPI (weighted variable * censored variable/divided by RMPI). The biggest contributor is the consumption indicator, followed by the two education indicators, suggesting that food security and education would be important dimensions for reducing poverty. A lack of income in the past month is also a relatively high contributor, suggesting that the ability to earn a living remains an issue in this refugee context.

Figure 8: Percentage Contribution by Indicator



Comparing the main contributors to poverty by ESSN group, we see similarities and differences across these groups. Two observations are worth highlighting. First, for all three groups, the two education indicators and the consumption indicator are the main drivers of poverty. However, for the ineligible applicants it is noteworthy that the contribution of the two education indicators is relatively less pronounced than for the other groups, while consumption contributes with 20.5% the most to poverty across all groups. Secondly, for the non-applicants we find that in addition to the education indicators and the consumption indicator, the second indicator of the food security dimension contributes quite substantially to poverty as well, namely when the dietary diversity score of the household is below 6, with 14.5%.

Figure 9: Eligible Applicants

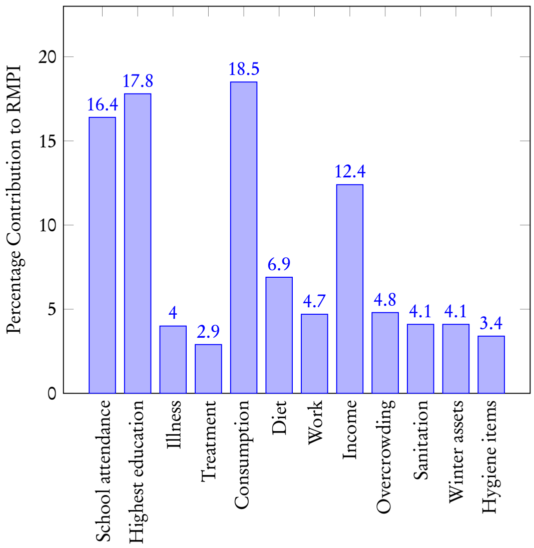


Figure 10: Ineligible Applicants

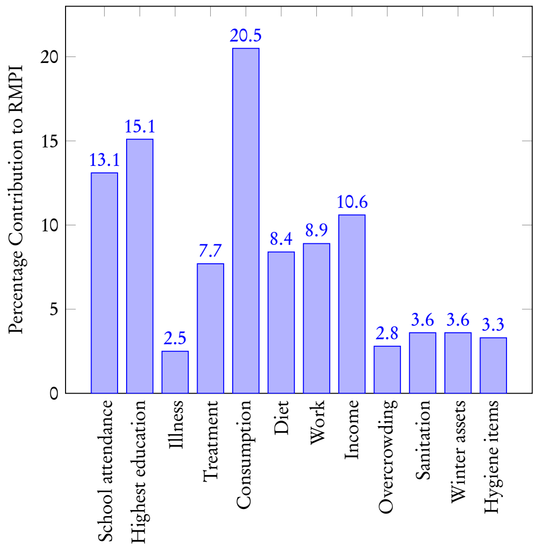
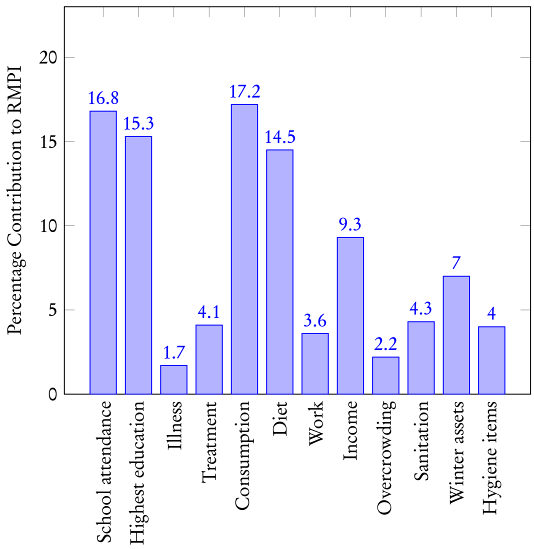


Figure 11: Non-Applicants



A.4 Geographical Variation in RMPI, Incidence and Intensity

While the CVME data used for our main analysis is representative at the national level, the following results are presented at the regional and provincial level to highlight geographical variation in poverty.[38] This is important as the focus of this study is on the ESSN supporting refugees living outside of camps among the host population, where the majority of refugees in Turkey live (Cuevas et al., 2019).

Figures 12, 13 and 14 respectively show the RMPI, incidence and intensity of poverty by province in Turkey. Darker colours show higher RMPI, incidence or intensity rates. Table 8 additionally shows RMPI, headcount ratios and average intensities, alongside the population share and sample size, by regions.

At the province level, Nevsehir and Kirsehir (located in the Anatolian region) are the poorest according to the RMPI, with values of 0.298 and 0.268, above the national level of 0.143. The poorest four provinces by headcount ratio (incidence) were found to be Nevsehir, Kirsehir, Yalova and Sanliurfa, with incidences of poverty of 83%, 78.4%, 75.8% and 73.2% respectively (the first three located in the Anatolian region, Sanliurfa in the South-East region).[39] They are, however, not the provinces with the highest intensities of poverty, which are Corum, Nigde, Bolu and Tekirdag (47.7%, 41.4%, 40% and 40%), all located in the Anatolian region.[40] Notwithstanding, Nevsehir and Kirsehir also depict intensities above 30% that classifies them among the provinces with the highest intensities of poverty. This explains why Nevsehir and Kirsehir were found to be the poorest by the adjusted headcount ratio, the RMPI.

The results fit with the regional level, where the Anatolian region, at an overall sample size of 46.4%, showed the highest RMPI of all of the five regions, at 0.155, as well as the highest intensity of poverty, with 31.2%. Istanbul has the next highest RMPI (0.153), followed by South-East (0.146), Mediterranean (0.133) and finally Aegean (0.109). However, note that generally, the confidence intervals between the regions overlap.

[38] After data cleaning, the sample size at provincial level reduced to 3,880. Across 81 provinces of Turkey, poverty was calculated for 59. The sample varied from 23 in Tekirdag to 350 in Ankara. 28 provinces had a sample size of 25, 15 between 50 and 75, and 12 above 100. Recall that 25 households were the minimum data collection point by GPS location, while some provinces were sampled more than once (WFP, 2019b, 2020a).
[39] The first two with a sample size of 25, Yalova with 75 and Sanliurfa with 175.
[40] At sample sizes of 25, 75, 50 and 23 respectively.

Figure 12: Refugee Multidimensional Poverty Index (RMPI)

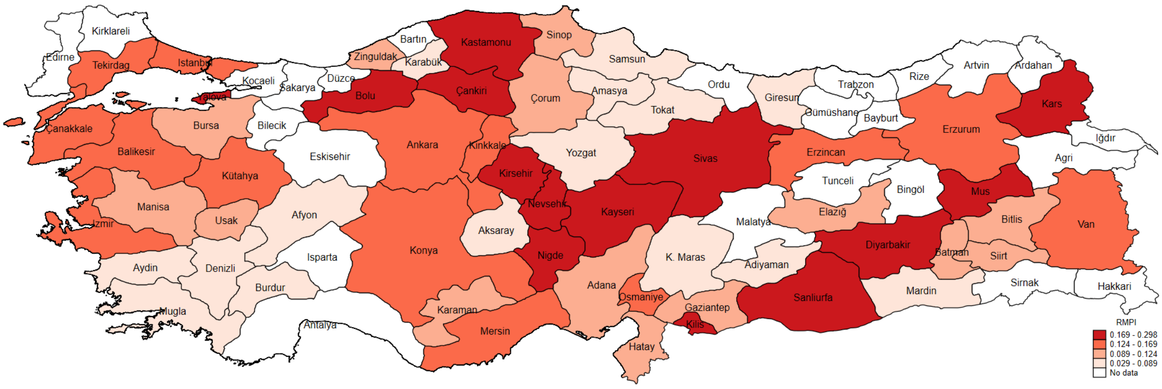


Figure 13: Incidence of Multidimensional Poverty among Refugees in Turkey

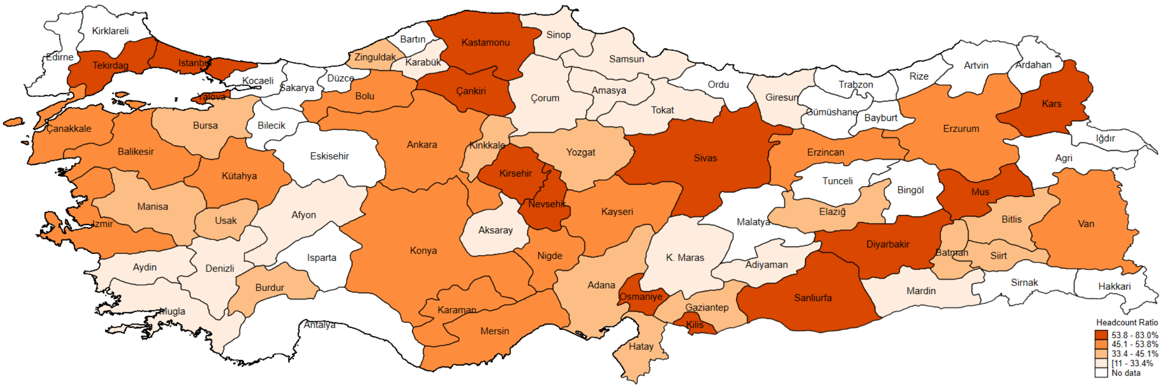


Figure 14: Average Intensity of Multidimensional Poverty among Refugees in Turkey

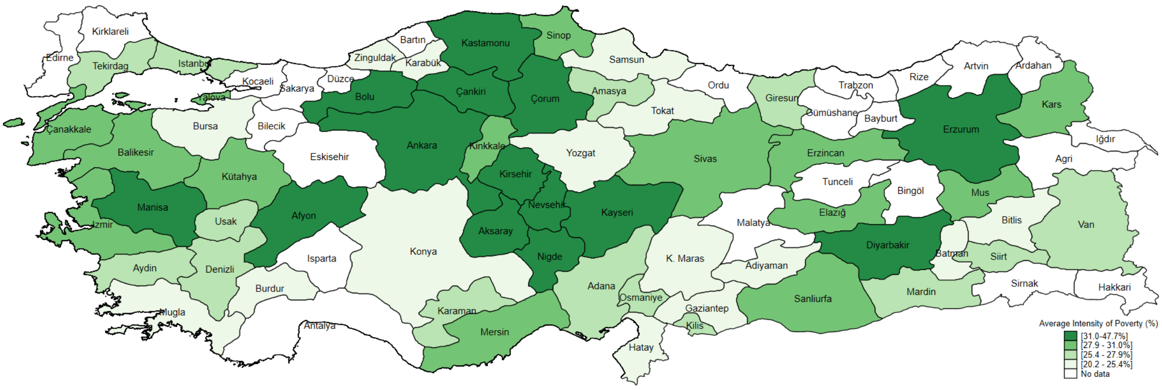


Table 8: RMPI, Headcount Ratio and Average Intensity of Poverty, by Region

Region	Population Share (Percentage)	Sample size (Frequency)	RMPI (95% CI)	H (95% CI)	A (95% CI)
Istanbul	5.5%	225	0.153	57.7%	26.5%
			(0.120, 0.185)	(46.5%, 68.9%)	(23.5%, 29.5%)
Aegean	20.1%	826	0.109	41.1%	26.5%
			(0.084, 0.133)	(32.2%, 50.0%)	(24.5%, 28.5%)
Mediterranean	13.4%	550	0.133	48.3%	27.4%
			(0.102, 0.163)	(38.3%, 58.4%)	(24.8%, 30.1%)
Anatolia	46.4%	1905	0.155	49.5%	31.2%
			(0.126, 0.183)	(43.0%, 56.0%)	(27.6%, 34.8%)
South-East	14.6%	600	0.146	52.4%	27.9%
			(0.111, 0.181)	(39.1%, 65.7%)	(25.0%, 30.7%)

A.5 Sensitivity

We conduct sensitivity analysis on RMPI deprivation scores to identify whether the treatment effect is robust to multiple specifications. The base case is shown in Model (1), the Average Treatment on the Treated (ATET) with a reduction in the deprivation score of -0.0577. In (2), the Average Treatment Effect (ATE) shows the average effect across the whole population, rather than just the treatment group; this effect is still significant but reduced to -0.0323. Model (3) reveals the result when only the applicants are used in the analysis; here the control group are those who applied but are non-beneficiaries, not including the non-applicants or pending households, as is in the main analysis. These results show there is still a large and significant effect, with a magnitude comparable to the base case. Model (4) shows the results when no sample weights are used; here again the result is in the same direction and highly significant, but the size of the effect is reduced to -0.0288. Finally, results are shown when no controls are used within the inverse-probability-weighted regression adjustment (i.e. this is a raw difference in averages between the control and treated group). This difference is not significantly different from zero.

Table 9: Sensitivity Analysis: Treatment Effects, Deprivation Score

	(1) **ATET** Coef./S.E.	(2) **ATE** Coef./S.E.	(3) **Applicants Only** Coef./S.E.	(4) **No S. Weights** Coef./S.E.	(5) **Unweighted** Coef./S.E.
Treatment Effect	-0.0577***	-0.0323***	-0.0437***	-0.0288***	-0.0070
	(0.0139)	(0.0110)	(0.0134)	(0.0069)	(0.0158)
PO: No Treatment	0.2328***	0.2061***	0.2188***	0.2063***	0.1833***
	(0.0147)	(0.0111)	(0.0138)	(0.0065)	(0.0127)
N	4042	4042	3504	4042	4106

Note: Inverse-probability-weighted regression-adjustments are used to estimate coefficients, the treatment effects and the expected potential outcome without treatment. The outcome across all models is the deprivation score. Model (1) shows the average treatment on the treated and (2) shows the average treatment effects. Model (3) excludes non-applicants and pending households, whilst (4) has no sample weights and (5) does not use inverse-probability weights. p-values: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

A.6 Inverse Probability Weighting: Balance and Overlap

A.6.1 Balance

Table 10 shows the balance between observable characteristics for the treatment (eligible applicants) and control (ineligible applicants, non-applicants and pending) groups, with and without the inverse-probability weights. The variables shown are those used to estimate propensity weights, for region, arrival time, characteristics of the household head, ESSN eligibility criteria and household composition. Before weighting, there are clear and significant imbalances between certain characteristics, but after weighting these differences are small and insignificant across all variables.

There are several notable differences between groups before the weights are applied. The largest differences relate to the eligibility criteria. In the treatment group, there are fewer households with lower dependency ratios and more with higher, and there are more single parents and households with higher numbers of children and elderly members. These imbalances are expected, as it is on these criteria that households are selected as being eligible for the ESSN. We also observe that there are fewer households located in Istanbul in the treated group. In the treatment group, the household heads are more often female, less likely to be younger (<30) or older (50+) and more likely to be unskilled. In terms of arrival time, the treatment group has more households who arrived 3–6 years ago and fewer households who have recently arrived (<12 months).

After matching, the differences between groups are insignificant. Hence, on these *observable* characteristics, these groups appear balanced. As in all propensity weighting methods, we cannot control for unobservable differences between these groups, and so our results are contingent upon the assumption of no relevant unobservable differences.

Table 10: Balance: Unweighted and Weighted

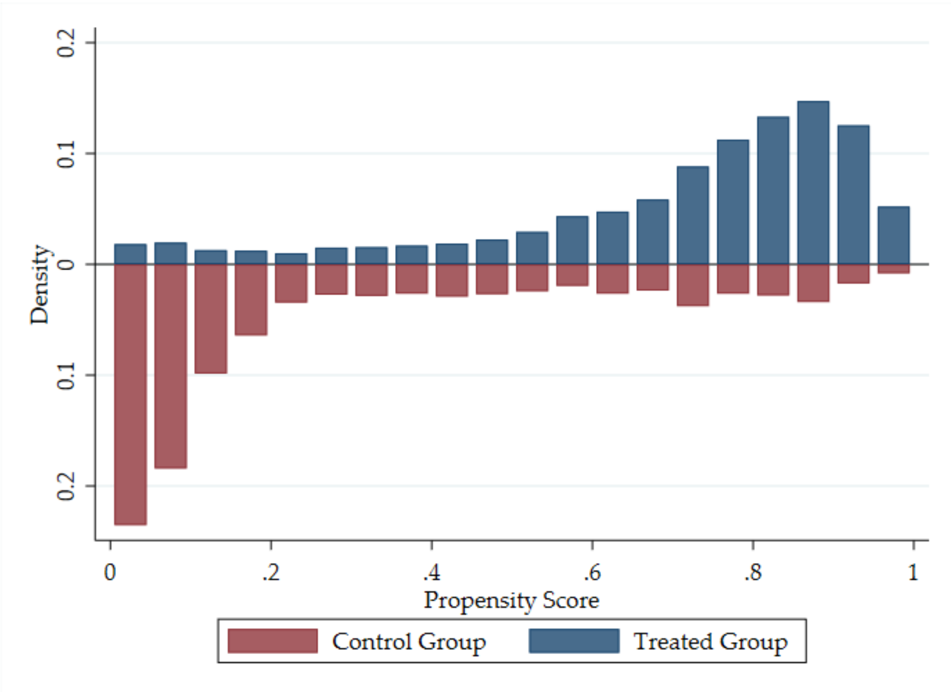
	(1) Unweighted		(2) Weighted	
	Coef.	Std. err.	Coef.	Std. err.
Region				
Istanbul	-0.027*	(0.015)	0.030	(0.028)
Aegean	-0.012	(0.021)	-0.021	(0.037)
Mediterranean	0.021	(0.033)	0.001	(0.037)
Anatolia	-0.037	(0.046)	-0.067	(0.055)
South-east	0.054	(0.069)	0.056	(0.077)
Arrival Time				
<12 months	-0.047*	(0.029)	-0.001	(0.039)
1-3 years	-0.068	(0.054)	-0.015	(0.045)
3-6 years	0.144**	(0.057)	-0.032	(0.067)
Before conflict	-0.029	(0.048)	0.048	(0.063)
Household Head				
Female Headed	0.100**	(0.039)	-0.032	(0.043)
Age (years) < 30	-0.173***	(0.053)	0.035	(0.065)
Age (years) 30-39	0.178***	(0.047)	-0.040	(0.057)
Age (years) 40-49	0.088**	(0.043)	-0.008	(0.060)
Age (years) 50+	-0.093*	(0.056)	0.012	(0.050)
Afghan	-0.020	(0.021)	-0.004	(0.015)
Iraqi	0.009	(0.015)	0.003	(0.019)
Syrian	0.010	(0.026)	0.017	(0.028)
Highly Skilled	-0.027	(0.024)	-0.015	(0.014)
Skilled	-0.039	(0.047)	-0.001	(0.035)
Semi-Skilled	-0.076	(0.048)	0.103	(0.072)
Unskilled	0.148***	(0.056)	-0.072	(0.064)
Eligibility Criteria				
Dep. Ratio < 1.5	-0.486***	(0.057)	0.069	(0.065)
Dep. Ratio 1.5 - 2.5	0.308***	(0.057)	-0.051	(0.059)
Dep. Ratio > 1.5	0.178***	(0.041)	-0.017	(0.058)
Single Female	0.031	(0.026)	0.003	(0.015)
Single Parent	0.042*	(0.024)	-0.024	(0.034)
Disabled Members	-0.007	(0.046)	0.021	(0.036)
Num. Children >= 4	0.391***	(0.049)	-0.052	(0.063)
Number of Children				
0	-0.066*	(0.038)	0.005	(0.038)
1-2	-0.483***	(0.046)	0.058	(0.066)
3-4	0.328***	(0.047)	-0.041	(0.060)
5+	0.221***	(0.041)	-0.022	(0.058)
Number of Adults				
1	0.026	(0.042)	-0.044	(0.039)
2	0.053	(0.058)	-0.001	(0.063)
3	-0.044	(0.047)	0.014	(0.045)
4+	-0.036	(0.030)	0.031	(0.036)
Number of Elderly				
0	-0.001	(0.034)	0.022	(0.033)
1	-0.021	(0.033)	-0.026	(0.032)
2+	0.021**	(0.009)	0.004	(0.007)
N				

Note: Each coefficient shows the difference in observable characteristics between the treatment (eligible applicants) and control (ineligible applicants and non-applicants) groups. Standard errors are robust, estimated from separate univariate regressions. The unweighted columns show the balance before inverse-probability weights are applied; the weighted shows the balance after. In both cases, sample weights are used. p-values: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

A.6.2 Overlap

The second assumption necessary for propensity matching is common support, or the overlap assumption. This requires the estimated propensity scores to overlap between the control and treated groups. This ensure that each observation in the treatment group has a 'nearby' observation in the control group. Figure 15 shows the distribution of the estimated propensity weights for both the control and treatment groups. As expected, there are generally lower propensity weights for the control group compared with the treated group. However, there is full support for both the control and treated groups across the whole distribution.

Figure 15: Overlap: Propensity Scores for Untreated and Treated



A.7 RMPI Indicator Treatment Effects

To further decompose treatment effects, we can estimate the impact of the ESSN on each (uncensored) indicator. In the main analysis, Table 5 shows significant reductions in deprivation in the dimensions of Food Security, Living Standards and Education, while no significant effects were found for Health or Income Resources. Table 11 delves deeper into the individual indicators.

Table 11: RMPI Indicators: Uncensored

	Education		Health		Food Security	
	SCH Coef./S.E.	**EDU** Coef./S.E.	**ILL** Coef./S.E.	**TRE** Coef./S.E.	**CSI** Coef./S.E.	**DDS** Coef./S.E.
Treatment Effect	-0.1337***	0.0077	0.0297	-0.0565*	-0.1469***	-0.1640***
	(0.0422)	(0.0525)	(0.0208)	(0.0317)	(0.0474)	(0.0407)
PO: No Treatment	0.4754***	0.2561***	0.0355***	0.1088***	0.4777***	0.2729***
	(0.0412)	(0.0379)	(0.0135)	(0.0306)	(0.0379)	(0.0445)
N	4042	4042	4042	4042	4042	4042

	Income Resources		Living Standards			
	PRC Coef./S.E.	**WRK** Coef./S.E.	**OVC** Coef./S.E.	**SAN** Coef./S.E.	**WAS** Coef./S.E.	**HYG** Coef./S.E.
Treatment Effect	-0.0201	0.0630**	-0.0578	-0.0668	-0.1292***	-0.0592
	(0.0277)	(0.0298)	(0.0440)	(0.0456)	(0.0463)	(0.0374)
PO: No Treatment	0.0909***	0.1412***	0.2455***	0.2073***	0.3079***	0.1786***
	(0.0216)	(0.0328)	(0.0427)	(0.0350)	(0.0415)	(0.0358)
N	4042	4042	4042	4042	4042	4042

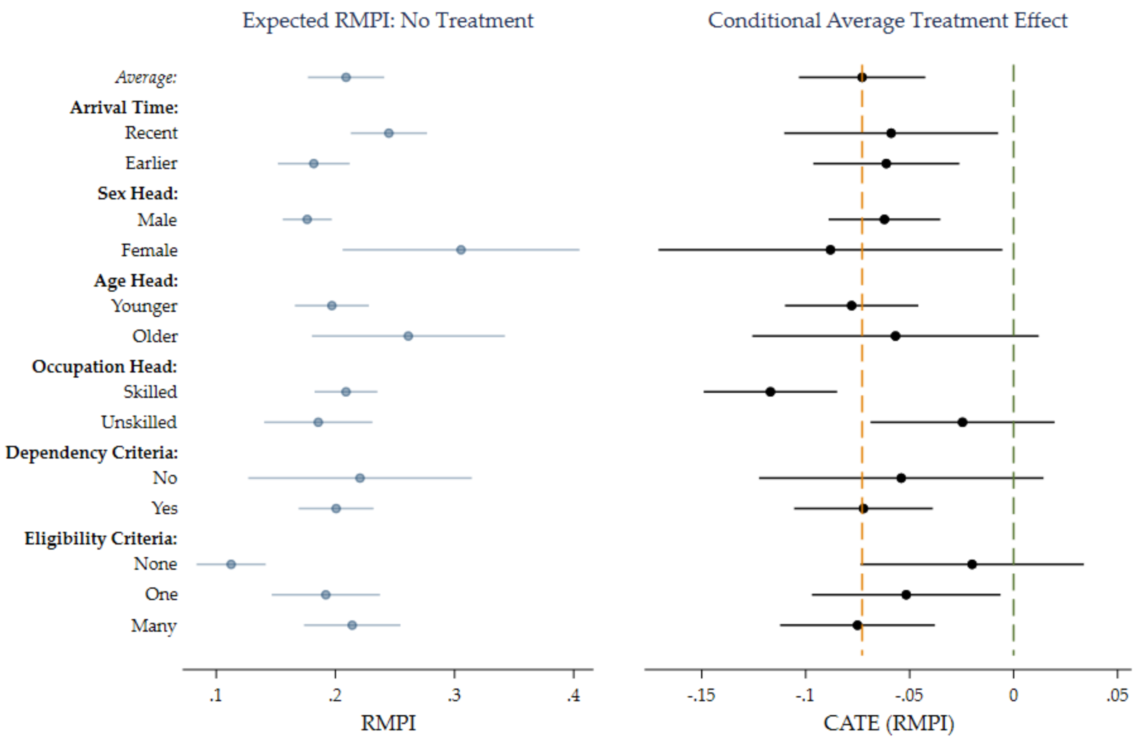
Note: Inverse-probability-weighted regression-adjustments are used to estimate coefficients, the Average Treatment on the Treated and expected potential outcome without treatment. Models show deprivation in each of the 12 indicators: 1 if deprived, 0 if not. p-values: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The effect on education is shown to be driven by reductions in deprivation in school attendance, but not by improvements in the education of the household head. For health, which has no significant effect overall, we observe a significant reduction in those not receiving treatment when sick, but no significant effect on the number of those reporting sick. For food security, deprivations in both the Coping Strategy Index and Dietary Diversity Score are significantly reduced. For income resources, we find that while there is no significant effect overall, deprivation in no income has increased, showing that significantly fewer households have any members who have worked in the last 30 days. The effect on precarious work, while negative, is insignificant. For living standards, we observe negative effects for each of the four indicators. However, this effect is only significant for winter assets.

A.8 Conditional Average Treatment Effects

Whilst the main analysis focuses on average effects, heterogeneity in treatment effects can be estimated by running IPWRA within subgroups. Figure 16 shows estimated Conditional Average Treatment Effects (CATEs) by arrival time; the sex, age and occupation of the household head; dependency criteria; and the number of simultaneous eligibility criteria.[41] These CATEs are shown for RMPI, alongside the expected RMPI (potential outcome) with no treatment. Note that all effects are 'on the treated'.

Figure 16: RMPI Conditional Average Treatment Effects: Forest Plot



Note: Estimates obtained with IPWRA, within each subgroup. Point estimates and 95% confidence intervals are shown. Predictions (on the treated) of expected RMPI, without treatment, are in the left panel, while CATEs are shown in the right panel. Subgroups are defined as follows. Arrival time: Recent <3 years, Earlier >3 years. Sex: Male or Female household head. Age: Younger ≤ 40, Older >40. Occupation: Skilled are highly skilled, skilled or semi-skilled, Unskilled are unskilled workers. Dependency Criteria: No <1.5, Yes ≥ 1.5. Eligibility Criteria: None $=0$, One $=1$, Many ≥ 2.

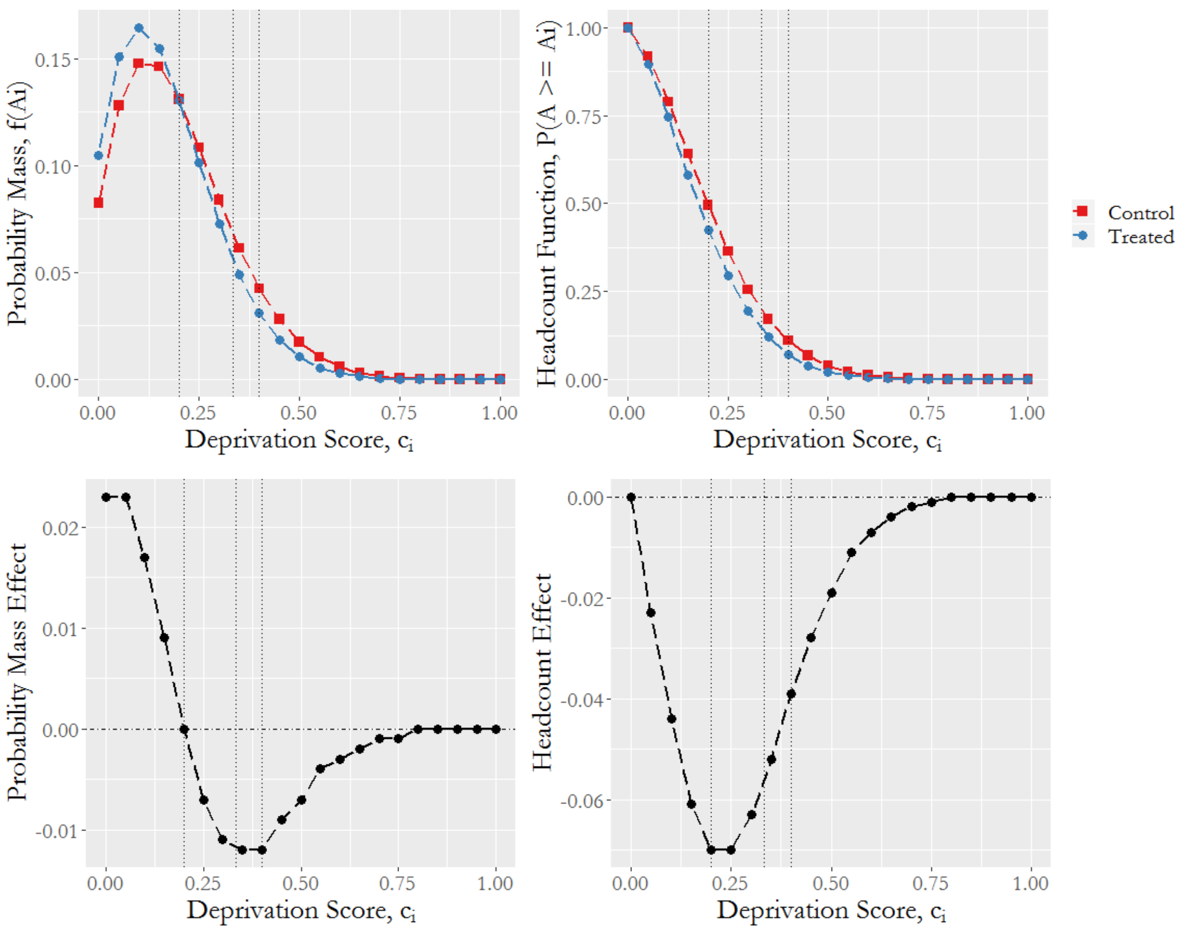
[41] Whilst other conditioning variables are available, those shown here are those for which balance within each subgroup was achieved. This balance was not met for subgroups defined by household size, number of children, region, disabled member, single parents, single female and elderly headed households.

These results show several differences by subgroups. More recent arrivals face higher levels of RMPI (without treatment) than earlier arrivals, but the effect of the ESSN is similar for both groups. Female-headed households have much higher RMPI than male-headed households, and the ESSN appears to have a greater effect on them (though not significantly so). Households with younger household heads have lower levels of poverty and the effects of the ESSN are greater, compared with older heads (not significant). Skilled and unskilled household heads face similar levels of RMPI (without treatment), but the ESSN has a significantly larger effect on reducing poverty for skilled households, and has lower (and insignificant) effects on unskilled households. Households with higher dependency ratios have similar expected RMPIs, and while the effects of the ESSN are higher, the difference is not significant. Finally, households who meet more eligibility criteria simultaneously have significantly higher RMPI and the ESSN has larger effects on reducing their poverty.

A.9 Distributional Estimation: Deprivation Score

In the main analysis, in Section 4.2.3, we estimate probabilistic potential outcomes for deprivation scores *indirectly*, by estimating marginal distributions of each dimension separately before estimating the joint distribution. An alternative is to directly estimate deprivation scores using the beta-binomial distribution with $T = 20$. This appendix shows and discusses the results from this method. Figure 17 shows the causal effect of the ESSN on the distribution of the deprivation score, for eligible applicants, estimated via the direct method. The probabilistic potential outcomes with (blue) and without (red) treatment, for eligible applicants, are plotted. The top-left panel shows the probability mass function, the top-right shows the headcount function. Results show that the ESSN has caused a downwards shift in deprivation scores, with an expected deprivation score of 0.2659 without the ESSN down to 0.2393 with the ESSN.

Figure 17: Probability Mass and Headcount Functions and Effects: Deprivation Score Estimation



The bottom-left graph shows how the distribution of the deprivation score has changed, with the ESSN increasing the probability of having lower intensities of poverty (below 0.25) and lowering the probability of higher intensities of poverty (above 0.25). The headcount effect, in the bottom-left, shows that at each threshold, the number of households classified as poor has reduced. This effect is greatest between the 20% and 30% threshold. At the 20% threshold the number of poor households has reduced from 69.7% to 63.7%. At the 33.3% threshold, the reduction is from 43.8% to 36.8%, and at the 40% threshold this is 22.7% to 17.2%.

The differences between the direct and indirect method is clear from the smoothness of the curves. While the smoothness of the underlying beta-binomial leads to a smooth curve for the direct estimation method, by estimating deprivation scores indirectly, via marginal distributions on dimensions, a closer fit to the true distribution of the RMPI is revealed.